Minimal Polynomials and Jordan Canonical Forms

Yang, Jia-Jin

In this article, $V$ denotes a finite-dimensional vector space over a field $F$ and $T$ is a linear operator on $V$.

1 Introduction

Let $T$ be a linear operator on $V$. If there exists a basis $\beta$ such that the matrix representation $[T]_\beta$ is as simple as possible, sometimes the problem can be simplified. However, not every linear operator on $V$ is diagonalizable, even if its characteristic polynomial splits. The purpose of this article is to introduce the Jordan canonical form (or simply Jordan form) of a linear operator. This kind of canonical form is “almost” a diagonal matrix (possibly some 1’s at $(i, i+1)$-entry). Fortunately, every linear operator on a $\mathbb{C}$-vector space has a Jordan form. Because it is “almost” a diagonal matrix, its matrix power is not hard to calculate. It will also be shown that two matrices are similar if and only if they have the same Jordan form. This is the reason why Jordan form is called a “canonical” form.

2 Minimal Polynomials

Lemma 2.1. There exists a unique monic polynomial $m \in P(F)$ of smallest degree such that $m(T) = 0$. This unique polynomial $m$ is called the minimal polynomial of $T$.

Proof. By Cayley-Hamilton theorem there exists a nonzero polynomial $f \in P(F)$ satisfying $f(T) = 0$. So there exists a nonzero polynomial $m \in P(F)$ of smallest degree such that $m(T)=0$.

If $m_1$ and $m_2$ are both monic polynomials of smallest degree such that $m_1(T) = m_2(T) = 0$, then $m_1 = m_2$, otherwise $m = (m_1 - m_2)/c$, where $c$ is the leading coefficient of $m_1 - m_2$, would be a monic polynomial of degree less than that of $m_1$ with $m(T) = 0$. This is a contradiction.

The minimal polynomial $m$ of a matrix $A \in M_{n \times n}(F)$ is similarly defined to be the unique monic polynomial of smallest degree that satisfies $m(A) = 0$.

Remark. In Ring Theory, the subset $I = \{f \in F[t] : f(T) = 0\}$ is an ideal of $F[t]$, called the annihilator of $T$. Since $F[t]$ is a principal ideal domain, $I$ is of the form $I = (m)$, where $m$ can be chosen to be a monic polynomial. The generator $m$ is exactly the minimal polynomial of $T$.

Theorem 2.2. Let $f \in P(F)$. Then $f(T) = 0$ if and only if the minimal polynomial $m$ of $T$ divides $f$. In particular, $m$ divides the characteristic polynomial of $T$.

Proof. Suppose $f(T) = 0$. By definition of $m$, $\deg m \leq \deg f$. By long division, there exist $q, r \in P(F)$ with $\deg r < \deg m$ such that $f = mq + r$. Since $0 = f(T) = m(T)q(T) + r(T) = r(T)$, $r$ must be 0. Therefore $m$ divides $f$.

Conversely if $f = mq$ for some $q \in P(F)$, then $f(T) = m(T)q(T) = 0$. □
**Theorem 2.3.** Let \( m \) be the minimal polynomial of \( T \) and \( \lambda \in F \). Then \( \lambda \) is an eigenvalue of \( T \) if and only if \( m(\lambda) = 0 \). Hence the minimal polynomial and the characteristic polynomial of \( T \) have the same zeros.

**Proof.** Suppose \( \lambda \) is an eigenvalue of \( T \). Let \( x \in V \) be an eigenvector of \( T \) such that \( T(x) = \lambda x \). Then \( m(T)(x) = m(\lambda)x = 0 \). So \( m(\lambda) = 0 \).

Conversely, suppose \( m(\lambda) = 0 \). Let \( p \) be the characteristic polynomial of \( T \). By Theorem 2.2, \( p(t) = m(t)q(t) \) for some \( q \in P(F) \). Then \( p(\lambda) = m(\lambda)q(\lambda) = 0 \). Therefore \( \lambda \) is an eigenvalue of \( T \).

**Lemma 2.4.** Let \( W \) be a \( T \)-invariant subspace of \( V \) (that is, \( T(W) \subseteq W \)).

1. The characteristic polynomial of \( T|_W \) divides the characteristic polynomial of \( T \).
2. The minimal polynomial of \( T|_W \) divides the minimal polynomial of \( T \).

**Proof.** Exercise 2.

**Example 2.5.** Find the minimal polynomials of the following \( n \times n \) matrices

\[
A = J_{\lambda,n} = \begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda & 1 \\
0 & 0 & \cdots & 0 & \lambda
\end{bmatrix}, \quad B = \lambda I_n, \quad C = \begin{bmatrix}
C_1 & 0 & \cdots & 0 \\
0 & C_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & C_k
\end{bmatrix},
\]

where each \( C_j \) is a square matrix with minimal polynomial \( m_j, j = 1, \ldots, k \).

**Solution.** The characteristic polynomial of \( A \) is \( p_A(t) = (-1)^n(t - \lambda)^n \). By Theorem 2.3, the minimal polynomial \( m_A \) of \( A \) must be of the form \( (t - \lambda)^l \), where \( l \leq n \). It can be checked that \( m_A(t) = (t - \lambda)^n \) since \( (A - \lambda I_n)^l \neq 0 \) for any \( l < n \).

The characteristic polynomial of \( B \) is \( p_B(t) = (-1)^n(t - \lambda)^n \), and the minimal polynomial of \( B \) is \( m_B(t) = t - \lambda \).

If \( m_C(t) \) is the minimal polynomial of \( C \), then

\[
m_C(C) = \begin{bmatrix}
m_C(C_1) & 0 & \cdots & 0 \\
0 & m_C(C_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & m_C(C_k)
\end{bmatrix} = 0.
\]

By Theorem 2.2, \( m_C(C) = 0 \) if and only if \( m_C(C_j) = 0 \) for all \( j = 1, \ldots, k \) if and only if \( m_j \) divides \( m_C \) for all \( j = 1, \ldots, k \). We conclude that \( m_C \) is the least common multiple of \( m_1, \ldots, m_k \).

