UNIT 2 JORDAN FORM

Structure **Page No.**

2.1	Introduction Objectives	25
2.2	Primary Decomposition	25
2.3	Jordan Canonical Form	33
2.4	Summary	40
2.5	Solutions/Answers	40

2.1 INTRODUCTION

In the previous unit, you have studied similar matrices. In this unit, we lead on to a related concept, that of similarity transforms. A is called a similarity transform of B if $B = P^{-1}AP$ for some invertible matrix P. In this unit, we shall be working towards categorising linear operators (or square matrices) in terms of near diagonal operator or matrices that they are similar to.

To start with, in Sec. 2.2, we will introduce you to an extremely useful theorem, which tells us how a vector space can be decomposed in terms of eigenspaces of an operator. At the same time, we shall discuss this result for matrices. This result also leads us to another condition for diagonalisability of a linear operator or a square matrix.

In the next section, Sec. 2.3, we introduce you to the Jordan canonical form. This is made up of 'near diagonal' blocks of matrices related to the eigenvalues of a linear operator (or matrix). In fact, the strength of this form is that it is easy to apply, and it can be used to immediately determine whether two matrices are similar or not. As you will see in the next unit, this way of presenting a matrix also has several applications in mathematics and other areas.

Objectives

After studying this unit, you should be able to

- prove, and apply, the primary decomposition theorem;
- obtain the Jordan canonical form of a linear operator or a square matrix;
- use the Jordan form to determine whether two operators/matrices are similar.

2.2 PRIMARY DECOMPOSITION

Throughout this section, unless mentioned otherwise, we will take V to be an n-dimensional vector space over **C**.

Let λ be an eigenvalue of a linear operator $T \in L(V)$. Let $E_T(\lambda)$ be the **eigenspace of T associated with** λ, i.e., $E_T(\lambda) = \text{Ker}(T - \lambda I) = \{v \in V \mid (T - \lambda I)(v) = \mathbf{0}\}$.

Now, notice that if $v \in E_T(\lambda)$, then $(T - \lambda I)(Tv) = T(T - \lambda I)(v) = \mathbf{0}$, and so $Tv \in E_T(\lambda)$. This means that **if** $\mathbf{v \in E_T(\lambda)}$ **then** $\mathbf{Tv \in E_T(\lambda)}$. This property of T is more generally defined below.

Definition: For $T \in L(V)$, a subspace W of V is called **T-invariant** if $Tw \in W$ $\forall w \in W$, that is, $T(W) \subseteq W$.

Can you think of such a vector space? You can check that the zero subspace of V, and the full space V, are T-invariant subspaces for any $T \in L(V)$. Another example is given below.

Example 1: Let $V = \{\alpha_0 + \alpha_1 t + \cdots + \alpha_5 t^5 \mid \alpha_i \in \mathbf{Q}, i = 0, 1, \ldots, 5\}$.
Let D be the differential operator on V. Show that

i) W, the subspace consisting of all polynomials in V of degree at most 3, is D-invariant;

ii) U, the subspace of V consisting of all polynomials in V with constant term 0, is not D-invariant.

Solution: i) Since the derivative of a polynomial of degree at most 3 is a polynomial of degree at most 2, W is D-invariant.

ii) V is not D-invariant because, for example, the polynomial $t \in U$ and $Dt = 1 \notin U$.

We now define a concept that gives us generic examples of T-invariance.

Definition: Let $T \in L(V)$, λ be an eigenvalue of T and r be a positive integer. Let $E_T^r(\lambda) = \text{Ker}(T - \lambda I)^r$. Then the subspace $E_T^r(\lambda)$ is called the **generalized eigenspace of T** associated with λ. The non-zero elements of $E_T^r(\lambda)$ are called **generalized eigenvectors** of T corresponding to λ. An element $x \in E_T^r(\lambda) \setminus E_T^{r-1}(\lambda)$ is called a **generalized eigenvector of T of order r corresponding to λ**.

Note: 1) $E_T^r(\lambda)$ is T-invariant $\forall\, r > 0$.

2) $E_T^r(\lambda) \subseteq E_T^{r+1}(\lambda)\ \forall\, r > 0$.

3) If x is a generalized eigenvector of T of order r corresponding to λ, then $(T - \lambda I)(x)$ is a generalized eigenvector of order $r - 1$ (if $r > 1$).

Let us consider some examples of generalized eigenspaces.

Example 2: For D and V as in Example 1, obtain $E_D(0)$ and $E_D^2(0)$.

Solution: $E_D(0) = \text{Ker}\, D = \{\alpha_0 + \alpha_1 t + \cdots + \alpha_5 t^5 \mid \alpha_i \in \mathbf{Q}, \alpha_1 + \alpha_2 t + \cdots + 5\alpha_5 t^4 = 0\}$
$= \{\alpha_0 \mid \alpha_0 \in \mathbf{Q}\} = \mathbf{Q}$.

$E_D^2(0) = \text{Ker}\, D^2 = \{\alpha_0 + \cdots + \alpha_5 t^5 \mid \alpha_i \in \mathbf{Q}, 2\alpha_2 + 6\alpha_3 t + 12\alpha_4 t^2 + 20\alpha_5 t^3 = 0\}$
$= \{\alpha_0 + \alpha_1 t \mid \alpha_0, \alpha_1 \in \mathbf{Q}\}$.

You can see that $E_D(0) \subseteq E_D^2(0)$.

Why not try some related exercises now!

E1) For T given in E3, Unit 1, obtain $E_T^r(\lambda)$, $r = 1, 2, 3, 4$, for all its eigenvalues λ.

E2) Prove (1), (2) and (3) of the note given above.

Now, let W be a T-invariant subspace of V. Then $T(w) \in W\ \forall\, w \in W$. Let T' be the restriction of the linear operator T to W, that is, $T'(w) = T(w)\ \forall\, w \in W$.

Then $T':W \to W$, i.e., $T' \in L(W)$.

Let $B' = \{w_1, \ldots, w_m\}$ be a fixed ordered basis of W. We can extend B' to form a basis B of V. Let $B = B' \cup \{v_{m+1}, \ldots, v_n\}$.

Since W is T- invariant, $T(w_j) \in W \; \forall \, j = 1, \ldots, m$. Let $T(w_j) = \alpha_{1j}w_1 + \cdots + \alpha_{mj}w_m$ $\forall \, j = 1, \cdots, m$, and $T(v_j) = \beta_{1j}w_1 + \cdots + \beta_{mj}w_m + \gamma_{m+1j}v_{m+1} + \cdots + \gamma_{nj}v_n$, $\forall j = m+1, \cdots, n$, where $\alpha_{rs}, \beta_{rs}, \gamma_{rs}$ are scalars.

So,

$$[T]_B = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1m} & \beta_{1m+1} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} & \beta_{mm+1} & \cdots & \beta_{mn} \\ 0 & \cdots & 0 & \gamma_{m+1m+1} & \cdots & \gamma_{m+1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_{nm+1} & \cdots & \gamma_{nn} \end{bmatrix} \qquad \ldots (1)$$

Note that the matrix $\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{bmatrix}$ is actually $[T']_{B'}$. The representation above of $[T]_B$ gives us important information, which we state as the following theorem.

