Honours Algebra - Week 10 - The Jordan Normal Form

Antonio León Villares

March 2022

Contents

Based on the notes by Iain Gordon, Section 6.1-6.4

1 Jordan Normal Form Basics

1.1 Recap: Triangularisability

Let $f : V \to V$ be an ***endomorphism*** *of a* ***finite dimensional*** *F-vector space V.*
The following are equivalent:

1. *There exists an ordered basis:*

$$\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$$

   such that:

$$f(\underline{v}_j) = \sum_{i=1}^{j} a_{ij}\underline{v}_i, \qquad i \in [1, n]$$

   In particular, this means that ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ will be a ***triangular matrix****, with entries a_{ij}:*

$$A = {}_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

   This means that f is ***triangularisable****.*

2. *The* ***characteristic polynomial****, $\mathcal{X}_f$, decomposes into* ***linear factors*** *in $F[x]$*

[Proposition 4.6.1]

Throughout we operate over an algebraically closed field (i.e $F = \mathbb{C}$). For triangularisability, this means that any endomorphism f will be triangularisable, since the algebraic closure ensures the linear decomposition of $\mathcal{X}_f$.

- **What is a nilpotent Jordan Block?**
  - let $r \in \mathbb{N}, r \geq 1$

- define a **nilpotent Jordan Block of size** r as the matrix:

$$(J(r))_{ij} = \begin{cases} 1, & j = i+1 \\ 0, & otherwise \end{cases} \implies J(r) = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

- **What is a Jordan Block?**
  - let $r \in \mathbb{N}, r \geq 1$ and $\lambda \in F$
  - define a **Jordan Block of size** r **and eigenvalue** λ as the matrix:

$$J(r,\lambda) = \lambda I_r + J(r) \implies J(r,\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

  - notice, λI_r and $J(r,\lambda)$ **commute**

1.2 Theorem: Jordan Normal Form

*Let F be an **algebraically closed field**.*
*Let V be a **finite dimensional** vector space.*
Let:

$$\phi : V \to V$$

*be an **endomorphism** with **characteristic polynomial:***

$$\mathcal{X}_\phi(x) = \prod_{i=1}^{s} (x - \lambda_s)^{a_i} \in F[x]$$

*where a_i denotes the **algebraic multiplicity** of the **distinct** eigenvalues λ_i, such that if $n = dim(V)$:*

$$a_i \geq 1 \qquad and \qquad \sum_{i=1}^{s} a_i = n$$

*Then, there exists an **ordered basis** $\mathcal{B}$ of V, such that the **representing matrix** ${}_{\mathcal{B}}[\phi]_{\mathcal{B}}$ is in **Jordan Normal Form**.*
*That is, ${}_{\mathcal{B}}[\phi]_{\mathcal{B}}$ is a **block diagonal** matrix, with **Jordan Blocks** on its diagonal:*

$$ {}_{\mathcal{B}}[\phi]_{\mathcal{B}} = \left(\begin{array}{c|c|c|c|c} J(r_{11}, \lambda_1) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & J(r_{1m_1}, \lambda_1) & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & J(r_{sm_s}, \lambda_s) \end{array}\right)$$

where $r_{ij} \geq 1$ and:

$$a_i = \sum_{j=1}^{m_i} r_{ij}$$

[Theorem 6.2.2]

1.3 Jordan Normal Form and Triangularisation

A matrix in ***Jordan Normal Form*** *is a special case of* ***upper triangular matrix****, with the restriction:*

$$a_{ij} = \begin{cases} 0 \textit{ or } 1, & i = j-1 \\ 0, & i < j-1 \end{cases}$$

such that for a given a basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$*:*

$$\begin{aligned} \phi(\underline{v}_1) &= a_{11}\underline{v}_1 \\ \phi(\underline{v}_2) &= a_{12}\underline{v}_1 + a_{22}\underline{v}_2 \\ &\vdots \\ \phi(\underline{v}_n) &= a_{(n-1)n}\underline{v}_{n-1} + a_{nn}\underline{v}_n \end{aligned}$$

1.4 Jordan Block and Endomorphism Properties

For Jordan Blocks, the above can be more specific. Given a basis $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$*, and endomorphism* $f : V \to V$ *with eigenvalue* $\lambda \in F$ *satisfies:*

$$\begin{aligned} f(\underline{v}_1) &= \lambda \underline{v}_1 \\ f(\underline{v}_2) &= \underline{v}_1 + \lambda \underline{v}_2 \\ &\vdots \\ f(\underline{v}_n) &= \underline{v}_{n-1} + \lambda \underline{v}_n \end{aligned}$$

so that:

$$_{\mathcal{B}}[f]_{\mathcal{B}} = J(r, \lambda)$$

Moreover, if we define an ***endomorphism****:*

$$e = f - \lambda \, \mathrm{id}_V \qquad e(\underline{v}) = f(\underline{v}) - \lambda \underline{v}$$

then in particular:

$$e(\underline{v}_i) = f(\underline{v}_i) - \lambda \underline{v}_i = (\underline{v}_{i-1} + \lambda \underline{v}_i) - \lambda \underline{v}_i = \underline{v}_{i-1}$$

The endomorphism e is quite interesting (and useful, as we will see below). In particular:

- $e^r = 0$
- $e^j \neq 0, j \in [1, r-1]$
- $V_j = ker(e^j) = \langle \underline{v}_i, \ldots, \underline{v}_j \rangle$
- $f(V_j) \subseteq V_j$

2 Proving Jordan Normal Form

Most of this will be directly copied from the notes provided by the course: the proof is long and tedious, and to be fair, it is pretty well explained and I don't think I can add much insight. I'll still try to add some stylistic adaptations/explanations where needed.

2.1 Intuition for Jordan Normal Form

2.1.1 Step 1: Decomposing Vector Space by Using Terms in Characteristic Polynomial

Step 1: The first step is to decompose the vector space V into a direct sum $V = \oplus_{i=1}^{s} V_i$ according to the factorization of the characteristic polynomial as a product of linear factors

$$\chi_\phi(x) \;=\; (x - \lambda_1)^{a_1}(x - \lambda_2)^{a_2} \ldots (x - \lambda_s)^{a_s} \in F[x]$$

for distinct scalars $\lambda_1, \lambda_2, \ldots, \lambda_s \in F$, where for each i:

- $V_i = \ker\left((\phi - \lambda_i \,\mathrm{id}_V)^{a_i} : V \to V\right) \subseteq V$, and
- $\phi(V_i) \subseteq V_i$, and
- $(\phi - \lambda_i id_{V_i})^{m_i}$ is zero on V_i for m_i large enough.