### 3 Generalized Eigenspaces

From now on, we assume that \( F = \mathbb{C} \) (and thus every polynomial over \( \mathbb{C} \) splits) and that if a polynomial is of the form \( (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k} \), the zeros \( \lambda_1, \ldots, \lambda_k \) are assumed to be all distinct.
Theorem 3.3. Let \( m(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k} \) be the minimal polynomial of \( T \). Then there exist \( p_1, \ldots, p_k \in P(\mathbb{C}) \) such that
\[
p_1(t)q_1(t) + \cdots + p_k(t)q_k(t) = 1,
\]
where \( q_j(t) = m(t)/(t - \lambda_j)^{j_i} = \prod_{i \neq j}(t - \lambda_i)^{j_i} \).

Proof. Consider
\[
\frac{1}{m(t)} = \frac{1}{(t - \lambda_1)^{j_1} \cdots (t - \lambda_k)^{j_k}}.
\]
By partial fraction decomposition, there exist \( p_1, \ldots, p_k \in P(\mathbb{C}) \) such that
\[
\frac{1}{m(t)} = \frac{p_1(t)}{(t - \lambda_1)^{j_1}} + \cdots + \frac{p_k(t)}{(t - \lambda_k)^{j_k}}.
\]
Multiplying both sides by \( m(t) \), the result follows.

Recall that if \( \lambda \) is an eigenvalue of \( T \), then the eigenspace \( E_\lambda \) is defined to be the nullspace of \( T - \lambda I \). In fact, we have
\[
\{0\} \subseteq N(T - \lambda I) \subseteq N((T - \lambda I)^2) \subseteq \cdots.
\]
Since \( V \) is finite-dimensional, there exists an integer \( l \) such that \( N((T - \lambda I)^k) = N((T - \lambda I)^{k+1}) \) for all \( k \geq l \).

Definition 3.2. Let \( \lambda \) be an eigenvalue of \( T \). The generalized eigenspace of \( T \) corresponding to \( \lambda \) is a subset of \( V \) defined by
\[
K_\lambda = \{ x \in V : (T - \lambda I)^k = 0 \text{ for some } k \in \mathbb{N} \} = \bigcup_{k \in \mathbb{N}} N((T - \lambda I)^k).
\]
It is not difficult to prove that \( K_\lambda \) is a \( T \)-invariant subspace. See Exercise [3]

Theorem 3.3. Let \( m(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k} \) be the minimal polynomial of \( T \). Then for each \( j = 1, \ldots, k \), \( l_j \) is the smallest integer such that \( N((T - \lambda_j I)^{l_j}) = N((T - \lambda_j I)^{l_j+1}) \).

Hence
\[
K_{\lambda_j} = N((T - \lambda_j I)^{l_j}).
\]
Proof. It suffices to prove that
\[
N((T - \lambda_1 I)^{l_1-1}) \subseteq N((T - \lambda_1 I)^{l_1}) = N((T - \lambda_1 I)^{l_1+1}).
\]
We illustrate the case \( j = 1 \) for convenience. The rest is similar.

By definition of \( m \), there exists \( x \in V \) such that
\[
(T - \lambda_1 I)^{l_1-1}(T - \lambda_2 I)^{l_2} \cdots (T - \lambda_k I)^{l_k}(x) \neq 0.
\]
Let \( y = (T - \lambda_2 I)^{l_2} \cdots (T - \lambda_k I)^{l_k}(x) \). Then \( y \in N((T - \lambda_1 I)^{l_1}) \) but \( y \notin N((T - \lambda_1 I)^{l_1-1}) \).

Hence \( N((T - \lambda_1 I)^{l_1-1}) \subseteq N((T - \lambda_1 I)^{l_1}) \).

Let \( x \in N((T - \lambda_1 I)^{l_1+1}) \). By Theorem 3.1 there exist \( p_1, \ldots, p_k \in P(\mathbb{C}) \) such that
\[
h_1(t) + \cdots + h_k(t) = 1,
\]
where \( h_i(t) = p_i(t)q_i(t) \), \( q_i(t) = m(t)/(t - \lambda_i)^{j_i} \). So
\[
h_1(T)(x) + \cdots + h_k(T)(x) = x.
\]
We claim that after applying \( (T - \lambda_1 I)^{l_1} \), the left hand side of (2) becomes 0. For \( i = 2, \ldots, k \), \( (T - \lambda_1 I)^{l_1}h_i(T)(x) = 0 \) since \( x \in N((T - \lambda_1 I)^{l_1+1}) \) and \( (t - \lambda_i)^{j_i+1} \) divides \( (t - \lambda_1)^{j_1}h_i(t) \). For \( i = 1 \), \( (T - \lambda_1 I)^{l_1}h_1(T)(x) = 0 \) since \( (t - \lambda_1)^{j_1}h_1(t) = p_1(t)m(t) \). Therefore \( N((T - \lambda_1 I)^{l_1+1}) \subseteq N((T - \lambda_1 I)^{l_1}) \). This completes the proof.
4 Primary Decomposition Theorem

We recall the definition of direct sum of vector spaces and its equivalent statements.

**Theorem 4.1.** Let $W_1, \ldots, W_k$ be subspaces of $V$, and let $W = W_1 + \cdots + W_k$. The followings are equivalent.

1. $x_1 + \cdots + x_k = 0$ implies $x_j = 0$, where $x_j \in W_j$, $j = 1, \ldots, k$.
2. Every vector $x \in W$ has a unique expression $x = x_1 + \cdots + x_k$ for $x_j \in W_j$.
3. If $\beta_j$ is a basis for $W_j$, $j = 1, \ldots, k$, then $\beta = \beta_1 \cup \cdots \cup \beta_k$ is a basis for $W$.
4. $\dim W = \dim W_1 + \cdots + \dim W_k$.
5. $W_j \cap \sum_{i \neq j} W_i = \{0\}$ for all $j = 1, \ldots, k$.