Theorem 1: Let $T \in L(V)$, W be a T-invariant subspace of V, and T' be the restriction of T to W. Then $T' \in L(W)$, and the characteristic polynomial of T' divides the characteristic polynomial of T.

Proof: We have already seen why $T' \in L(W)$. Let us now prove the second part. The characteristic polynomial of T is $C_T(t) = \det(tI_n - [T]_B)$

$$= \det \begin{bmatrix} t-\alpha_{11} & \cdots & -\alpha_{1m} & -\beta_{1m+1} & \cdots & -\beta_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\alpha_{m1} & \cdots & t-\alpha_{mm} & -\beta_{mm+1} & \cdots & -\beta_{mn} \\ 0 & \cdots & 0 & t-\gamma_{m+1m+1} & \cdots & -\gamma_{m+1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -\gamma_{nm+1} & \cdots & t-\gamma_{nn} \end{bmatrix}$$

$$= \det \begin{bmatrix} t-\alpha_{11} & \cdots & -\alpha_{1m} \\ \vdots & \ddots & \vdots \\ -\alpha_{m1} & \cdots & t-\alpha_{mm} \end{bmatrix} \det \begin{bmatrix} t-\gamma_{m+1m+1} & \cdots & -\gamma_{m+1n} \\ \vdots & \ddots & \vdots \\ -\gamma_{nm+1} & \cdots & t-\gamma_{nn} \end{bmatrix}$$

$$= \det(tI_m - [T']_{B'}) \det(tI_{n-m} - C), \text{ where } C = \begin{bmatrix} \gamma_{m+1\,m+1} & \cdots & \gamma_{m+1n} \\ \vdots & \ddots & \vdots \\ \gamma_{nm+1} & \cdots & \gamma_{nn} \end{bmatrix}$$

$$= C_{T'}(t) \det(tI_{n-m} - C).$$

We get a better result than Theorem 1 if V is a direct sum of two T-invariant subspaces.

Theorem 2: Let T be a linear operator on V. If W_1 and W_2 are T-invariant subspaces of V such that $V = W_1 \oplus W_2$, and if T_i is the restriction of T to W_i, $i = 1, 2$, then the characteristic polynomial of T is the product of the characteristic polynomials of T_1 and T_2, i.e., $C_T(t) = C_{T_1}(t) C_{T_2}(t)$.

Proof: Let $B_1 = \{x_1, \ldots, x_r\}$ and $B_2 = \{y_1, \ldots, y_s\}$ be ordered bases of W_1 and W_2, respectively. Since $V = W_1 \oplus W_2$, $B = B_1 \cup B_2$ is a basis of V. As the subspaces W_1 and W_2 are T-invariant, we have

$T(x_j) = \alpha_{1j}x_1 + \cdots + \alpha_{rj}x_r$, $\alpha_{ij} \in F$ for $j = 1, \ldots, r$, and

$T(y_j) = \beta_{1j}y_1 + \cdots + \beta_{sj}y_s$, $\beta_{ij} \in F$, for $j = 1, \ldots, s$.

Also, by considering the definitions of T_1 and T_2, we see that

$$[T]_B = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1r} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \beta_{11} & \cdots & \beta_{1s} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta_{s1} & \cdots & \beta_{ss} \end{bmatrix} = \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{bmatrix},$$

where $C = [T_1]_{B_1}$ and $D = [T_2]_{B_2}$. Hence

$$\begin{aligned} C_T(t) &= \det(tI_n - [T]_B) \\ &= \det(tI_r - C)\det(tI_s - D) \\ &= C_{T_1}(t)\, C_{T_2}(t). \end{aligned}$$

The advantage of Theorem 2 is that we can breakdown the study of a linear operator on a space of a large dimension to a study of linear operators on spaces of smaller dimensions. Further, if the T-invariant subspaces are easy to handle, then we have a better representation for linear operators. We see this in the next theorem.

Theorem 3 (Primary decomposition): Let T be a linear operator on V, with its minimal polynomial given by $m_T(t) = (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$, where $\lambda_1, \ldots, \lambda_k$ are distinct and $m_1, \ldots, m_k$ are positive integers.
Then

i) $V = E_T^{m_1}(\lambda_1) \oplus \cdots \oplus E_T^{m_k}(\lambda_k)$, a direct sum of generalized eigenspaces of T.

ii) If, for each i, T_i is the restriction of T to $E_T^{m_i}(\lambda_i)$, then the minimal polynomial of T_i is $(t-\lambda_i)^{m_i}$.

Proof: i) Firstly, recall that $E_T^{m_i}(\lambda_i) = (T - \lambda_i I)^{m_i}$, and $E_T^{m_i}(\lambda_i)$ is a T-invariant non-zero subspace of V for all $i = 1, \ldots, k$.

Now, writing $[m_T(t)]^{-1}$ as partial fractions, we have

$\dfrac{1}{m_T(t)} = \dfrac{p_1(t)}{(t-\lambda_1)^{m_1}} + \cdots + \dfrac{p_k(t)}{(t-\lambda_k)^{m_k}}$, where $p_i(t)$ is a polynomial in t of degree at most $m_i - 1$ $\forall$ $i = 1, \ldots, k$. Therefore,

$$1 = p_1(t)\frac{m_T(t)}{(t-\lambda_1)^{m_1}} + \cdots + p_k(t)\frac{m_T(t)}{(t-\lambda_k)^{m_k}}.$$

For each $i = 1, \ldots, k$, we write $q_i(t) = \dfrac{m_T(T)}{(t-\lambda_i)^{m_i}}$.

Then $q_i(t)$ is a polynomial for each i, and

$1 = p_1(t)q_1(t) + \cdots + p_k(t)q_k(t)$

So, $I = p_1(T)q_1(T) + \cdots + p_k(T)q_k(T)$. ...(2)

Thus, for every $v \in V$, we have

$v = p_1(T)q_1(T)(v) + \cdots + p_k(T)q_k(T)(v)$.

If we write $w_i = p_i(T)q_i(T)(v)$, then

$(T - \lambda_i I)^{m_i}(w_i) = (T - \lambda_i I)^{m_i} p_i(T)q_i(T)(v) = p_i(T)(T - \lambda_i I)^{m_i} q_i(T)(v)$

$= p_i(T)m_T(T)(v) = \mathbf{0}$.

So, $w_i \in E_T^{m_i}(\lambda_i)\ \forall i = 1, \cdots, k$.

Therefore, $V = E_T^{m_1}(\lambda_1) + \cdots + E_T^{m_k}(\lambda_k)$. ...(3)

Next, we need to see whether this sum is a direct sum or not. For this, assume $u_1 + \cdots + u_k = \mathbf{0},\ u_i \in E_T^{m_i}(\lambda_i)$. Suppose some of these u's are non-zero, say $u_j \neq \mathbf{0}$.