This behaviour is an example of a general phenomenon from module theory called the Krull-Remak-Schmidt Decomposition.

2.1.2 Step 2: Analysing Terms in Vector Space Decomposition

Step 2: The outcome of the first step is to focus attention on the **individual spaces** V_i instead of V.

These spaces have the advantage that a **power** of the endomorphism $(\phi - \lambda_i id_{V_i}) : V_i \to V_i$ is **zero**: in other words:

$$\psi := \phi - \lambda_i id_{V_i}$$

is a **nilpotent linear mapping** on V_i. I already showed you this situation in Exercise 39:

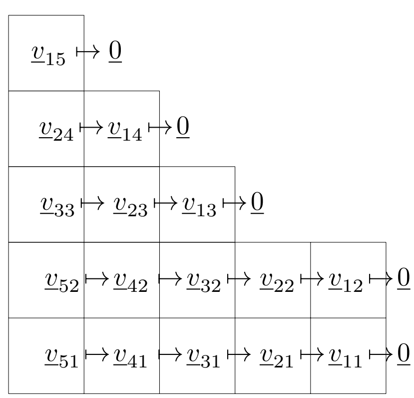
> *An **endomorphism** $f : V \to V$ of an f-vector space is **nilpotent** if and only if $\exists d \in \mathbb{N} : f^d = 0$. Let f be **nilpotent**.*
> *Show that the vector space V has an **ordered basis** $\mathcal{A}$ such that the **representing matrix** ${}_{\mathcal{A}}[f]_{\mathcal{A}}$ is **upper triangular**, and with 0s along the main diagonal.*
> *Show that any $n \times n$ matrix M that is **upper triangular** with 0s along the main diagonal satisfies $M^n = 0$.*

The proof will study a finite dimensional vector space W together with a **nilpotent** endomorphism

$$\psi : W \to W$$

I will show that there is an ordered basis of W, written $\{\underline{v}_{11}, \underline{v}_{21}, \underline{v}_{31}, \ldots, \underline{v}_{12}, \underline{v}_{22}, \underline{v}_{32}, \ldots\}$ such that the matrix of ψ with respect to this basis is **block diagonal** with **nilpotent** Jordan blocks of various sizes along the diagonal.

In my head, I picture such a basis together with ψ as follows:



This picture would describe an example where $\dim W = 16$ because each of the 16 boxes represents one basis vector; the mapping ψ moves from left to right through the boxes, vanishing when it reaches the outer edge of the diagram. In the example the matrix would have the form $\mathrm{diag}(J(5), J(5), J(3), J(2), J(1))$.

The vector space W has ordered basis (partitioned to match the Jordan blocks)

$$\begin{aligned}\mathcal{B} &= \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4 \cup \mathcal{B}_5 \\ &= (\underline{v}_{11}, \underline{v}_{21}, \underline{v}_{31}, \underline{v}_{41}, \underline{v}_{51}) \cup (\underline{v}_{12}, \underline{v}_{22}, \underline{v}_{32}, \underline{v}_{42}, \underline{v}_{52}) \cup (\underline{v}_{13}, \underline{v}_{23}, \underline{v}_{33}) \cup (\underline{v}_{14}, \underline{v}_{24}) \cup (\underline{v}_{15})\end{aligned}$$

and $\psi : W \to W$ is given by

$$\begin{aligned}
&\psi(\underline{v}_{11}) = \underline{0}\ ,\ \psi(\underline{v}_{21}) = \underline{v}_{11}\ ,\ \psi(\underline{v}_{31}) = \underline{v}_{21}\ ,\ \psi(\underline{v}_{41}) = \underline{v}_{31}\ ,\ \psi(\underline{v}_{51}) = \underline{v}_{41}\ ,\\
&\psi(\underline{v}_{12}) = \underline{0}\ ,\ \psi(\underline{v}_{22}) = \underline{v}_{12}\ ,\ \psi(\underline{v}_{32}) = \underline{v}_{22}\ ,\ \psi(\underline{v}_{42}) = \underline{v}_{32}\ ,\ \psi(\underline{v}_{52}) = \underline{v}_{42}\ ,\\
&\psi(\underline{v}_{13}) = \underline{0}\ ,\ \psi(\underline{v}_{23}) = \underline{v}_{13}\ ,\ \psi(\underline{v}_{33}) = \underline{v}_{21}\ ,\\
&\psi(\underline{v}_{14}) = \underline{0}\ ,\ \psi(\underline{v}_{24}) = \underline{v}_{14}\ ,\\
&\psi(\underline{v}_{15}) = \underline{0}\ .
\end{aligned}$$

Define an increasing sequence of subspaces

$$W_0 = \{0\} \subset W_1 \subset W_2 \subset W_3 \subset W_4 \subset W_5 = W$$

with

$$W_k = \ker(\psi^k : W \to W) = \{\underline{w} \in W \,|\, \psi^k(\underline{w}) = \underline{0}\}\ .$$