If one of the above condition holds, we say $W$ is the direct sum of $W_1, \ldots, W_k$, denoted by

$$W = W_1 \oplus \cdots \oplus W_k.$$  

**Proof.** We only prove that 2 implies 3. The rest is left as an exercise. (Exercise 4)

Let $\beta_j = \{u_{j1}, \ldots, u_{jn_j}\}$ be a basis for $W_j$, $j = 1, \ldots, k$. Suppose

$$c_{j1}u_{j1} + \cdots + c_{jn_j}u_{jn_j} = 0.$$

Note that for each $j$, $c_{j1}u_{j1} + \cdots + c_{jn_j}u_{jn_j} \in W_j$. Since, by assumption, the only combination of 0 is $0 = 0 + \cdots + 0$, we have $c_{j1}u_{j1} + \cdots + c_{jn_j}u_{jn_j} = 0$ for all $j$. Since $\beta_j$ is a basis, $c_{j1} = \cdots = c_{jn_j} = 0$ for all $j$. Therefore $\beta$ is linearly independent.

If $x \in W$, then by assumption, there exist $x_j \in W_j$, $j = 1, \ldots, k$, such that $x = x_1 + \cdots + x_k$. Since each $x_j$ is a linear combination of vectors in $\beta_j$, $x$ is a linear combination of vectors in $\beta = \beta_1 \cup \cdots \cup \beta_k$. Therefore $\beta$ spans $W$. \qed

The following theorem is one of the main theorems in this article. It plays an important role in the theory of Jordan forms.

**Theorem 4.2** (Primary Decomposition Theorem). Let $m(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k}$ be the minimal polynomial of $T$. Then

$$V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_k}.$$  

Moreover, if $T_j$ is the restriction of $T$ on $K_{\lambda_j}$, then the minimal polynomial of $T_j$ is $(t - \lambda_j)^{l_j}$.

We need the following theorem before proving the Primary Decomposition Theorem.

**Theorem 4.3.** Let $E_1, \ldots, E_k$ be linear operators on $V$, and let $W_j = R(E_j)$. Suppose

1. $I = E_1 + \cdots + E_k$,
2. $E_iE_j = 0$ for $i \neq j$,
3. $E_j^2 = E_j$ for $j = 1, \ldots, k$.

Then

$$V = W_1 \oplus \cdots \oplus W_k.$$
Proof. For each \( x \in V \), by condition 1, \( x = E_1(x) + \cdots + E_k(x) \in W_1 + \cdots + W_k \).

Suppose \( y_j \in W_j \) for \( j = 1, \ldots, k \) and \( y_1 + \cdots + y_k = 0 \). For each \( j \), since \( y_j \in R(E_j) \), there exists \( x_j \in V \) such that \( y_j = E_j(x_j) \). Then \( E_1(x_1) + \cdots + E_k(x_k) = 0 \). Applying \( E_j \) and by condition 2, we get \( E_j E_j(x_j) = 0 \). Therefore, by condition 3, \( y_j = E_j(x_j) = E_j^2(x_j) = 0 \). By Theorem 4.1, \( V = W_1 \oplus \cdots \oplus W_k \).

Proof of Theorem 4.2. We claim that there exist \( h_1, \ldots, h_k \in P(\mathbb{C}) \) that satisfy the following statements.

1. \( I = h_1(T) + \cdots + h_k(T) \),
2. \( h_i(T)h_j(T) = 0 \) for \( i \neq j \),
3. \( h_j(T)^2 = h_j(T) \) for \( j = 1, \ldots, k \),
4. \( R(h_j(T)) = K_{\lambda_j} \), where \( K_{\lambda_j} = \mathbb{N}((T - \lambda_j I)^l), j = 1, \ldots, k \).

The existence of \( h_1, \ldots, h_k \) and statement 1 follow from equation \([1]\) in the proof of Theorem 3.3. To prove statement 2, note that \( m \) divides \( h_i h_j \) and apply Theorem 2.2. Statement 3 follows from statements 1 and 2.

We prove statement 4. If \( x \in R(h_j(T)) \), then there exists \( u \in V \) such that \( x = h_j(T)(u) \). So

\[
(T - \lambda_j I)^l(x) = (T - \lambda_j I)^l h_j(T)(u) = p_j(T)m(T)(u) = 0.
\]

This implies \( x \in K_{\lambda_j} \).

If \( x \in K_{\lambda_j} \), then \( (T - \lambda_j I)^l(x) = 0 \). Note that for all \( i \neq j \), \( (T - \lambda_j I)^l \) divides \( h_i(t) \). Hence \( h_i(T)(x) = 0 \) and by condition 1,

\[
x = h_1(T)(x) + \cdots + h_k(T)(x) = h_j(T)(x) \in R(h_j(T)).
\]

Finally we apply Theorem 4.3 to conclude that \( V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_k} \).

Let \( m_j \) be the minimal polynomial of \( T_j \). Since \( (T_j - \lambda_j I)^l = 0 \) on \( K_{\lambda_j} \), by Theorem 2.2 \( m_j \) divides \( (T - \lambda_j I)^l \). By Theorem 3.3 \( (T_j - \lambda_j I)^s \neq 0 \) on \( K_{\lambda_j} \) for all \( s < l_j \). Hence \( m_j(t) = (T - \lambda_j I)^l \).

Definition 4.4. A linear operator \( T \) on \( V \) is said to be nilpotent if there exists a positive integer \( k \) such that \( T^k = 0 \). The smallest positive integer \( k \) such that \( T^k = 0 \) is called the index of \( T \).

Theorem 4.5 (Jordan-Chevalley Decomposition). Suppose the minimal polynomial of \( T \) splits. Then there exist a diagonalizable operator \( D \) and a nilpotent operator \( N \) such that \( T = D + N \) and \( DN = ND \). Moreover, this decomposition is unique.

Sketch of proof. Suppose the minimal polynomial of \( T \) is \( m(t) = (t - \lambda_1)^{\nu_1} \cdots (t - \lambda_k)^{\nu_k} \). As in the proof of Theorem 4.2, there exist \( h_1, \ldots, h_k \in P(\mathbb{C}) \) such that \( h_1(T), \ldots, h_k(T) \) satisfy that 4 conditions. Then \( T = Th_1(T) + \cdots + Th_k(T) \). Define \( D = \lambda_1 h_1(T) + \cdots + \lambda_k h_k(T) \), and define \( N = T - D \). Then \( D \) is diagonalizable and \( N = (T - \lambda_1 I)h_1(T) + \cdots + (T - \lambda_k I)h_k(T) \) is nilpotent. Since \( D \) and \( N \) are polynomials in \( T \), they commute.