Then, $u_j = -(u_1 + \cdots + u_{j-1} + u_{j+1} + \cdots + u_k)$. ...(4)

Next, $m_T(T) = (T - \lambda_j I)^{m_j} q_j(T)$, where

$q_j(T) = (T - \lambda_1 I)^{m_1} \cdots (T - \lambda_{j-1} I)^{m_{j-1}} (T - \lambda_{j+1} I)^{m_{j+1}} \cdots (T - \lambda_k I)^{m_k}$.

So, from (2) it follows that there are polynomials $r(t)$ and $s(t)$ such that

$I = r(T)q_j(T) + s(T)(T - \lambda_j I)^{m_j}$.

Hence, from (4) it follows that

$u_j = r(T)q_j(T)u_j + s(T)(T - \lambda_j I)^{m_j} u_j = \mathbf{0}$.

This, together with (3) above, tells us that

$V = E_T^{m_1}(\lambda_1) \oplus \cdots \oplus E_T^{m_k}(\lambda_k)$

ii) Now, $(T - \lambda_i I)^{m_i} u = \mathbf{0}\ \ \forall\ u \in E_T^{m_i}(\lambda_i)$ and $i = 1, \ldots, k$. So, $(T_i - \lambda_i I)^{m_i} = \mathbf{0}$.

So, the minimal polynomial of T_i is $(t - \lambda_i)^r$ for some $r \leq m_i$. We now show that $r = m_i$. If, for some $r < m_i$, the minimal polynomial of T_i is $(t - \lambda_i)^r$, then $(T_i - \lambda_i I)^r = \mathbf{0}$. So, for every $u_i \in E_T^{m_i}(\lambda_i), (T_i - \lambda_i I)^r\ u_i = \mathbf{0}$. But then the minimal polynomial of T will be $(t - \lambda_1)^{m_1} \cdots (t - \lambda_{i-1})^{m_{i-1}} (t - \lambda_i)^r (t - \lambda_{i+1})^{m_{i+1}} \cdots (t - \lambda_k)^{m_k}$ with $r < m_i$, a contradiction.

Hence, $(t - \lambda_i)^{m_i}$ is the minimal polynomial of T_i.

What does Theorem 3 really say? It states that if the minimal polynomial of T is $(t - \lambda_1)^{m_1} \ldots (t - \lambda_k)^{m_k}$, for $m_i \in \mathbf{Z}^+$ and $\lambda_1, \ldots, \lambda_k$ distinct eigenvalues of T, then V is a direct sum of generalized eigenspaces, i.e.,

$V = V_1 \oplus \cdots \oplus V_k$, where $V_i = E_T^{m_i}(\lambda_i)$ for $i = 1, \ldots, k$.

We have seen that each V_i is T-invariant. Let B_i be a basis of V_i. The elements of B_i are generalized eigenvectors of T, and $B = \bigcup_{i=1}^{k} B_i$ is a basis of V. It follows from the primary decomposition theorem that V has a basis consisting of generalised eigenvectors. Since each V_i is T-invariant, a repeated use of Theorem 3 gives

$$[T]_B = \begin{bmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{bmatrix}, \text{ where each } A_i = [T_i]_{B_i}, \text{ and } T_i \text{ is the restriction of } T \text{ to } V_i.$$

In fact, A_i is an $n_i \times n_i$ matrix, with $n_i = \dim V_i$, that is, n_i is the algebraic multiplicity of λ_i, which you studied about in Section 1.4.

Translating Theorem 3 to matrices gives us:

Theorem 4: If A is an $n \times n$ matrix with characteristic polynomial $(t-\lambda_1)^{n_1} \ldots (t-\lambda_k)^{n_k}$ and minimal polynomial $(t-\lambda_1)^{m_1} \ldots (t-\lambda_k)^{m_k}$, then A is similar to a block diagonal matrix of the form $\begin{bmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{bmatrix}$, with each A_i an $n_i \times n_i$ matrix, and $A_i - \lambda_i I_{n_i}$ being a nilpotent matrix with nilpotency index m_i.

An important consequence of Theorem 3 is the following characterisation of diagonalisable operators.

Theorem 5: Let $T \in L(V)$ and $\lambda_1, \ldots, \lambda_k$ be all the distinct eigenvalues of T. Then the following are equivalent:

i) T is diagonalisable.
ii) The minimal polynomial of T is $(t-\lambda_1)\ldots(t-\lambda_k)$.
iii) $V = E_T(\lambda_1) \oplus \cdots \oplus E_T(\lambda_k)$, a direct sum of eigenspaces.

Proof ((i) ⇒ (ii)): Let T be diagonalisable. Then there is a basis B such that

$$[T]_B = \begin{bmatrix} \overbrace{\lambda_1 \; \ddots \; \lambda_1}^{n_1 \text{ times}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \overbrace{\lambda_k \; \ddots \; \lambda_k}^{n_k \text{ times}} \end{bmatrix}_{n \times n}, \text{ where } n_i \text{ is the algebraic multiplicity}$$

of λ_i.

Then $([T]_B - \lambda_1 I_n) \cdots ([T]_B - \lambda_k I_n)$

$$= \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & 0 & & & & \\ & & & \lambda_2 & & & \\ & & & & \ddots & & \\ & & & & & \lambda_2 & \\ & & & & & & \ddots \\ & \mathbf{0} & & & & & & \lambda_k \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_k \end{bmatrix} \cdots \begin{bmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_{k-1} & & \\ & & & & & \ddots & \\ & & & & & & \lambda_{k-1} \\ & \mathbf{0} & & & & & & 0 \\ & & & & & & & & \ddots \\ & & & & & & & & & 0 \end{bmatrix}$$

$= \mathbf{0}$, that is,

$(T - \lambda_1 I) \cdots (T - \lambda_k I) = \mathbf{0}$.

Hence, the minimal polynomial of T is $(t - \lambda_1) \cdots (t - \lambda_k)$.

$((\mathbf{ii}) \Rightarrow (\mathbf{iii}))$: This is the primary decomposition theorem.

$((\mathbf{iii}) \Rightarrow (\mathbf{i}))$: Let B_i be a basis of $E_T(\lambda_i)$. Then the elements of B_i are eigenvectors of T corresponding to the eigenvalue $\lambda_i \forall\ i = 1, \ldots, k$. Thus, $B = B_1 \cup \cdots \cup B_k$ is a basis of V consisting of eigenvectors of T. Hence, T is diagonalisable. ■

Let us see how we can use this result.

Example 3: Consider the matrix $A = \begin{bmatrix} 3 & -3 & 0 & -2 \\ -3 & 5 & 0 & 3 \\ -2 & 9 & 1 & 4 \\ 4 & -6 & 0 & -3 \end{bmatrix}$. Find a matrix P such that $P^{-1}AP$ is in block diagonal form, given that A has eigenvalues 1 and 2 only.

Solution: Step I: To find the least positive integer k such that $\text{rank}(A - \lambda I)^k = \text{rank}(A - \lambda I)^{k+1}$, for each eigenvalue λ of A.