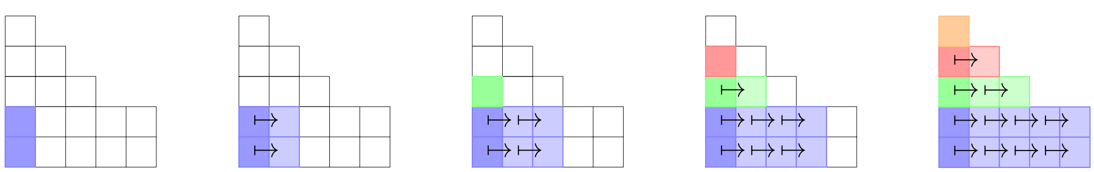
the kernel of the k-fold iteration of ψ

$$\psi^k = \psi \circ \psi \circ \cdots \circ \psi : W \to W\ .$$

Thus W_k is spanned by the basis elements $\underline{v}_{ij} \in \mathcal{B}$ with $\psi^k(\underline{v}_{ij}) = \underline{0} \in W$, and

$$\begin{aligned}
W_1 &= \langle \underline{v}_{11}, \underline{v}_{12}, \underline{v}_{13}, \underline{v}_{14}, \underline{v}_{15} \rangle\ ,\\
W_2 &= \langle \underline{v}_{24}, \underline{v}_{23}, \underline{v}_{22}, \underline{v}_{21} \rangle \oplus W_1\ ,\\
W_3 &= \langle \underline{v}_{33}, \underline{v}_{32}, \underline{v}_{31} \rangle \oplus W_2\ ,\\
W_4 &= \langle \underline{v}_{42}, \underline{v}_{41} \rangle \oplus W_3\ ,\\
W_5 &= \langle \underline{v}_{52}, \underline{v}_{51} \rangle \oplus W_4\ .
\end{aligned}$$

The subspaces are most easily illustrated by the coloured sequence of boxes in the diagram:



The first diagram on the left has W_4 in white, the next one has W_3 in white, the next one W_2, the next W_1, and the final one $\{0\}$. The coloured boxes produce a basis for the quotient vector spaces

$$\begin{aligned}
W/W_4 &= \langle W_4 + \underline{v}_{52}, W_4 + \underline{v}_{51} \rangle \ , \\
W/W_3 &= \langle W_3 + \underline{v}_{52}, W_3 + \underline{v}_{51}, W_3 + \underline{v}_{41}, W_3 + \underline{v}_{42} \rangle \ , \\
W/W_2 &= \langle W_2 + \underline{v}_{52}, W_2 + \underline{v}_{51}, W_2 + \underline{v}_{41}, W_2 + \underline{v}_{42}, W_2 + \underline{v}_{33}, W_2 + \underline{v}_{32}, W_2 + \underline{v}_{31} \rangle \ , \\
W/W_1 &= \langle W_1 + \underline{v}_{52}, W_1 + \underline{v}_{51}, W_1 + \underline{v}_{41}, W_1 + \underline{v}_{42}, W_1 + \underline{v}_{33}, W_1 + \underline{v}_{32}, W_1 + \underline{v}_{31}, \\
&\qquad W_1 + \underline{v}_{24}, W_1 + \underline{v}_{23}, W_1 + \underline{v}_{22}, W_1 + \underline{v}_{21} \rangle \ .
\end{aligned}$$

The more darkly coloured boxes (all on the left column) are "generators" from which all other basis vectors (more lightly coloured with the same colour) are produced using the mapping ψ.

2.2 Proof of Jordan Normal Form: Step 1

Let $\phi : V \to V$ be an endomorphism of the finite dimensional F-vector space V. Since F is algebraically closed, the characteristic polynomial $\chi_\phi(x)$ decomposes into linear factors by Theorem 3.3.14. I write it as follows

$$\chi_\phi(x) = \prod_{i=1}^{s} (x - \lambda_i)^{a_i} \in F[x]$$

where each a_i is a positive integer, $\lambda_i \neq \lambda_j$ for $i \neq j$, and the λ_i are the eigenvalues of ϕ. For $1 \leqslant j \leqslant s$ define

$$P_j(x) = \prod_{\substack{i=1 \\ i \neq j}}^{s} (x - \lambda_i)^{a_i}$$

2.2.1 Lemma: Polynomial Sum

There exists polynomials $Q_j(x) \in F[x]$ such that

$$\sum_{j=1}^{s} P_j(x) Q_j(x) = 1.$$

[Lemma 6.3.1]

Proof. This is an application of the extended Euclidean algorithm for $F[x]$, based on Theorem 3.3.4. This algorithm computes the highest common factor of a set of polynomials in terms of the polynomials themselves and some subsidiary polynomials $Q_j(x)$:

$$\sum_{j=1}^{s} P_j(x) Q_j(x) = \text{h.c.f.}\{P_1(x), \ldots, P_s(x)\}$$

Since the highest common factor of the set of polynomials $\{P_1(x), P_2(x), \ldots, P_s(x)\}$ is 1, the lemma follows. □

The extended Euclidean algorithm for $F[x]$ works in exactly the same way as for $\mathbb{Z}$, but using Theorem 3.3.4 here, the division algorithm for polynomials with coefficients in the field F.

2.2.2 Defining the Generalised Eigenspace

- **What is the generalised eigenspace of an endomorphism?**
  - the **generalized eigenspace** of ϕ with eigenvalue λ_i, $E^{\mathsf{gen}}(\lambda_i, \phi)$, is the following subspace of V

$$E^{\mathsf{gen}}(\lambda_i, \phi) = \{\underline{v} \in V \mid (\phi - \lambda_i \, \mathrm{id}_V)^{a_i}(\underline{v}) = \underline{0}\}$$

  - notice, the **standard eigenspace**:

$$E(\lambda_i, \phi) \; = \; \{\underline{v} \in V \mid (\phi - \lambda_i \, \mathrm{id}_V)(\underline{v}) = \underline{0}\} \; .$$

    is nothing but a **subset** of the generalised eigenspace

- **What is the algebraic multiplicity of an endomorphism?**
  - the **dimension** of $E^{\mathsf{gen}}(\lambda_i, \phi)$ is the **algebraic multiplicity** of ϕ with eigenvalue λ_i

- **What is the geometric multiplicity of an endomorphism?**
  - the **dimension** of the eigenspace $E(\lambda_i, \phi)$ is called the **geometric multiplicity of ϕ with eigenvalue λ_i**
  - notice, the **algebraic multiplicity** of λ_i is greater than its **geometric multiplicity**

2.2.3 Stable Endomorphisms

- **When is an endomorphism stable?**
  - let $f : X \to X$ be a mapping from a set X to itself.
  - a subset $Y \subseteq X$ is **stable under** f precisely when

$$f(Y) \subseteq Y$$

  - that is, if $y \in Y$ then $f(y) \in Y$.