Suppose \( T = D' + N' \), where \( D' \) is diagonalizable, \( N' \) is nilpotent, and \( D'N' = N'D' \). Since \( D' \) and \( N' \) commute with each other, they commute with \( T \). Thus \( D' \) and \( N' \) commute with any polynomial in \( T \); in particular, with \( D \) and with \( N \). Since \( D + N = D' + N' \), we have \( D - D' = N' - N \).
Since $D$ and $D'$ are diagonalizable and they commute, they are simultaneously diagonalizable (see Friedberg, Linear Algebra). Hence $D - D'$ is diagonalizable. Since $N$ and $N'$ commute and they are nilpotent, $N' - N$ is nilpotent. Since the only linear operator that is diagonalizable and nilpotent is 0, we conclude that $D - D' = N' - N = 0$. \hfill \qed

**Theorem 4.6.** Suppose $\dim V = n$. Let $p(t) = (-1)^n(t - \lambda_1)^{n_1}\cdots(t - \lambda_k)^{n_k}$ be the characteristic polynomial of $T$. Then for each $j$, $\dim K_{\lambda_j} = m_{\lambda_j}$.

Moreover, if $T_j$ is the restriction of $T$ on $K_{\lambda_j}$, then the characteristic polynomial of $T_j$ is $(t - \lambda_j)^{n_j}$.

Recall that the geometric multiplicity of an eigenvalue $\lambda$ is defined to be the dimension of the eigenspace $E_{\lambda}$. **Theorem 4.6** tells us that the algebraic multiplicity of $\lambda$ is equal to the dimension of the generalized eigenspace $K_{\lambda}$. This also implies that the geometric multiplicity of $\lambda$ does not exceed the algebraic multiplicity of $\lambda$.

**Proof.** For each $j$, by **Theorem 4.2**, the minimal polynomial of $T_j$ is $p_j(t) = (t - \lambda_j)^{d_j}$. By **Theorem 2.3**, the characteristic polynomial of $T_j$ is of the form $p_j(t) = (-1)^{d_j}(t - \lambda_j)^{d_j}$, where $d_j = \dim K_{\lambda_j}$. By **Lemma 2.4**, $p_j$ divides $p$. Therefore $d_j \leq n_j$. By **Theorems 4.1** and 4.2, $n = d_1 + \cdots + d_k \leq n_1 + \cdots + n_k \leq n$. Hence $d_j = n_j$ for all $j = 1, \ldots, k$. \hfill \qed

The following theorem gives a necessary and sufficient condition for diagonalizability of $T$.

**Theorem 4.7.** $T$ is diagonalizable if and only if the minimal polynomial $m$ of $T$ is of the form

$$m(t) = (t - \lambda_1)\cdots(t - \lambda_k),$$

where $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $T$.

**Proof.** Suppose $T$ is diagonalizable. Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of $T$ and define $f(t) = (t - \lambda_1)\cdots(t - \lambda_k)$. By **Theorem 2.3**, $f$ divides $m$. We claim that $f(T) = 0$, then, by **Theorem 2.2**, $m$ divides $f$ and thus $f = m$.

By assumption, there exists a basis $\beta$ for $V$ consisting of eigenvectors of $T$. For any $v \in \beta$, let $\lambda$ be an eigenvalue of $T$ such that $T(v) = \lambda v$. Then $f(T)(v) = f(\lambda)v = 0$. Therefore $f(T) = 0$.

Conversely, suppose $m(t) = (t - \lambda_1)\cdots(t - \lambda_k)$. By **Theorem 4.2**, $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$, where $E_{\lambda_j} = N(T - \lambda_j I)$. We find that $V$ is the direct sum of eigenspaces of $T$. Therefore $T$ is diagonalizable. \hfill \qed

**Example 4.8.** Every projection (a linear operator $P$ satisfying $P^2 = P$) on a finite-dimensional $F$-vector space is, by **Theorem 4.7**, diagonalizable since its minimal polynomial divides $f(t) = t(t - 1)$.

## 5 Jordan Canonical Forms and Jordan bases

**Definition 5.1.** Let $A_1, \ldots, A_k$ be matrices (may have different sizes). The direct sum of $A_1, \ldots, A_k$ is defined to be the block matrix

$$A_1 \oplus \cdots \oplus A_k = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}.$$
Definition 5.2. 1. A Jordan block of size \( n \) corresponding to the eigenvalue \( \lambda \) is the \( n \times n \) matrix
\[
J_{\lambda,n} = \begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda & 1 \\
0 & 0 & \cdots & 0 & \lambda
\end{bmatrix}.
\]

2. A Jordan matrix is a direct sum of some Jordan blocks.

3. If there exists a basis \( \beta \) for \( V \) such that \([T]_{\beta}\) is a Jordan matrix, we say \( T \) has a Jordan form, and \( \beta \) is called a Jordan basis for \( T \).

4. We say a matrix \( A \in M_{n \times n}(F) \) has a Jordan form if \( A \) is similar to a Jordan matrix (or, equivalently, if \( L_A \) has a Jordan form).

Lemma 5.3. Suppose \( V = W_1 \oplus \cdots \oplus W_k \) and each \( W_j \) has a basis \( \beta_j \). If each \( W_j \) is \( T \)-invariant, then \([T]_{\beta} = [T|_{W_1}]_{\beta_1} \oplus \cdots \oplus [T|_{W_k}]_{\beta_k}\), where \( \beta = \beta_1 \cup \cdots \cup \beta_k \) is a basis for \( V \).

Proof. Exercise 7.

Suppose \( V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_k} \). It is obvious that the restriction of \( T - \lambda_j I \) on \( K_{\lambda_j} \) is nilpotent for each \( j \). If there exists a basis \( \beta_j \) for \( K_{\lambda_j} \) such that \([T - \lambda_j I]|_{K_{\lambda_j}}|_{\beta_j}\) is a Jordan matrix, then by Theorem 4.1 and Lemma 5.3, \( \beta = \beta_1 \cup \cdots \cup \beta_k \) is a basis for \( V \) and \([T]_{\beta}\) is a Jordan matrix. This observation suggests that we focus on a linear operator on \( V \) which is nilpotent.