For the eigenvalue $\lambda = 1$, we compute powers of $A - I$.

$$A - I = \begin{bmatrix} 2 & -3 & 0 & -2 \\ -3 & 4 & 0 & 3 \\ -2 & 9 & 0 & 4 \\ 4 & -6 & 0 & -4 \end{bmatrix},\ (A - I)^2 = \begin{bmatrix} 5 & -6 & 0 & -5 \\ -6 & 7 & 0 & 6 \\ -15 & 18 & 0 & 15 \\ 10 & -12 & 0 & -10 \end{bmatrix},$$

$$(A - I)^3 = \begin{bmatrix} 8 & -9 & 0 & -8 \\ -9 & 10 & 0 & 9 \\ -24 & 27 & 0 & 24 \\ 16 & -18 & 0 & -16 \end{bmatrix}$$

Now, $\text{rank}(A - I) = 3$, and $\text{rank}(A - I)^2 = 2 = \text{rank}(A - I)^3$.

We now repeat the same procedure for the eigenvalue 2.

$$A - 2I = \begin{bmatrix} 1 & -3 & 0 & -2 \\ -3 & 3 & 0 & 3 \\ -2 & 9 & -1 & 4 \\ 4 & -6 & 0 & -5 \end{bmatrix},\ \text{rank}(A - 2I) = 3$$

$$(A - 2I)^2 = \begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -11 & 0 & 1 & 7 \\ 2 & 0 & 0 & -1 \end{bmatrix},\ \text{rank}(A - 2I)^2 = 2$$

$$(A - 2I)^3 = \begin{bmatrix} -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 15 & 0 & -1 & -9 \\ -2 & 0 & 0 & 1 \end{bmatrix},\ \text{rank}(A - 2I)^3 = 2.$$

So, 2 is the least power of $A - I$ as well as $A - 2I$ for which $\text{rank}(A - I)^k = \text{rank}(A - I)^{k+1}$, and $\text{rank}(A - 2I)^k = \text{rank}(A - 2I)^{k+1}$.

Step II: Find all linearly independent solutions of the homogeneous system $(A-I)^2 X = \mathbf{0}$, that is,

$$\begin{bmatrix} 5 & -6 & 0 & -5 \\ -6 & 7 & 0 & 6 \\ -15 & 18 & 0 & 15 \\ 10 & -12 & 0 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \mathbf{0} \qquad \ldots(5)$$

Since rank $(A-I)^2 = 2$, nullity $(A-I)^2 = 4-2=2$. So this system will have 2 linearly independent solutions, and the algebraic multiplicity of the eigenvalue 1 is 2.

You should verify that $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent solutions of (5).

Similarly, we need to find linearly independent solutions of $(A-2I)^2 X = \mathbf{0}$. Since nullity $(A-2I)^2 = 2$, this system will have 2 linearly independent solutions, and the algebraic multiplicity of 2 is 2. You can check that $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ -3 \\ 2 \end{bmatrix}$ are linearly independent solutions of $(A-2I)^2 X = \mathbf{0}$.

Step III: Obtaining an invertible matrix P such that $P^{-1}AP$ is in block diagonal form. The matrix P will have columns consisting of the 4 linearly independent solution vectors obtained above. So, $P = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 2 \end{bmatrix}$. Hence

$$P^{-1}AP = \left[\begin{array}{cc:cc} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 5 & 3 \\ 0 & 0 & -3 & -1 \end{array}\right]$$

We have seen that the primary decomposition theorem helps us to transform a given $n \times n$ matrix A to the block diagonal form. Before giving more examples of this, we write the steps which are generally needed to perform this similarity transform.

Step I: For each eigenvalue λ, find the least positive integer k such that $E_A^k(\lambda) = E_A^{k+1}(\lambda)$. Equivalently, find the least positive integer k such that nullity $(A-\lambda I)^k =$ nullity $(A-\lambda I)^{k+1}$, or rank $(A-\lambda I)^k = \text{rank}(A-\lambda I)^{k+1}$. This k is the power of $(t-\lambda)$ in the minimal polynomial of A.

Step II: Solve the system $(A-\lambda I)^k X = \mathbf{0}$ for each eigenvalue λ of A. The number of linearly independent solutions of this equation is actually equal to the algebraic multiplicity of λ.

Step III: Write a matrix P whose columns are the linearly independent solutions of the linear systems obtained in Step II, corresponding to each λ. Then P will be invertible (why?), and $P^{-1}AP$ is the required block diagonal form.

Note that the columns of P are actually a basis of F^n, consisting of generalized eigenvectors.

Here are some exercises now, which will give you a chance to apply the similarity transform as above.

E3) Let $A = \begin{bmatrix} 3 & 0 & 2 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$. Find an invertible matrix P so that $P^{-1}AP$ is in block diagonal form.

(**Hint:** Note that the sum of the entries in each row is 5. Thus, in this case 5 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Use this to find other eigenvalues. In general, if the sum of the entries of each row of an $n \times n$ matrix is r, then r is an eigenvalue, and the corresponding eigenvector is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.)

E4) Let $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & 1 & 3 \end{bmatrix}$. Find an invertible matrix P such that $P^{-1}AP$ is in block diagonal form.

(**Hint:** The fourth column of A is $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix}$, that is, $Ae_4 = 3e_4$, where $e_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$. This shows that e_4 is an eigenvector corresponding to the eigenvalue 3.)

Let us now see how to transform a square matrix into a special kind of block diagonal form.

2.3 JORDAN CANONICAL FORM

Camille Jordan
(1838-1922)

Let T be a linear operator on V and let λ be an eigenvalue of T. You have just seen that there is a basis B of V such that $[T]_B$ is in block diagonal form. Now we show you that we can, in fact, choose B in such a manner that each block in $[T]_B$ is nearly diagonal, which is very convenient for calculations. Let us see how we do this.

We first start with choosing a sequence of basis vectors that help us to get the required form.

Definition: A sequence of non-zero vectors $x_1, \ldots, x_k$ in V is called a **Jordan chain of length k** associated with an eigenvalue λ of V if

$$Tx_1 = \lambda x_1$$
$$Tx_2 = \lambda x_2 + x_1$$
$$\vdots$$
$$Tx_k = \lambda x_k + x_{k-1},$$

that is, for each $i = 2,\ldots,k, (T-\lambda I)x_i = x_{i-1}$ and $(T-\lambda I)x_1 = \mathbf{0}$.

For example, for T given by A in Example 3,

$$A\begin{bmatrix}0\\0\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, A\begin{bmatrix}-1\\1\\2\\2\end{bmatrix}=\begin{bmatrix}-1\\1\\2\\2\end{bmatrix}+\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \text{ and so on.}$$

Note: $(\mathbf{T}-\lambda\mathbf{I})^{i-1}\mathbf{x}_i = \mathbf{x}_1$, that is, $\mathbf{x}_i$ **is a generalized eigenvector of T of order i.**

Let us now see how a Jordan chain is useful.

Theorem 6: A Jordan chain consists of linearly independent vectors.