2.2.4 Proposition: Step 1 - Direct Sum Decomposition

For each $1 \leqslant i \leqslant s$, let

$$\mathcal{B}_i \ = \ \{\underline{v}_{ij} \in V \,|\, 1 \leqslant j \leqslant a_i\}$$

*be a **basis** of $E^{\mathsf{gen}}(\lambda_i, \phi)$, where a_i is the **algebraic multiplicity** of ϕ with **eigenvalue** λ_i, such that $\sum_{i=1}^{s} a_i = n$ is the dimension of V.*

Then:

1. *Each $E^{\mathsf{gen}}(\lambda_i, \phi)$ is **stable** under ϕ.*
2. *For each $\underline{v} \in V$ there exist **unique** $\underline{v}_i \in E^{\mathsf{gen}}(\lambda_i, \phi)$ such that $\underline{v} = \sum_{i=1}^{s} \underline{v}_i$. In other words, there is a **direct sum decomposition***

$$V \ = \ \bigoplus_{i=1}^{s} E^{\mathsf{gen}}(\lambda_i, \phi)$$

with ϕ restricting to endomorphisms of the summands

$$\phi_i \ = \ \phi| \ : \ E^{\mathsf{gen}}(\lambda_i, \phi) \to E^{\mathsf{gen}}(\lambda_i, \phi) \ .$$

3. *Then*

$$\mathcal{B} \ = \ \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_s \ = \ \{\underline{v}_{ij} \,|\, 1 \leqslant i \leqslant s, 1 \leqslant j \leqslant a_i\}$$

*is a **basis** of V. The matrix of the endomorphism ϕ with respect to this basis is given by the block diagonal matrix*

$$_{\mathcal{B}}[\phi]_{\mathcal{B}} \ = \ \left(\begin{array}{c|c|c|c} B_1 & 0 & 0 & 0 \\ \hline 0 & B_2 & 0 & 0 \\ \hline 0 & 0 & \ddots & 0 \\ \hline 0 & 0 & 0 & B_s \end{array}\right) \in \mathsf{Mat}(n; F)$$

with $B_i =_{\mathcal{B}_i} [\phi_i]_{\mathcal{B}_i} \in \mathsf{Mat}(a_i; F)$.

[Proposition 6.3.5]

Proof. (1) Let $\underline{v} \in E^{\mathsf{gen}}(\lambda_i, \phi)$ so that $(\phi - \lambda_i \operatorname{id}_V)^{a_i}(\underline{v}) = \underline{0}$. Then

$$\phi(\phi - \lambda_i \operatorname{id}_V) \ = \ \phi^2 - \lambda_i \phi \ = \ (\phi - \lambda_i \operatorname{id}_V)\phi \ : \ V \to V \ ,$$

so I deduce that for all $\underline{v} \in E^{\mathsf{gen}}(\lambda_i, \phi)$

$$(\phi - \lambda_i \operatorname{id}_V)^{a_i}\phi(\underline{v}) \ = \ \phi(\phi - \lambda_i \operatorname{id}_V)^{a_i}(\underline{v}) \ = \ \phi(\underline{0}) \ = \ \underline{0} \in V \ .$$

This shows that $\phi(\underline{v}) \in E^{\mathsf{gen}}(\lambda_i, \phi)$ so that $E^{\mathsf{gen}}(\lambda_i, \phi)$ is indeed stable under ϕ.

---

(2) By (2.2.1) I have $1 = \sum_{j=1}^{s} P_j(x)Q_j(x)$ and so evaluating this at the endomorphism ϕ gives

$$\mathrm{id}_V = \sum_{j=1}^{s} P_j(\phi) \circ Q_j(\phi) \tag{1}$$

Therefore, for all $\underline{v} \in V$ I have

$$\underline{v} = \sum_{j=1}^{s} P_j(\phi) \circ Q_j(\phi)(\underline{v})$$

Now I observe that

$$(\phi - \lambda_j \,\mathrm{id}_V)^{a_j} \circ P_j(\phi) \circ Q_j(\phi)(\underline{v}) = \chi_\phi(\phi) \circ Q_j(\phi)(\underline{v}) = 0(\underline{v}) = \underline{0}$$

where I used the Cayley-Hamilton Theorem for the second equality. Setting

$$\underline{v}_j \;:=\; P_j(\phi) \circ Q_j(\phi)(\underline{v}) \in E^{\mathsf{gen}}(\lambda_j, \phi)$$

we have

$$\underline{v} \;=\; \sum_{j=1}^{s} \underline{v}_j \;,$$

demonstrating that $V = \sum_{j=1}^{s} E^{\mathsf{gen}}(\lambda_j, \phi)$.

It remains to check uniqueness in this decomposition. So suppose that $\sum_{j=1}^{s} \underline{v}_i = \sum_{i=1}^{s} \underline{w}_i$ with $\underline{v}_i, \underline{w}_i \in E^{\mathsf{gen}}(\lambda_i, \phi)$ for each i. This means that $\sum_{i=1}^{s}(\underline{v}_i - \underline{w}_i) = \underline{0}$. Given any $\underline{x}_j \in E^{\mathsf{gen}}(\lambda_j, \phi)$ I have for $k \neq j$

$$P_k(\phi)(\underline{x}_j) = \prod_{\substack{\ell=1 \\ \ell \neq k}}^{s} (\phi - \lambda_\ell \,\mathrm{id}_V)^{a_\ell}(\underline{x}_j) = \underline{0}$$

since $(\phi - \lambda_j \,\mathrm{id}_V)^{a_j}(\underline{x}_j) = \underline{0}$ and $(\phi - \lambda_j \,\mathrm{id}_V)^{a_j}$ is a factor of $P_k(\phi)$. So, on applying (1), I find

$$\underline{x}_j = \sum_{k=1}^{s} P_k(\phi) \circ Q_k(\phi)(\underline{x}_j) = P_j(\phi) \circ Q_j(\phi)(\underline{x}_j)$$

I apply this to the equality $\sum_{i=1}^{s}(\underline{v}_i - \underline{w}_i) = \underline{0}$. For each j this gives

$$\underline{0} = P_j(\phi)Q_j(\phi)\left(\sum_{i=1}^{s}(\underline{v}_i - \underline{w}_i)\right) = \sum_{i=1}^{s} P_j(\phi)Q_j(\phi)(\underline{v}_i - \underline{w}_i) = \underline{v}_j - \underline{w}_j.$$

It follows that $\underline{v}_j = \underline{w}_j$ for each j, as required.