Before the discussion about nilpotent linear operators, we begin from some definitions.

Definition 5.4. Let \( W \) be a subspace of \( V \).

1. If \( W' \) is a subspace of \( V \) such that \( V = W \oplus W' \), we say \( W' \) is a complement of \( W \).

2. We say \( \{u_1, \ldots, u_k\} \) is linearly independent over \( W \) (or \( \{u_1, \ldots, u_k\} \) is linearly independent in \( V/W \)) if \( c_1 u_1 + \cdots + c_k u_k \in W \) implies \( c_1 = \cdots = c_k = 0 \).

3. We say \( \{u_1, \ldots, u_k\} \) is a basis for \( V \) over \( W \) (or \( \{u_1, \ldots, u_k\} \) is a basis for \( V/W \)) if \( \{u_1, \ldots, u_k\} \) is a basis for a complement of \( W \).

4. The dimension of \( V \) over \( W \) (or the dimension of \( V/W \)) is defined to be the number \( \dim V - \dim W \).

Note that if \( W = \{0\} \), the above definitions coincide with the usual definitions of basis and linear independence.

Remark. Using the language of quotient spaces, saying \( u_1, \ldots, u_k \in V \) is linearly independent over \( W \) is equivalent to saying \( \bar{u}_1, \ldots, \bar{u}_k \in V/W \) is linearly independent.

Lemma 5.5. Let \( W \) be a subspace of \( V \). If \( \{u_1, \ldots, u_k\} \) is linearly independent over \( W \), then it can be extended to a basis for \( V \) over \( W \).

Proof. Exercise 8.
Besides Theorem 4.2, the following theorem is another main theorem in this article. It also plays an important role in both theory and computation of Jordan forms.

**Theorem 5.6.** Let $T$ be a nilpotent linear operator on $V$. Then $T$ has a Jordan basis.

**Proof.** Let $T$ have index $k$. Let $N_j = N(T^j)$, $j = 0, \ldots, k$. We have

$$\{0\} = N_0 \subset N_1 \subset N_2 \subset \cdots \subset N_{k-1} \subset N_k = V.$$ 

Our goal is to find vectors $u_{11}, \ldots, u_{1m_1}, u_{21}, \ldots, u_{2m_2}, \ldots, u_{k1}, \ldots, u_{km_k}$ such that

$$\beta = \{ T^{k-1}(u_{11}), \ldots, T(u_{11}), u_{11} \} \cup \cdots \cup \{ T^{k-1}(u_{1m_1}), \ldots, T(u_{1m_1}), u_{1m_1} \}$$

$$\cup \{ T^{k-2}(u_{21}), \ldots, T(u_{21}), u_{21} \} \cup \cdots \cup \{ T^{k-2}(u_{2m_2}), \ldots, T(u_{2m_2}), u_{2m_2} \}$$

$$\cup \cdots$$

$$\cup \{ u_{k1} \} \cup \cdots \cup \{ u_{km_k} \}$$

forms a basis for $V$ and $[T]_\beta$ is a Jordan matrix.

First we choose $u_{11}, \ldots, u_{1m_1} \in N_k$, a basis for $N_k = V$ over $N_{k-1}$. We claim that

$$T(u_{11}), \ldots, T(u_{1m_1}) \in N_{k-1}$$

are linearly independent over $N_{k-2}$. Suppose

$$c_1 T(u_{11}) + \cdots + c_{m_1} T(u_{1m_1}) \in N_{k-2}.$$ 

Apply $T^{k-2}$ to get

$$T^{k-1}(c_1 u_{11} + \cdots + c_{m_1} u_{1m_1}) = 0.$$

So $c_1 u_{11} + \cdots + c_{m_1} u_{1m_1} \in N_{k-1}$. Since $u_{11}, \ldots, u_{1m_1}$ are linearly independent over $N_{k-1}$, $c_1 = \cdots = c_{m_1} = 0$. This proves the claim.

Now, by Lemma 5.5 we extend (if needed) $T(u_{11}), \ldots, T(u_{1m_1})$ to

$$T(u_{11}), \ldots, T(u_{1m_1}), u_{21}, \ldots, u_{2m_2} \in N_{k-1},$$

a basis for $N_{k-1}$ over $N_{k-2}$. By the same argument,

$$T^2(u_{11}), \ldots, T^2(u_{1m_1}), T(u_{21}), \ldots, T(u_{2m_2}) \in N_{k-2}$$

are linearly independent over $N_{k-3}$. By Lemma 5.5 again, they can be extended to

$$T^2(u_{11}), \ldots, T^2(u_{1m_1}), T(u_{21}), \ldots, T(u_{2m_2}), u_{31}, \ldots, u_{3m_3} \in N_{k-2}$$

a basis for $N_{k-2}$ over $N_{k-3}$. Continue this process until we get a basis for $N_1$.

For each $j = 1, \ldots, k$, let $W_j$ be a subspace of $N_j$ such that $N_j = N_{j-1} \oplus W_j$ and each $W_j$ has a basis as listed below:

- Basis for $W_k$: $u_{11}, \ldots, u_{1m_1}$
- Basis for $W_{k-1}$: $T(u_{11}), \ldots, T(u_{1m_1})$, $u_{21}, \ldots, u_{2m_2}$
- $\cdots$
- Basis for $W_1$: $T^{k-1}(u_{11}), \ldots, T^{k-1}(u_{1m_1})$, $T^{k-2}(u_{21}), \ldots, T^{k-2}(u_{2m_2})$, $\ldots$, $u_{k1}, \ldots, u_{km_k}$