Proof: Assume that $x_1,\ldots,x_k$ is a Jordan chain of $T \in L(V)$ associated with the eigenvalue λ of T. Suppose these vectors are linearly dependent, and $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$, where $\alpha_i \in F\ \forall\ i$. Let r be the largest integer such that $\alpha_r \neq 0$. Then $r>1$ (Why?).

Now, we have $x_r = \sum_{i=1}^{r-1}(-\alpha_r^{-1}\alpha_i)x_i$.

Operating $(T-\lambda I)^{r-1}$ on both sides, we get

$x_1 = (T-\lambda I)^{r-1}x_r = \sum_{i=1}^{r-1}(-\alpha_r^{-1}\alpha_i)(T-\lambda I)^{r-1}x_i = \mathbf{0}$, a contradiction, since $x_1 \neq \mathbf{0}$.

Therefore, $\{x_i\}_i$ are linearly independent. ■

You would now be in a position to appreciate the utility of a Jordan chain. If $x_1,\ldots,x_k$ is a Jordan chain of $T \in L(V)$, associated with the eigenvalue λ, then the subspace of V generated by $x_1, \ldots, x_k$, say W, is a T-invariant subspace of V, since $Tx_i = \lambda x_i + x_{i-1} \in W$. If $\hat{T}$ is the linear operator induced by T on W and $B = \{x_1,\cdots,x_k\}$, then

$$\left[\hat{T}\right]_B = \begin{bmatrix}\lambda & 1 & \mathbf{0} & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & \mathbf{0} & & \lambda\end{bmatrix}.$$

This $k\times k$ matrix is called the **Jordan block** of size k associated with λ, and we denote it by $\mathbf{J_k(\lambda)}$.

This leads us to the concept that the whole unit focusses on.

Definition: If V has a basis which is a disjoint union of Jordan chains of $T \in L(V)$, then the matrix representation of T with respect to this basis is a block diagonal matrix with Jordan blocks of various sizes along the diagonal. This basis is called **a**

Jordan basis of V, and the corresponding matrix representation is called **a Jordan canonical form** of T.

Remark: If T is a linear operator on V, where $V = E_{\lambda_1}^{r_1}(T) \oplus \cdots \oplus E_{\lambda_k}^{r_k}(T)$, then $[T]_B = A_1 \oplus \cdots \oplus A_k$, where each A_i is the Jordan block of size r_i associated with λ_i. **This representation as a Jordan matrix is unique up to permutation of the λ_is.**

These representations are dependent on whether V has a Jordan basis for every $T \in L(V)$. And, in fact, this is true, as we see from the following result.

Theorem 7 (Existence of Jordan canonical form): Let V be a vector space over F of dimension n, and let T be a linear operator on V. Let the characteristic polynomial of T be $C_T(t) = (t-\lambda_1)^{n_1} \cdots (t-\lambda_k)^{n_k}$, where $\lambda_1, \ldots, \lambda_k$ are distinct elements in F and $n_1, \ldots, n_k$ are positive integers. Then V has a Jordan basis for T.

We shall not give a proof of this result here. If you are interested in a proof, it is given in "Linear Algebra and its Applications" by G. Strang, available at your programme study centre, or in any other book on Linear Algebra. Over here we shall first give an analogue of this result for matrices, and then an algorithm for finding the Jordan canonical form of a matrix.

For matrices we have the following analogue of Theorem 7.

Theorem 8: Let $A \in \mathbf{M}_n(F)$ have characteristic polynomial $(t-\lambda_1)^{n_1} \cdots (t-\lambda_k)^{n_k}$. Then there is an invertible matrix P such that

$$P^{-1}AP = J = \begin{bmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{bmatrix}, \text{ where each } A_i \text{ is of the form } \begin{bmatrix} J_{p_1}(\lambda_i) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_{p_s}(\lambda_i) \end{bmatrix},$$

$p_1 \geq \cdots \geq p_s$.

The matrix J is called the **Jordan form** of A, and the A_is are the Jordan blocks of A.

In particular, **A is similar to its Jordan form.**

Note: We have seen in Unit 1 that not all linear operators (square matrices) are diagonalisable. Theorems 7 and 8 say that all of them are **near** diagonalisable, i.e., represented by a Jordan form.

Let us now write **an algorithm,** i.e., **a step-by-step procedure,** for finding the Jordan canonical form for a matrix A.

Step 1: For each eigenvalue λ of A, we determine the geometric and algebraic multiplicities of λ.

Step 2: The total number of Jordan blocks associated with λ equals the geometric multiplicity of λ. The sum of the sizes of the Jordan blocks associated with λ is the algebraic multiplicity of λ.

Step 3: The number of Jordan blocks of size exactly s associated with λ is $n_s = 2\,\mathrm{rank}(A-\lambda I)^s - \mathrm{rank}(A-\lambda I)^{s+1} - \mathrm{rank}(A-\lambda I)^{s-1}$.

However, for matrices of smaller sizes, we usually do not need to use Step 3. We shall see this in the following examples.

Example 4: Let A be a 5×5 matrix with characteristic polynomial $(t-1)^2(t-2)^3$, the geometric multiplicity of 1 being 1 and the geometric multiplicity of 2 being 2. Find the Jordan form of A.

Solution: There will be one Jordan block associated with the eigenvalue 1, as the geometric multiplicity of 1 is one. Similarly, as the geometric multiplicity of the eigenvalue 2 is 2, there are two Jordan blocks associated with 2.

Since the algebraic multiplicity of 1 is 2 and the algebraic multiplicity of 2 is 3, the Jordan form of A will be

$$\begin{bmatrix} J_2(1) & \vdots & & \vdots & \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ & \vdots & J_2(2) & \vdots & \\ \cdots & \vdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & J_1(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \vdots & & & \vdots & \\ 0 & 1 & & & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \vdots & 2 & 1 & \vdots & \\ & & & 0 & 2 & & \\ & \cdots & \vdots & \cdots & & \cdots & \cdots \\ & & \vdots & & & \vdots & 2 \end{bmatrix}$$

Remark: This way of writing a matrix, only showing the elements we are interested in, is a convention. The rest of the elements in the matrix are zero.

Before doing any more examples, let us see why the algorithm given above works. For this we look into the structure of the Jordan basis B of V corresponding to the linear operator T.

Let λ be an eigenvalue of T. Since $E_T^r(\lambda) \subseteq E_T^{r+1}(\lambda)$ (see E2) and V is finite-dimensional, there exists a positive integer p such that
$\{0\} \subseteq E_T^1(\lambda) \subseteq E_T^2(\lambda) \subseteq \cdots \subseteq E_T^p(\lambda) = E_T^{p+1}(\lambda) = \ldots$
Thus, the maximal length of a Jordan chain associated with λ will be p.