---

(3) Since the set $\{\underline{v}_{ij} : 1 \leqslant j \leqslant a_i\}$ is a basis of $E^{\mathsf{gen}}(\lambda_i, \phi)$ for each i, it should be clear to you that the union of these bases is a basis of $\oplus_{i=1}^{s} E^{\mathsf{gen}}(\lambda_i, \phi)$. If you're not sure, it is proved in the solution to Exercise 6:

> *Given F-vector spaces $V_1, \ldots, V_n$ show that the dimension of their* ***cartesian product*** *is given by:*
>
> $$dim(V_1 \oplus \ldots \oplus V_n) = dim(V_1) + \ldots + dim(V_n)$$

Since $V = \oplus_{i=1}^s E^{\mathsf{gen}}(\lambda_i, \phi)$ by Part (2), that deals with the basis $\mathcal{B}$ of V.

What is the matrix with the respect to this basis? (I really need to take an ordered basis: I will take $\underline{v}_{11}, \underline{v}_{12}, \ldots, \underline{v}_{1n_1}, \underline{v}_{21}, \ldots, \underline{v}_{2n_2}, \ldots, \underline{v}_{sn_s}$ as the ordering.) If I calculate the matrix ${}_{\mathcal{B}}[\phi]_{\mathcal{B}}$ by the usual method of Theorem 2.3.1, I see that since $\phi(\underline{v}_{ij}) \in E^{\mathsf{gen}}(\lambda_i, \phi)$ by Part (1), $\phi(\underline{v}_{ij})$ can be expressed as a linear combination of the vectors $\underline{v}_{ij}$ where $1 \leqslant j \leqslant a_i$. Therefore the matrix is block diagonal with the i-th block having size $(a_i \times a_i)$. □

That completes the first step of the strategy. Each matrix B_i appearing in Part (3) of the Theorem in (2.2.4) represents the restriction of ϕ to $E^{\mathsf{gen}}(\lambda_i, \phi)$. This endomorphism of $E^{\mathsf{gen}}(\lambda_i, \phi)$ is special because it has the property that a power of $\phi - \lambda_i \, \mathrm{id}_{E^{\mathsf{gen}}(\lambda_i, \phi)}$ is zero.

2.2.5 Exercises (TODO)

1. **Using the Section above** (2.2.4) **show that:**

   1. **each matrix $A \in \mathsf{Mat}(n; F)$ can be written as $A = D + N$ where D is a diagonalisable matrix and N is a nilpotent matrix and $DN = ND$;**
   2. **the decomposition $A = D + N$ is unique.**

   This decomposition is called the Jordan decomposition of A; it plays a basic role in the theory of Lie algebras.

So now to the next step, studying nilpotent endomorphisms.

2.3 Proof of Jordan Normal Form: Step 2

Let W be a finite dimensional vector space and $\psi : W \to W$ an endomorphism such that some power of ψ is zero, that is $\psi^m = 0$ for some m. This should remind you of Exercise 39.

I will fix m to be **minimal**: $\psi^m = 0$ but $\psi^{m-1} \neq 0$. For $0 \leqslant i \leqslant m$ define

$$W_i = \ker(\psi^i)$$

If $\underline{w} \in W_i$ then

$$\psi^{i+1}(\underline{w}) = \psi \circ \psi^i(\underline{w}) = \psi(\underline{0}) = \underline{0}$$

so that $\underline{w} \in W_{i+1}$. It follows that

$$W_i \subseteq W_{i+1}$$

Moreover, since $\psi^0 = \mathrm{id}_W$ and $\psi^m = 0$ I also see that $W_0 = 0$ and $W_m = W$. Therefore I get a chain of subspaces

$$0 = W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_{m-1} \subseteq W_m = W$$

2.3.1 Lemma: Injective, Well-Define Mapping Between Quotient Spaces

*For each i, define a **linear mapping***

$$\psi_i : \frac{W_i}{W_{i-1}} \to \frac{W_{i-1}}{W_{i-2}}$$

by

$$\psi_i(\underline{w} + W_{i-1}) = \psi(\underline{w}) + W_{i-2}, \qquad \underline{w} \in W_i$$

*. Then ψ_i is **well-defined** and **injective**. [Lemma 6.3.6]*

Proof. Let $\underline{w}, \underline{w}' \in W_i$. First, $\psi(\underline{w}) \in W_{i-1}$ since

$$\psi^{i-1}(\psi(\underline{w})) = \psi^i(\underline{w}) = \underline{0}$$

Second, I check that the mapping is **well-defined**. That is:

$$\underline{w} + W_{i-1} = \underline{w}' + W_{i-1} \iff \psi(\underline{w}) + W_{i-2} = \psi(\underline{w}') + W_{i-2}$$

If

$$\underline{w} + W_{i-1} = \underline{w}' + W_{i-1}$$

then $\underline{w} - \underline{w}' \in W_{i-1}$. Therefore

$$\psi^{i-1}(\underline{w} - \underline{w}') = \underline{0}$$

and so

$$\underline{0} = \psi^{i-2} \circ \psi(\underline{w} - \underline{w}') = \psi^{i-2} \circ (\psi(\underline{w}) - \psi(\underline{w}'))$$

Therefore, $\psi(\underline{w}) - \psi(\underline{w}') \in W_{i-2}$ so that

$$\psi(\underline{w}) + W_{i-2} = \psi(\underline{w}') + W_{i-2}$$

This confirms that the mapping ψ_i is well-defined.

I now have to prove that ψ_i is **injective**. If

$$\psi_i(\underline{w} + W_{i-1}) = \underline{0} + W_{i-2}$$

then

$$\psi(\underline{w}) \in W_{i-2}$$

which means that

$$\underline{0} = \psi^{i-2}(\psi(\underline{w})) = \psi^{i-1}(\underline{w})$$

so that $\underline{w} \in W_{i-1}$, or, in other words, that $\underline{w} + W_{i-1} = \underline{0} + W_{i-1}$. This proves that $\ker \psi_i$ is zero and hence that ψ_i is injective. □

This result shows me that if I define

$$d_i = \dim\left(\frac{W_i}{W_{i-1}}\right) \qquad 1 \leqslant i \leqslant m$$

then $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_m$. This is because ψ will map basis elements of $\frac{W_i}{W_{i-1}}$ to basis elements of $\frac{W_{i-1}}{W_{i-2}}$. The mapping being injective means that $\left|\frac{W_i}{W_{i-1}}\right| \leq \left|\frac{W_{i-1}}{W_{i-2}}\right|$ so that $d_i \leq d_{i-1}$ or equivalently $d_i \geq d_{i+1}$.