By Theorem 4.1, these vectors form a basis for $V$ since $V = W_1 \oplus \cdots \oplus W_k$. Finally we reorder the vectors to get the desired $\beta$ as described in equation (3). This completes the proof. \qed
Here are some observations from the proof of the previous theorem. First we observe that \( u_{11}, \ldots, u_{1m_1} \) form a basis for \( N_k \) over \( N_{k-1} \). So the dimension of \( N_k \) over \( N_{k-1} \) is \( m_1 \). Similarly, \( T(u_{11}), \ldots, T(u_{1m_1}), u_{21}, \ldots, u_{2m_2} \) form a basis for \( N_{k-1} \) over \( N_{k-2} \). So the dimension of \( N_{k-1} \) over \( N_{k-2} \) is \( m_1 + m_2 \). And so on. Thus we have

\[
\begin{align*}
m_1 &= \dim W_k = \dim N_k - \dim N_{k-1}, \\
m_1 + m_2 &= \dim W_{k-1} = \dim N_{k-1} - \dim N_{k-2}, \\
\vdots \\
m_1 + m_2 + \cdots + m_k &= \dim W_1 = \dim N_1.
\end{align*}
\]

On the other hand, denote

\[
\beta_{ij} = \{T^{k-i}(u_{ij}), \ldots, T(u_{ij}), u_{ij}\}, \quad i = 1, \ldots, k, \quad j = 1, \ldots, m_i
\]

for convenience. Let \( W_{ij} = \text{span} \beta_{ij} \) be a \( T \)-cyclic subspace of \( V \) generated by \( u_{ij} \). Then \( [T|_{W_{ij}}]_{\beta ij} = J_{0,k-i+1} \). There are \( m_i \) such blocks in \( [T]_{\beta} \). Therefore

1. \( \dim N_1 \) is the number of Jordan blocks (of size \( \geq 1 \)).
2. \( \dim N_i - \dim N_{i-1} \) is the number of Jordan blocks of size \( \geq i \), \( i = 2, \ldots, k \).

Note that once a nilpotent linear operator \( T \) is given, the dimension of each \( N_j \) is determined. By the above observation, the number of Jordan blocks of each size are also determined. This tells us that we are able to determine the Jordan form of \( T \) without finding its Jordan basis explicitly. All we have to do is to calculate the nullity of \( T^j \) for each \( j \). Hence the Jordan form of a nilpotent linear operator is unique up to permutation of Jordan blocks.

**Example 5.7.** Let \( V \) be a \( \mathbb{C} \)-vector space of dimension 9. Let \( T \) be a linear operator on \( V \) with characteristic polynomial \( p(t) = (-1)^9(t - 2)^9 \). Suppose the dimensions of \( N((T - 2I)^j) \) are listed as the following table.

<table>
<thead>
<tr>
<th>( j )</th>
<th>( \nullity (T - 2I)^j )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

Find the Jordan form of \( T \).

**Proof.** Clearly, \( T - 2I \) is a nilpotent linear operator on \( V \). We follow the step in the proof of Theorem 5.6 to find the Jordan form of \( S \).

Let \( N_j = N((T - 2I)^j), \quad j = 0, 1, 2, 3, 4 \).

Let \( u_{11} \) be a basis for \( N_4 \) over \( N_3 \). Then \( (T - 2I)(u_{11}) \in N_3 \) is linearly independent over \( N_2 \). Extend it to a basis \( (T - 2I)(u_{11}), u_{21} \) for \( N_3 \) over \( N_2 \). Continuing the process as in the proof of Theorem 5.6 we obtain a Jordan basis for \( T - 2I \):

- Basis for \( N_4/N_3 \): \( u_{11} \)
- Basis for \( N_3/N_2 \): \( (T - 2I)(u_{11}) \), \( u_{21} \)
- Basis for \( N_2/N_1 \): \( (T - 2I)^2(u_{11}) \), \( (T - 2I)(u_{21}) \)
- Basis for \( N_1 \): \( (T - 2I)^3(u_{11}) \), \( (T - 2I)^2(u_{21}) \), \( u_{41} \), \( u_{42} \)

Let

\[
\beta = \{(T - 2I)^3(u_{11}), (T - 2I)^2(u_{11}), (T - 2I)(u_{11}), u_{11}\} \\
\cup \{(T - 2I)^2(u_{21}), (T - 2I)(u_{21}), u_{21}\} \cup \{u_{41}\} \cup \{u_{42}\}.
\]
Then $[T - 2I]_\beta = J_{0,4} \oplus J_{0,3} \oplus J_{0,1} \oplus J_{0,1}$. Therefore we conclude that the Jordan form of $T$ is

$$[T]_\beta = [T - 2I]_\beta + 2I_9 = J_{2,4} \oplus J_{2,3} \oplus J_{2,1} \oplus J_{2,1} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 2 & 1 \\ 2 & 0 & \end{bmatrix}.$$

Note that once we have the dimension of each $N_j$, we can drop the label of each vector and directly draw the helpful “dot diagram”, like the following first figure.

There are much information in the dot diagram. As for the second figure, each rectangle represents the space $N_j$, and the dots in the rectangle form a basis for $N_j$. For example, there are 6 dots enclosed by rectangle $N_2$; these 6 dots form a basis for $N_2$.

As for the third figure, each column represents a Jordan block, and the number of dots in each column represents the size of the Jordan block. For example, the 4 dots of the first column represents the Jordan block $J_{2,4}$.

Now we discuss the general case. Let $T$ be a linear operator on $V$ with minimal polynomial $m(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k}$. By Theorem 4.2, $V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_k}$. For an eigenvalue $\lambda$ of $T$, $(T - \lambda I)|_{K_{\lambda}}$ is a nilpotent linear operator on $K_{\lambda}$. By Theorem 5.6, there exists a basis $\beta_{\lambda}$ for $K_{\lambda}$ such that $[(T - \lambda I)|_{K_{\lambda}}]_{\beta_{\lambda}}$ is a Jordan matrix. This implies $[T|_{K_{\lambda}}]_{\beta_{\lambda}}$ is a Jordan matrix. If $\beta = \beta_{\lambda_1} \cup \cdots \cup \beta_{\lambda_k}$, then by Theorem 4.1 and Lemma 5.3, $\beta$ is a basis for $V$ such that $[T]_\beta$ is a Jordan matrix.

With the aid of the following lemma, the nullity of each $(T - \lambda I)|_{K_{\lambda}}$ can be computed by directly computing the nullity of $(T - \lambda I)^j$. This is helpful for drawing dot diagrams.