Next, recall that if $\{x_1, \ldots, x_r\}$ is a Jordan chain of length r associated with λ, then each x_i is a generalised eigenvector of order i, $i = 1, \ldots, r$. Therefore, in particular, the number of Jordan chains of T associated with λ is $\dim \operatorname{Ker}(T - \lambda I)$, which is the geometric multiplicity of λ. If q_s denotes the number of Jordan chains of length at least s, then it is equal to the number of generalised eigenvectors of T of order s associated with λ, and so
$q_s = \dim \operatorname{Ker}(T - \lambda I)^s - \dim \operatorname{Ker}(T - \lambda I)^{s-1}$.
Hence the number n_s, of Jordan chains of length exactly s which are associated with λ, is given by
$n_s = q_{s+1} - q_s$
$= \dim \operatorname{Ker}(T - \lambda I)^{s+1} + \dim \operatorname{Ker}(T - \lambda I)^{s-1} - 2 \dim \operatorname{Ker}(T - \lambda I)^s$
$= \left[n - \operatorname{rank}(T - \lambda I)^{s+1}\right] + \left[n - \operatorname{rank}(T - \lambda I)^{s-1}\right] - 2\left[n - \operatorname{rank}(T - \lambda I)^s\right]$, using Sylvester's Rank-Nullity theorem.
This gives us the formula in Step 3 of the algorithm.

Let us do a few more examples now.

Example 5: Let A be a matrix whose characteristic polynomial is $(t-1)^2(t-2)^3(t+1)^4$, and the minimal polynomial is $(t-1)^2(t-2)^2(t+1)$. Write the Jordan form of A.

Solution: Here we will use the fact that if for an eigenvalue λ, r is the power of $t - \lambda$ appearing in the minimal polynomial, then the size of the largest Jordan block associated with the eigenvalue λ is r.

Now, as the size of the largest Jordan block of A associated with the eigenvalue 1 is 2, the size of the largest Jordan block associated with the eigenvalue 2 is 2 and the size of the largest Jordan block associated with the eigenvalue -1 is 1, we have the following:

- The Jordan block corresponding to 1 is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.
- The Jordan blocks corresponding to 2 are $\begin{bmatrix} 2 & 1 & \vdots & \\ 0 & 2 & \vdots & \\ \cdots & \cdots & \cdots & \cdots \\ & & \vdots & 2 \end{bmatrix}$.
- The Jordan blocks corresponding to -1 are $\begin{bmatrix} -1 & \vdots & & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & -1 & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & -1 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & & \vdots & -1 \end{bmatrix}$.

Therefore, the Jordan form for A is

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

Example 6: Find the Jordan form of the matrix $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -3 & 3 & 1 & 0 \\ -2 & 2 & 2 & 0 \\ 0 & -3 & 2 & 3 \end{bmatrix}$.

Solution: The characteristic polynomial of A is $(t-2)^3(t-3)$. (Verify this!) Next, rank $(A-2I)=3$, and so $\dim \text{Ker}(A-2I)$ is 1, that is, the geometric multiplicity of 2 is 1. Since the blocks corresponding to $\lambda = 2$ are together 3×3, the block corresponding to $\lambda = 3$ has to be 1×1. Therefore, the Jordan form of A is

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Example 7: Compute the Jordan form of $A = \begin{bmatrix} -2 & 5 & 1 & 0 \\ -2 & 4 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{bmatrix}$, and find a matrix Q such that $Q^{-1}AQ$ is the Jordan form of A.

Solution: The characteristic polynomial of A is $(x-1)^4$. (Check this!)
As $\text{rank}(A-I)=2$, the geometric multiplicity of the eigenvalue 1 is $4-2=2$. Therefore, there are two Jordan blocks in the Jordan form of A, namely, either diag $(J_2(1), J_2(1))$ or diag $(J_3(1), J_1(1))$. Further,

$$(A-I)^2 = \begin{bmatrix} -2 & 2 & 2 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{bmatrix},$$ which is non-zero, and so the minimal polynomial of A cannot be $(x-1)^2$. So, diag $(J_2(1), J_2(1))$ cannot be the Jordan form of A. Thus, the Jordan form of A is

$$J = \text{diag}(J_3(1), J_1(1)) = \left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right],$$ and $Q^{-1}AQ = J$ for some invertible matrix Q. We now explicitly find this Q.

Let $Q = [q_1 \; q_2 \; q_3 \; q_4]$.

Since $AQ = QJ$, on equating the corresponding columns we have
$Aq_1 = (AQ)e_1 = (QJ)e_1 = Qe_1 = q_1$,
$Aq_2 = (AQ)e_2 = (QJ)e_2 = Q(e_1+e_2) = q_1 + q_2$,
$Aq_3 = (AQ)e_3 = (QJ)e_3 = Q(e_2+e_3) = q_2 + q_3$,
$Aq_4 = (AQ)e_4 = (QJ)e_4 = Qe_4 = q_4$,

So, the vectors q_1 and q_4 are eigenvectors of A, and q_2 and q_3 are generalised eigenvectors of orders 2 and 3, respectively. Let us pick them accordingly now.
q_4 can be any eigenvector of A, so we take $q_4 = e_4$.
Next, q_3 is a generalised eigenvector of order 3, that is, $q_3 \in \mathbf{R}^4 \setminus \text{Ker}(A-I)^2$. To find q_3, we will first find a basis of $\text{Ker}(A-I)^2$. For this we solve the equation $(A-I)^2 x = \mathbf{0}$, that is,

$$\begin{bmatrix} -2 & 2 & 2 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus, $\text{Ker}(A-I)^2$ has a basis $\{u_1 = [1\;1\;0\;0]^t, u_2 = [1\;0\;1\;0]^t, u_3 = e_4\}$.
Now, $q_3 \in \mathbf{R}^4 \setminus \text{Ker}(A-I)^2$ if and only if q_3, u_1, u_2, u_3 are linearly independent. Thus, we may take $q_3 = e_1$. (Note that there are many choices for q_3. However, we take $q_3 = e_1$, as the computations are easy with such vectors.)

This gives

$q_2=(A-I)q_3=(A-I)e_1=[-3\ -2\ -1\ -1]^t$,

$q_1=(A-I)q_2=(A-I)^2e_1=[-2\ -1\ -1\ -1]^t$.

Hence, $Q=\begin{bmatrix}-2 & -3 & 1 & 0\\ -1 & -2 & 0 & 0\\ -1 & -1 & 0 & 0\\ -1 & -1 & 0 & 1\end{bmatrix}$

You should check that $Q^{-1}AQ=J$.

Here are some related exercises now.

E5) Find all possible Jordan canonical forms for a matrix whose characteristic polynomial is $(t-1)^3(t+1)^2$, and the minimal polynomial is $(t^2-1)^2$.

E6) Write all possible Jordan canonical forms for a 5×5 matrix having characteristic polynomial $(t-1)^3(t-2)^2$, with the geometric multiplicity of 1 being 2 and the geometric multiplicity of 2 being 1.

E7) Write all possible Jordan canonical forms of a 5×5 matrix whose minimal polynomial is $(t-1)(t-2)(t-3)$.

E8) Compute the Jordan canonical form of the matrix E3. Also find the matrix P such that $P^{-1}AP$ is in Jordan form.

E9) Show that the Jordan block of size k for the eigenvalue λ of T, $J_k(\lambda)$, is such that $(J_k(\lambda)-\lambda I_k)$ is nilpotent with nilpotency index k.