2.3.2 Lemma: Mappings Conserving Linear Independence

To refine the above and help me to pick a good basis for W, I need a little technical lemma.

*Let $f : X \to Y$ be an **injective** linear mapping between the F-vector spaces X and Y. If $\{\underline{x}_1, \ldots, \underline{x}_t\}$ is a **linearly independent** set in X, then $\{f(\underline{x}_1), \ldots, f(\underline{x}_t)\}$ is a **linearly independent** set in Y. [Lemma 6.3.7]*

Proof. As is usual for most of the proofs of linear independence in an abstract setting, you just need to sniff the air and then follow your nose. So let $\alpha_1, \ldots, \alpha_t \in F$ be scalars. Suppose that

$$\alpha_1 f(\underline{x}_1) + \cdots + \alpha_t f(\underline{x}_t) = \underline{0}_Y$$

Then the linearity of f allows me to rewrite this equation as

$$f(\alpha_1 \underline{x}_1 + \cdots + \alpha_t \underline{x}_t) = \underline{0}_Y$$

Since f is assumed to be injective, this means that $\alpha_1 \underline{x}_1 + \cdots + \alpha_t \underline{x}_t = \underline{0}_X$. As the set $\{\underline{x}_1, \ldots, \underline{x}_t\}$ are linearly independent, this implies that $\alpha_1 = \cdots = \alpha_t = 0$. Thus $\{f(\underline{x}_1), \ldots, f(\underline{x}_t)\}$ is a linearly independent set. □

2.3.3 Algorithm for Basis Elements of Quotients

I can now develop an algorithm to construct a basis of each W_i/W_{i-1}. The algorithm goes as follows:

1. Choose an arbitrary basis for W_m/W_{m-1}, say

$$\{\underline{v}_{m,1} + W_{m-1}, \underline{v}_{m,2} + W_{m-1}, \ldots, \underline{v}_{m,d_m} + W_{m-1}\} .$$

2. Since

$$\psi_m : W_m/W_{m-1} \to W_{m-1}/W_{m-2}$$

   is injective by Lemmas 6.3.6, 6.3.7 above, this proves that

$$\{\psi(\underline{v}_{m,1}) + W_{m-2}, \psi(\underline{v}_{m,2}) + W_{m-2}, \ldots, \psi(\underline{v}_{m,d_m}) + W_{m-2}\}$$

   is a linearly independent set in W_{m-1}/W_{m-2}.
   Set

$$\underline{v}_{m-1,i} = \psi(\underline{v}_{m,i}) \qquad 1 \leqslant i \leqslant d_m$$

   .

3. Choose vectors

$$\{\underline{v}_{m-1,i} : d_m + 1 \leqslant i \leqslant d_{m-1}\}$$

   so that

$$\{\underline{v}_{m-1,i} + W_{m-2} : 1 \leqslant i \leqslant d_{m-1}\}$$

   is a basis of W_{m-1}/W_{m-2}.

4. Repeat!

Let me be explicit about what happens with a repetition.
At the i-th stage you will have chosen vectors

$$\underline{v}_{j,k} \text{ for } m+1-i \leqslant j \leqslant m, \quad 1 \leqslant k \leqslant d_j$$

so that $\{\underline{v}_{j,k} + W_{j-1} : 1 \leqslant k \leqslant d_j\}$ *is a basis of* W_j/W_{j-1}.
These vectors have the additional property that

$$\psi(\underline{v}_{j,k}) = \underline{v}_{j-1,k}, \qquad m+1-i < j \leqslant m$$

You'll then define $\underline{v}_{m-i,k} = \psi(\underline{v}_{m+1-i,k})$ *for* $1 \leqslant k \leqslant d_{m+1-i}$. *By Lemmas 6.3.6, 6.3.7*

$$\{\underline{v}_{m-i,k} + W_{m-i-1} : 1 \leqslant j \leqslant d_{m+1-i}\}$$

is a linearly independent set in W_{m-i}/W_{m-i-1}.
You now choose $\{\underline{v}_{m-i,k} : d_{m+1-i}+1 \leqslant k \leqslant d_{m-i}\}$ *so that* $\{\underline{v}_{m-i,k} + W_{m-i-1} : 1 \leqslant k \leqslant d_{m-i}\}$ *is a basis of* W_{m-i}/W_{m-i-1}.
You reach the end of the algorithm when you have completed the m-th stage: this produces a basis for $W_1/W_0 = W_1$. *Since* $W_1 = \ker(\psi)$ *all elements of this basis have the property that* $\psi(\underline{v}_{1,k}) = \underline{0}$.

2.3.4 Lemma: Algorithm Constructs Basis for W

The set of elements

$$\{\underline{v}_{j,k} : 1 \leqslant j \leqslant m, 1 \leqslant k \leqslant d_j\}$$

constructed in the algorithm above is a basis for W. *[Lemma 6.3.8]*

Proof. I check spanning first. I will show a finer statement:

For $1 \leqslant i \leqslant m$, the set of elements $\{\underline{v}_{j,k} : 1 \leqslant j \leqslant i, 1 \leqslant k \leqslant d_j\}$ spans W_i.

Of course, a statement like that is set up for a proof by induction.

① **Base Case**

It holds for $i = 1$ because $\{\underline{v}_{1,k} : 1 \leqslant k \leqslant d_1\}$ was constructed as a basis for W_1, so in particular a spanning set.

② **Inductive Hypothesis**

Assume that the finer statement holds for a given i.