**Lemma 5.8.** Let $\lambda$ be an eigenvalue of $T$. Then $N((T - \lambda I)^j) = N((T - \lambda I)^j|_{K_{\lambda}})$ for all $j$.

**Proof.** Exercise [10]

The conclusions are summarized in the following theorems, which are the goals of this article.

**Theorem 5.9.** Let $T$ be a linear operator on a $\mathbb{C}$-vector space $V$. Then $T$ has a Jordan basis. Moreover, the Jordan form of $T$ is unique up to permutation of Jordan blocks.
Theorem 5.10. Let $A, B \in M_{n \times n}(\mathbb{C})$. Then $A$ is similar to $B$ if and only if they have the same Jordan form.

Proof. Exercise 11

Example 5.11. Let $V$ be a $\mathbb{C}$-vector space of dimension 13. Let $T$ be a linear operator on $V$ with characteristic polynomial $p(t) = (-1)^{13}(t-2)^9(t-3)^4$. Suppose the dimensions of $N((T - 2I)^j)$ and $N((T - 3I)^j)$ are listed as the following table.

<table>
<thead>
<tr>
<th>$j$</th>
<th>nullity of $(T - 2I)^j$</th>
<th>nullity of $(T - 3I)^j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Find the Jordan form of $T$ and the minimal polynomial of $T$.

Solution. By Theorem 4.2, $V = K_2 \oplus K_3$. Let $T_2 = (T - 2I)|_{K_2}$ and $T_3 = (T - 3I)|_{K_3}$. Then $T_2$ and $T_3$ are nilpotent linear operators on $K_2$ and $K_3$, respectively. By Lemma 5.8, we may calculate each nullity$(T - \lambda I)^j|_{K\lambda}$ and draw the dot diagrams for $\lambda = 2$ and $\lambda = 3$:

Thus the Jordan form of $T|_{K_2}$ is $J_{2,4} \oplus J_{2,3} \oplus J_{2,1} \oplus J_{2,1}$, and the Jordan form of $T|_{K_3}$ is $J_{3,3} \oplus J_{3,1}$. Therefore the Jordan form of $T$ is

$$J_{2,4} \oplus J_{2,3} \oplus J_{2,1} \oplus J_{2,1} \oplus J_{3,3} \oplus J_{3,1}.$$ 

Since the minimal polynomial $m$ of $T$ has the same zeros as the characteristic polynomial of $T$, $m$ is of the form $(t - 2)^4(t - 3)^3$. By Theorem 3.3, since $\dim K_2 = 4$ and $\dim K_3 = 3$, we conclude that $m(t) = (t - 2)^4(t - 3)^3$.

We remark that the minimal polynomial can also be calculated from the Jordan form of $T$. As in the solution of Example 2.5, the minimal polynomial of $T$ is the least common multiple of the minimal polynomials of each Jordan block. That is,

$$m(t) = \text{lcm}((t - 2)^4, (t - 2)^3, (t - 2), (t - 2)^3, (t - 3)^3, (t - 3)) = (t - 2)^4(t - 3)^3. \square$$

Example 5.12. Let

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

1. Find the Jordan form $J$ of $A$.
2. Find the minimal polynomial of $A$.
3. Find an invertible matrix $Q$ such that $AQ = QJ$.
**Solution.** The characteristic polynomial of $A$ is $p(t) = (-1)^5t^4(t - 1)$. By Theorem 4.2, $V = K_0 \oplus K_1$. By Theorem 4.6, $\dim K_0 = 4$ and $\dim K_1 = 1$.

For $\lambda = 0$, we compute that
\[
A^2 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & -1 \\
-1 & 0 & 0 & 0 & 0
\end{bmatrix}, \\
A^3 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0
\end{bmatrix},
\]
and that nullity $A = 2$, nullity $A^2 = 3$, nullity $A^3 = 4$. We draw the dot diagram for $\lambda = 0$:

\[
\bullet \\
\bullet \\
\bullet \\
\bullet
\]

For $\lambda = 1$, we compute that
\[
A - I = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & -2 & 0 & 0 & -1 \\
1 & -1 & -1 & 0 & -1 \\
0 & 0 & 0 & -1 & -1 \\
-1 & 1 & 0 & 0 & 0
\end{bmatrix},
\]
and that nullity $(A - I) = 1$. (In fact, there is no need to directly compute it; just remember we already have $\dim K_1 = 1$.)

We conclude that the Jordan form of $A$ is
\[
J = J_{0,3} \oplus J_{0,1} \oplus J_{1,1}.
\]

The minimal polynomial $m$ of $A$ is $m(t) = t^3(t - 1)$.

To find a Jordan basis for $A$, we follow the steps in the proof of Theorem 5.6. We start from finding a basis for $K_0$. First we need to choose $u_{11} \in N(A^3)$ that is a basis for $N(A^3)$ over $N(A^2)$. We pick $u_{11} = (0, 1, 0, 0, 0)^t$. Then $Au_{11} \in N(A^2)$ is linearly independent over $N(A)$. Since the dimension of $N(A^2)$ over $N(A)$ is $3 - 2 = 1$, $Au_{11}$ is a basis for $N(A^2)$ over $N(A)$. Then $A^2u_{11} \in N(A)$ is linearly independent. Extend it to a basis for $N(A)$ by putting $u_{31} = (0, 0, 1, 0, 0)^t$. Then $\{A^2u_{11}, Au_{11}, u_{11}\} \cup \{u_{31}\}$ is a basis for $K_0$. For $K_1$, we choose $v = (1, 1, 1, 1, -1)$ to be a basis. Therefore $\{A^2u_{11}, Au_{11}, u_{11}\} \cup \{u_{31}\} \cup \{v\}$ forms a Jordan basis for $A$. Hence,
\[
Q = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 \\
0 & -1 & 1 & 0 & 1 \\
0 & -1 & 0 & 1 & 1 \\
-1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & -1
\end{bmatrix}
\]
is invertible and $AQ = QJ$.

**Example 5.13.** Let $p(t) = (-1)^5(t - 2)^2(t - 3)^3$ be the characteristic polynomial of $T$. Find all possible Jordan forms of $T$ explicitly.