E10) What is the Jordan form of a diagonalisable matrix, and why?

Now let us see what a major reason is for the importance of the Jordan form.

Theorem 9: Two matrices are similar if and only if they have the same Jordan form, up to permutation of the eigenvalues.

Proof: Suppose A and B are similar, i.e., $B = P^{-1}AP$, for some P. Also, let $Q^{-1}AQ = J$, the Jordan form of A. Then $Q^{-1}PBP^{-1}Q = J$, so that J is the Jordan form of B also.

Conversely, suppose A and B are two matrices such that $R^{-1}AR = J = S^{-1}BS$, for some nonsingular matrices R and S. Then $SR^{-1}ARS^{-1} = B$, so that A and B are similar.

Note: What Theorem 10 says is that the set of all matrices over F splits into several 'families', such that all the matrices in a 'family' can be represented by one Jordan matrix to which each of them is similar. Further, matrices from different 'families' are not similar.

Here's a related exercise.

E11) Check whether

(i) J^t is similar to J, where J is a Jordan matrix;

(ii) A^t is similar to A, for any square matrix A.

We have now finished our discussion on the Jordan canonical form. Let us briefly go over what we covered in it.

2.4 SUMMARY

Let V be an n-dimensional vector space over F and let $T \in L(V)$. Let the characteristic polynomial of T be $C_T(t) = (t-\lambda_1)^{n_1} \cdots (t-\lambda_k)^{n_k}$, where $\lambda_1, \ldots, \lambda_k$ are distinct elements of F and $n_1, \ldots, n_k$ are positive integers.

Let the minimal polynomial of T be $m_T(t) = (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$, where $1 \le m_i \le n_i$, $i = 1, \ldots, k$. Then, in this unit you have studied the following points.

1) By the primary decomposition theorem, $V = V_1 \oplus \cdots \oplus V_k$, a direct sum of T-invariant subspaces of V, where $V_i = E_T^{m_i}(\lambda_i)$. If T_i is the linear operator induced by T on V_i, then $m_{T_i}(t) = (t-\lambda_i)^{m_i}$. Also $c_T(t) = c_{T_1}(t) \cdots c_{T_k}(t)$. Thus, $c_{T_i}(t) = (t-\lambda_i)^{n_i}$. In particular, dim $V_i = n_i$.

2) Each T_i has the Jordan basis B_i associated with the eigenvalue λ_i. Let $B_i = C_1 \cup \cdots \cup C_s$ be a disjoint union of Jordan chains of length $p_1, \ldots, p_s$ such that $p_1 \ge \cdots \ge p_s \ge 1$. Then

$$[T_i]_{B_i} = \begin{bmatrix} J_{p_1}(\lambda_i) & & \\ & \ddots & \\ & & J_{p_s}(\lambda_i) \end{bmatrix}, \text{ where } J_{p_r}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}, r = 1, \cdots, s.$$

3) If $B = B_1 \cup \cdots \cup B_k$, then B is a Jordan basis of V, and

$$[T]_B = \begin{bmatrix} [T_1]_{B_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & [T_k]_{B_k} \end{bmatrix}.$$

This is the matrix representation of T with respect to a Jordan basis.

2.5 SOLUTIONS / ANSWERS

E1) $[T]_B = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$

So, the eigenvalues of T are $\pm 2 - 1$.

$E_T(2) = \text{Ker}(T-2I) = \{(x, y, z) \in \mathbf{R}^3 \mid -x+y+z=0, x-3y+z=0, x+y-3z=0\}$
$= \{(2t, t, t) \mid t \in \mathbf{R}\}$

$E_T^2(2) = \text{Ker}(T-2I)^2$
$= \{(x, y, z) \in \mathbf{R}^3 \mid 3x-3y-3z=0, -3x+11y-5z=0, -3x-5y+11z=0\}$
$= \{(2t, t, t) \mid t \in \mathbf{R}\}$.

You can check that $E_T^2(2) = E_T^3(2) = \ldots$

Now, you should similarly obtain $E_T^r(1)$ and $E_T^r(-1)$ for $r = 1, 2, 3, 4$. You will find that $E_T(1) = \{(0, 0, 0)\}$, $E_T^2(1) = \{(t, t, t) \mid t \in \mathbf{R}\} = E_T^3(1) = \ldots$
$E_T(-1) = \{(-t, t, t) \mid t \in \mathbf{R}\} = E_T^2(-1) = \ldots$.

E2) (i) Let $x \in E_T^r(\lambda) = \text{Ker}(T - \lambda I)^r$

$\Rightarrow (T - \lambda I)^r(x) = \mathbf{0}$

$\Rightarrow T(T - \lambda I)^r(x) = \mathbf{0}$

$\Rightarrow (T - \lambda I)^r(Tx) = \mathbf{0}$

$\Rightarrow Tx \in E_T^r(\lambda)$

This proves that $E_T^r(\lambda)$ is T-invariant.

(ii) Let $x \in E_T^r(\lambda)$

Then $(T - \lambda I)^r(x) = \mathbf{0}$

$\Rightarrow (T - \lambda I)^{r+1}(x) = \mathbf{0}$

$\Rightarrow x \in E_T^{r+1}(\lambda)$

Hence the statement.

(iii) $x \in E_T^r(\lambda) \Rightarrow (T - \lambda I)^r(x) = \mathbf{0} \Rightarrow (T - \lambda I)^{r-1}[(T - \lambda I)(x)] = \mathbf{0}$

$\Rightarrow (T - \lambda I)(x) E_T^{r-1}(\lambda)$.

E3) The eigenvalues of A are $5, 2, 2$.

$\text{rank}(A - 5I) = 2 = \text{rank}(A - 5I)^2$

$\text{rank}(A - 2I) = 2$, $\text{rank}(A - 2I)^2 = 1 = \text{rank}(A - 2I)^3$.

Next, the linearly independent solutions if $(A - 5I)x = \mathbf{0}$ and $(A - 2I)^2 x = \mathbf{0}$ are

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\} \text{ and } \left\{\begin{bmatrix}2\\-3\\0\end{bmatrix}, \begin{bmatrix}0\\-2\\1\end{bmatrix}\right\}, \text{ respectively.}$$

$$\therefore\ P = \begin{bmatrix}1 & 2 & 0\\1 & -3 & -2\\1 & 0 & 1\end{bmatrix} \text{ gives us } P^{-1}AP = \begin{bmatrix}5 & 0 & 0\\0 & 3 & 1\\0 & -1 & 1\end{bmatrix}, \text{ which is in block diagonal}$$

form.

E4) The eigenvalues of A are 3 and 2.

$\text{rank}(A - 3I) = 3 = \text{rank}(A - 3I)^2$

$\text{rank}(A - 2I) = 3$, $\text{rank}(A - 2I)^2 = 2$, $\text{rank}(A - 2I)^3 = 1 = \text{rank}(A - 2I)^4$.