③ **Inductive Step**

Let $\underline{v} \in W_{i+1}$ be an arbitrary element. Since $\{\underline{v}_{i+1,k} + W_i : 1 \leqslant k \leqslant d_{i+1}\}$ is a basis for W_{i+1}/W_i, there exist $\alpha_1, \ldots, \alpha_{d_{i+1}} \in F$ such that

$$\underline{v} + W_i = \alpha_1 \underline{v}_{i+1,1} + \cdots + \alpha_{d_{i+1}} \underline{v}_{i+1,d_{i+1}} + W_i$$

It follows that

$$\underline{v} - \alpha_1 \underline{v}_{i+1,1} - \cdots - \alpha_{d_{i+1}} \underline{v}_{i+1,d_{i+1}} \in W_i$$

By induction this element can be expressed as a linear combination of vectors from the set $\{\underline{v}_{j,k} : 1 \leqslant j \leqslant i, 1 \leqslant k \leqslant d_j\}$, and so $\underline{v}$ can be expressed as a linear combination of element of $\{\underline{v}_{j,k} : 1 \leqslant j \leqslant i+1, 1 \leqslant k \leqslant d_j\}$. This confirms the finer statement for $i+1$ and hence completes the induction.

Now I know that the set $\{\underline{v}_{j,k} : 1 \leqslant j \leqslant m, 1 \leqslant k \leqslant d_j\}$ spans $W = W_m$ and that it contains $\sum_{j=1}^m d_j$ elements. I'll now explain why $\dim W = \sum_{j=1}^m d_j$. With that fact in my pocket I can apply the **Cardinality Criterion for Bases, Part (2)**:

> *Let V be a **finitely generated** vector space. Then:*
>
> 1. - *each **linearly independent** subset $L \subset V$ has **at most** $dimV$ elements*
>   - *if $|L| = dimV$, then L is a **basis***
> 2. - *each **generating set** $E \subseteq V$ has **at least** $dimV$ elements*
>   - *if $|E| = dimV$, then E is a **basis***
>
> *[Corollary 1.6.7]*

to deduce that the set is a basis.

To calculate $\dim(W)$ I use repeatedly the general formula of Exercise 66: if M is an F-vector space and N a subspace of M then $\dim(M/N) = \dim(M) - \dim(N)$. This gives:

$$\begin{aligned}
\dim(W) = \dim(W_m) &= \dim(W_m/W_{m-1}) + \dim(W_{m-1}) \\
&= \dim(W_m/W_{m-1}) + \dim(W_{m-1}/W_{m-2}) + \dim(W_{m-2}) \\
&\vdots \\
&= \dim(W_m/W_{m-1}) + \dim(W_{m-1}/W_{m-2}) + \cdots + \dim(W_1/W_0) \\
&= \sum_{j=1}^{m} d_j.
\end{aligned}$$

□

This lemma gives me a basis of W which I will order via the ordering on subscripts $(j,k) < (j',k')$ if and only if $k < k'$ or $k = k'$ and $j < j'$. So for instance $(3,2) < (1,3)$ and $(1,3) < (2,3)$ so that $\underline{v}_{1,3}$ would appear in the list after $\underline{v}_{3,2}$ but before $\underline{v}_{2,3}$.

2.3.5 Proposition: Jordan Block from Basis

*Let $\mathcal{B}$ be the **ordered basis** of W constructed above*

$$(\underline{v}_{jk} : 1 \leqslant j \leqslant m, 1 \leqslant k \leqslant d_j)$$

Then

$$_{\mathcal{B}}[\psi]_{\mathcal{B}} = \operatorname{diag}(\underbrace{J(m), \ldots, J(m)}_{d_m \text{ times}}, \underbrace{J(m-1), \ldots, J(m-1)}_{d_{m-1}-d_m \text{ times}}, \ldots, \underbrace{J(1), \ldots, J(1)}_{d_1-d_2 \text{ times}})$$

*where $J(r)$ denotes the **nilpotent Jordan block of size** r. [Proposition 6.3.9]*

Proof. It follows from the explicit construction of the basis $\mathcal{B}$ that

$$\psi(\underline{v}_{i,j}) = \begin{cases} \underline{v}_{i-1,j} & \text{if } i > 1 \\ 0 & \text{otherwise} \end{cases}$$

This tells me that the entries of the (i, j)-th column of the matrix $_{\mathcal{B}}[\psi]_{\mathcal{B}}$ are all zero if $i = 1$ and otherwise are zero everywhere except for a 1 in the $(i-1, j)$-th row. This is the property that defines the nilpotent Jordan blocks, so I get the description I claimed. □

This completes Step 2 of the proof. Overall, we have shown that:

*"For all **nilpotent** endomorphisms there exists a basis such that the representing matrix can be written as a **block diagonal matrix** with **nilpotent Jordan blocks** along the diagonal."*

2.3.6 Exercises (TODO)

1. **Let $\psi : V \to V$ be a nilpotent endomorphism. Show that: the Jordan Normal Form of ψ is unique up to re-ordering of the nilpotent Jordan blocks. Explicitly, if $\mathcal{A}$ and $\mathcal{B}$ are bases of V such that**

$$_{\mathcal{A}}[\psi]_{\mathcal{A}} = \operatorname{diag}(J(a_1), \ldots, J(a_s)) \textbf{ and } _{\mathcal{B}}[\psi]_{\mathcal{B}} = \operatorname{diag}(J(b_1), \ldots, J(b_{s'}))$$

for some positive integers $a_1, \ldots, a_s$ and $b_1, \ldots b_{s'}$, then the multisets $\{a_1, \ldots, a_s\}$ and $\{b_1, \ldots, b_{s'}\}$ are equal.

2.4 Proof of Jordan Normal Form: Step 3

I now apply the outcome of Step 2 to each of the endomorphisms $(\phi - \lambda_i \operatorname{id}_V)$ restricted to $E^{\mathsf{gen}}(\lambda_i, \phi)$.

This means each such endomorphism can be written as a block diagonal matrix of the form stated in (2.3.5) for a suitable choice of basis.

The endomorphism $\lambda_i \operatorname{id}_V$ restricted to $E^{\mathsf{gen}}(\lambda_i, \phi)$ is of course $\lambda_i \operatorname{id}_{E^{\mathsf{gen}}(\lambda_i,\phi)}$ and so its matrix with respect to the chosen basis is just $\lambda_i I_{a_i}$. Therefore the matrix for $\phi = \lambda_i \operatorname{id}_V + (\phi - \lambda_i \operatorname{id}_V)$ is just $\lambda_i I_n$ plus the block diagonal matrix found above from (2.3.5).