**Solution.** The possible dot diagrams for each eigenvalue are listed below.
6 Exercises  

The possible 6 Jordan forms of $T$ are listed below.

<table>
<thead>
<tr>
<th>Λ = 2</th>
<th>Λ = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\begin{bmatrix} 2 &amp; 2 \ 3 &amp; 3 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2 &amp; 2 \ 1 &amp; 3 \end{bmatrix}$</td>
</tr>
<tr>
<td>$\begin{bmatrix} 2 &amp; 1 \ 3 &amp; 3 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2 &amp; 1 \ 2 &amp; 3 \end{bmatrix}$</td>
</tr>
<tr>
<td>$\begin{bmatrix} 2 &amp; 1 \ 2 &amp; 3 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2 &amp; 1 \ 2 &amp; 3 \end{bmatrix}$</td>
</tr>
<tr>
<td>$\begin{bmatrix} 2 &amp; 1 \ 3 &amp; 3 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2 &amp; 1 \ 3 &amp; 3 \end{bmatrix}$</td>
</tr>
<tr>
<td>$\begin{bmatrix} 2 &amp; 1 \ 3 &amp; 3 \end{bmatrix}$</td>
<td>$\begin{bmatrix} 2 &amp; 1 \ 3 &amp; 3 \end{bmatrix}$</td>
</tr>
</tbody>
</table>

Example 5.14. Find $J_{\lambda,n}^k$ explicitly for any $k \in \mathbb{N}$.

Solution. Write $J_{\lambda,n} = \lambda I_n + N$, where $N = J_{\lambda,n} - \lambda I_n$. Then $N$ is nilpotent. Hence we may expand $(\lambda I_n + N)^k$ by binomial theorem.

$$J_{\lambda,n}^k = (\lambda I_n + N)^k = \lambda^k I_n + \binom{k}{1} \lambda^{k-1} N + \binom{k}{2} \lambda^{k-2} N^2 + \cdots \text{ (finite terms)}$$

$$= \begin{bmatrix} \lambda & \binom{k}{1} \lambda^{k-1} & \binom{k}{2} \lambda^{k-2} & \cdots & \binom{k}{n-1} \lambda^{k-n+1} \\ 0 & \lambda & \binom{k}{1} \lambda^{k-1} & \cdots & \binom{k}{n-2} \lambda^{k-n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{k}{1} \lambda^{k-1} \\ 0 & 0 & 0 & \cdots & \lambda^k \end{bmatrix}$$

Here $\binom{k}{j}$ is defined to be 0 for all $j > k$.  

6 Exercises

Exercise 1. Prove the following statements.

1. If $A, B \in M_{n \times n}(F)$ are similar, then they have the same minimal polynomial.

2. Let $\beta$ be a basis for $V$. Then $T$ and $[T]_\beta$ have the same minimal polynomial.

3. Let $A \in M_{n \times n}(F)$. Then $A$ and $L_A$ have the same minimal polynomial.

Exercise 2. Prove Lemma 2.4.

Exercise 3. Let $T$ be a linear operator on an $F$-vector space $V$. Let $\lambda$ be an eigenvalue of $T$. Prove that $K_\lambda$ is a subspace of $V$. Then prove that $K_\lambda$ is $T$-invariant.

[Note that $T(x) = (T - \lambda I)(x) + \lambda x$.]
Exercise 4. Prove Theorem 4.1: \(1 \Rightarrow 2, 3 \Rightarrow 4, 4 \Rightarrow 5, 5 \Rightarrow 1\).

Exercise 5. Prove that for any \(k \in \mathbb{N}\), if \(A \in M_{n \times n}(\mathbb{C})\) satisfies \(A^k = I_n\), then \(A\) is diagonalizable.

Exercise 6. Let \(A \in M_2(\mathbb{R})\). If \(A^2 - 3A + 2I_2 = 0\), prove that either \(A = I_2, A = 2I_2\), or \(A\) is similar to \(\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\).

Exercise 7. Prove Lemma 5.3.

Exercise 8. Prove Lemma 5.5.

Exercise 9. Let \(A \in M_{n \times n}(\mathbb{C})\). Prove that the followings are equivalent.

1. \(A\) is nilpotent.
2. The characteristic polynomial of \(A\) is \(p(t) = (-1)^n t^n\).
3. The minimal polynomial of \(A\) is \(m(t) = t^k\) for some positive integer \(k\).
4. 0 is the only eigenvalue of \(A\).

(Hence if \(A\) has index \(k\), then \(k \leq n\).)


Exercise 11. Prove Theorem 5.10.

Exercise 12. Is \(A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}\) similar to \(B = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}\)? Justify your answer.

Exercise 13. Let
\[
A = \begin{bmatrix}
0 & -3 & 1 & 2 \\
-2 & 1 & -1 & 2 \\
-2 & 1 & -1 & 2 \\
-2 & -3 & 1 & 4
\end{bmatrix}.
\]
Suppose it is known that the characteristic polynomial of \(A\) is \(p(t) = (-1)^4 t^2 (t - 2)^2\).

1. Find the Jordan form \(J\) of \(A\).
2. Find the minimal polynomial of \(A\).
3. Find an invertible matrix \(Q\) such that \(AQ = QJ\).

Exercise 14. Let \(V = \{ax^2 + bxy + cy^2 + dx + ey + f : a, b, c, d, e, f \in \mathbb{R}\}\) be a vector space. Let \(T : V \to V\) be defined by \(T(f(x, y)) = \frac{\partial f}{\partial x}(x, y) + \frac{\partial f}{\partial y}(x, y)\).

1. Find the Jordan form of \(T\).
2. Find the minimal polynomial of \(T\).
3. Find a Jordan basis \(\beta\) for \(T\).

Exercise 15. Let \(V\) be a \(\mathbb{C}\)-vector space of dimension 8. Let \(T\) be a linear operator on \(V\) with characteristic polynomial \(p(t) = (-1)^8 (t - 3)^8\) and minimal polynomial \(m(t) = (t - 3)^3\). Find all possible Jordan forms of \(T\).

Exercise 16. Prove that for any \(A \in M_{n \times n}(\mathbb{C})\), \(A\) is similar to its transpose \(A^t\).