Next, the linearly independent solutions of $(A - 3I)x = \mathbf{0}$ and $(A - 2I)^3 x = \mathbf{0}$ are

$$\left\{\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\} \text{ and } \left\{\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\\-1\end{bmatrix}\right\}, \text{ respectively.}$$

$$\therefore\ P = \begin{bmatrix}0 & 0 & 0 & 1\\0 & 0 & 1 & 0\\0 & 1 & 0 & 0\\1 & 0 & 1 & -1\end{bmatrix} \text{ givens us } P^{-1}AP = \begin{bmatrix}3 & \vdots & & & \\ \cdots & \vdots & \cdots & \cdots & \cdots\\ & \vdots & 2 & 0 & 0\\ & \vdots & 1 & 3 & -1\\ & \vdots & 0 & 1 & 1\end{bmatrix}, \text{ which is in}$$

block diagonal form.

E5) The size of the largest Jordan block associated with $\lambda = 1$ or $\lambda = -1$ is 2.
Algebraic multiplicity of 1 and -1 are 3 and 2, respectively.
Now $\dim \ker(A - I)$ can be 1, 2 or 3 (by Theorem 4, Unit 1).
So the Jordan blocks corresponding to $\lambda = 1$ can be

$$J_1 = \begin{bmatrix} 1 & \vdots & & \\ \cdots & \cdots & \cdots & \cdots \\ & \vdots & 1 & 1 \\ & \vdots & 0 & 1 \end{bmatrix}, \text{ or } J_2 = \begin{bmatrix} 1 & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & 1 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & 1 \end{bmatrix}.$$

The Jordan blocks corresponding to $\lambda = -1$ can be

$$J_3 = \begin{bmatrix} 1 & \vdots & \\ \cdots & \cdots & \cdots \\ & \vdots & 1 \end{bmatrix} \text{ or } J_4 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}.$$

So the possible Jordan forms for the matrix are $\begin{bmatrix} J_i & \vdots & \\ \cdots & \cdots & \cdots \\ & \vdots & J_j \end{bmatrix}$, where $i = 1, 2$ and $j = 3, 4$.

E6) The largest Jordan block associated with 1 (or 2) is of size 2 (or 1, respectively).
The sum of the sizes of the Jordan blocks corresponding to 1 and 2 is 3 and 2, respectively.
So, the Jordan blocks corresponding to $\lambda = 1$ can be

$$J_1 = \begin{bmatrix} 1 & \vdots & & \\ \cdots & \cdots & \cdots & \cdots \\ & \vdots & 1 & 1 \\ & \vdots & 0 & 1 \end{bmatrix}, \text{ or } J_2 \begin{bmatrix} 1 & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & 1 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & 1 \end{bmatrix}.$$

The Jordan blocks corresponding to $\lambda = 2$ can be

$$J_3 = \begin{bmatrix} 2 & \vdots & \\ \cdots & \cdots & \cdots \\ & \vdots & 2 \end{bmatrix}, \text{ or } J_4 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

So, the Jordan canonical form of the matrix can be $\begin{bmatrix} J_i & \\ & J_j \end{bmatrix}$, $i = 1, 2$, $j = 3, 4$.

E7) Suppose the characteristic polynomial is $(t-1)^3(t-2)(t-3)$.
The geometric multiplicities of 1, 2, 3 are 1, 1, 1, respectively.
The algebraic multiplicities of 1, 2, 3 are 3, 1, 1, respectively.
So, the Jordan blocks corresponding to $\lambda = 1$ can only be

$$J_1 = \begin{bmatrix} 1 & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & 1 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & 1 \end{bmatrix}.$$

The Jordan block corresponding to $\lambda = 2$ can only be $J_2 = [2]$.

The Jordan block corresponding to $\lambda = 3$ can only be $J_3 = [3]$.

So, the Jordan form would be $\begin{bmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{bmatrix}$.

Similarly, you can find the Jordan forms for all the other possible characteristic polynomials.
Note that in all the cases, the Jordan matrices will be diagonal.

E8) The characteristic polynomial of A is $(t-2)^2(t-5)$.

$\dim\ker(A-2I)=1$, with basis $\left\{\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}\right\}$.

$\dim\ker(A-5I)=1$, with basis $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

So the geometric multiplicities of both eigenvalues are 1.
Therefore, there is only one Jordan block associated with each of them.

Thus, the Jordan form of A is $J=\begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

We now find P, such that $P^{-1}AP=J$.
Let $P=[p_1\ p_2\ p_3]$.

$AP=PJ \Rightarrow Ap_1=(AP)e_1=(PJ)e_1=5p_1$

$Ap_2=(PJ)e_2=2p_2$

$Ap_3=(PJ)e_3=P(e_2+2e_3)=p_2+2p_3$.

So, we can take p_1 to be an eigenvector of A corresponding to 5, say $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Similarly, we take $p_2=\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$. Now $p_3 \in \ker(A-2I)^2$ and

$(A-2I)^2=\begin{bmatrix} 3 & 2 & 4 \\ 3 & 2 & 4 \\ 3 & 2 & 4 \end{bmatrix}$. So, we take $p_3=\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$.

Then $P=\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}$ is such that $P^{-1}AP=J$.

E9) $J=J_k(\lambda)-\lambda I_k=\begin{bmatrix} 0 & 1 & & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & 0 & & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & & 0 \end{bmatrix}=\begin{bmatrix} J' & e_n \\ \mathbf{0} & 0 \end{bmatrix}$, where J' is the $(k-1)\times(k-1)$ matrix of the same form as J. Now, if $k=1$, we know that $J=0$, so that $J_1(\lambda)$ is nilpotent with nilpotency under 1.

Assume that the property holds for $J_{k-1}(\lambda)$. Then $J'^{(k-1)}=\mathbf{0}$.

Then $J^k = \begin{bmatrix} J'^k & J'^{(k-1)}e_n \\ \mathbf{0} & 0 \end{bmatrix}$.

Also, by the induction assumption $J'^k = \mathbf{0}$ and $J'^{(k-1)}e_n = \mathbf{0}$.

Therefore, $J^k = \mathbf{0}$.

Also k is the least such power, since $(k-1)$ is the least such power for J'.

Hence, by the principle of induction, we have proved the result.

E10) For a diagonalisable matrix, with eigenvalues $\lambda_1, \ldots, \lambda_n$, the Jordan form is $\begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}$. This is because the geometric multiplicity of each eigenvalue is the same as its algebraic multiplicity. And hence, the size of each Jordan block must be 1.

E11) i) J^t can be obtained from J by applying a series of row transformation $E_1, \ldots, E_s$ and the corresponding column transformation $E_1^{-1}, \ldots, E_s^{-1}$. So, $P^{-1}J\,P = J^t$, where $P = E_s \ldots E_1$.

ii) Suppose $A = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix}$, where each J_i is a Jordan block.

Then $A^t = \begin{bmatrix} J_i^t & & \\ & \ddots & \\ & & J_r^t \end{bmatrix}$, and $J_i^t = P_i^{-1}\,J_i\,P$, for some invertible P_i. Then

$P^{-1}AP = A^t$, where $P = \begin{bmatrix} P_i & & \\ & \ddots & \\ & & P_r \end{bmatrix}$. Hence A and A^t are similar.