In other words it is a block diagonal matrix of the form stated in (2.3.5) where I replace each $J(r)$ that appears with $J(r, \lambda_i)$. This means that each matrix $B_i \in \mathsf{Mat}(a_i; F)$ appearing in (2.2.4) has exactly the form that I'm looking for in the statement of the Jordan Normal Form Theorem.

3 Worked Examples

3.1 General Strategy From Proofs

This strategy follows both from what was proven, alongside the proofs provided. We consider $A \in Mat(n, F)$, where F is an algebraically closed field.

1. *Compute the characteristic polynomial $\mathcal{X}_A = \prod_{i=1}^{s}(x - \lambda_i)^{a_i}$*
2. *Given $P_j = \prod_{i=1, i\neq j}^{s}(x - \lambda_i)^{a_i}$, use the Euclidean Algorithm to determine $Q_j(x)$ such that*
$$1 = \sum_{j=1} P_j(x)Q_j(x)$$
   Then:
$$E^{gen}(\lambda_j, A) = col\ space(P_j(A)Q_j(A))$$
3. *For each $j \in [1, s]$:*
   (a) *Let $B = A - \lambda_j I_n$*
   (b) *We know that:*
$$\{\underline{0}\} \subseteq ker(B) \subseteq \ldots \subseteq ker(B^{a_j}) = E^{gen}(\lambda_j, A)$$
   (c) *We find a basis for each $ker(B^k)$ by applying the algorithm. Let $e_k = dim(ker(B^k))$ and set $e_0 = 0$.*
      - *Set $d_k = e_k - e_{k-1}$ and $\beta = \emptyset$*
      - *Set $j \in \mathbb{Z}$ as the largest integer with $d_j > 0$ (stop if j doesn't exist)*
      - *Let $\underline{v} \in ker(B^j) \setminus ker(B^{j-1})$ with $\underline{v} \notin \beta$*
      - *Update β via:*
$$\beta = \beta \cup \{B^{j-1}\underline{v}, \ldots, B\underline{v}, \underline{v}\}$$
      - *Set $d_i = d_i - 1$ for $1 \leq i \leq j$ and go to step 2*

If this is too obscure, this is very nicely explained.
Alternatively, the Wikipedia Entry for Jordan Normal Form is quite good.

3.2 Example from the Notes

Consider the matrix:

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

The characteristic polynomial can be computed by expanding along the third column:

$$\begin{aligned}
\mathcal{X}_A(x) = (-1-x)&\begin{vmatrix} 1-x & -1 & -1 \\ 0 & 2-x & 1 \\ 2 & -1 & -x \end{vmatrix} - 2\begin{vmatrix} 1-x & -1 & -1 \\ 0 & 2-x & 1 \\ -2 & 1 & 1 \end{vmatrix} \\
&= -(1+x)[(1-x)((2-x)(-x)+1) + 2(-1+(2-x)] - 2[(1-x)((2-x)-1) - 2(-1+(2-x)] \\
&= -(1+x)[(1-x)((2-x)(-x)+1) + 2(1-x)] + 2(1-x)(1+x) \\
&= -(1+x)(1-x)((2-x)(-x)+1) - 2(1+x)(1-x) + 2(1-x)(1+x) \\
&= -(1+x)(1-x)((2-x)(-x)+1) \\
&= -(1+x)(1-x)[x^2-2x+1] \\
&= -(1+x)(1-x)(1-x)^2 \\
&= -(1+x)(1-x)^3 \\
&= (1+x)(x-1)^3
\end{aligned}$$

Hence, we have 2 eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 1$.

$\lambda_1 = -1$ has algebraic multiplicity 1, so we expect a one dimensional generalised eigenspace, spanned by its corresponding eigenvector. We compute this:

$$(A+I_4)\underline{v}_1 = \begin{pmatrix} 2 & -1 & 0 & -1 \\ 0 & 3 & 0 & 1 \\ -2 & 1 & 0 & 1 \\ 2 & -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \underline{0} \implies \begin{pmatrix} 3v_2 + v_4 = 0 \\ 2v_1 - v_2 - v_4 = 0 \\ 2v_1 - v_2 + 2v_3 + v_4 = 0 \end{pmatrix}$$

Letting $v_2 = s$, the first equation tells us that $v_4 = -3s$. The second equation then says:

$$2v_1 - s + 3s = 0 \implies v_1 = -s$$

The third equation then says:

$$-2s - s + 2v_3 - 3s = 0 \implies v_3 = 3s$$

So it follows that that $ker(A+I_4)$ is spanned by:

$$\begin{pmatrix} -s \\ s \\ 3s \\ -3s \end{pmatrix} = s\begin{pmatrix} -1 \\ 1 \\ 3 \\ -3 \end{pmatrix}$$

$\lambda_2 = 1$ has algebraic multiplicity 3.
We work on the first eigenvalue equation:

$$(A - I_4)\underline{v}_{2,1} = \underline{0} \qquad (A - I_4)\underline{v}_{2,2} = \underline{v}_{2,1} \qquad (A - I_4)\underline{v}_{2,3} = \underline{v}_{2,2}$$

Indeed:

$$(A - I_4)\underline{v}_2 = \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ -2 & 1 & -2 & 1 \\ 2 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \underline{0} \implies \begin{pmatrix} v_2 + v_4 = 0 \\ v_1 + v_3 = 0 \end{pmatrix}$$

Letting $v_2 = s, v_1 = t$, it follows that $ker(A - I_4)$ is spanned by:

$$\begin{pmatrix} t \\ s \\ -t \\ -s \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

This is 2 dimensional, but we need a 3 dimensional generalised eigenspace.
We need to compute $(A - I_4)^2$:

$$(A - I_4)^2 = \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ -2 & 1 & -2 & 1 \\ 2 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ -2 & 1 & -2 & 1 \\ 2 & -1 & 2 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & -2 & 0 \\ 2 & 0 & 2 & 0 \\ 6 & 0 & 6 & 0 \\ -6 & 0 & -6 & 0 \end{pmatrix}$$

Thus to get the spanning vectors of $ker((A - I_4)^2)$:

$$(A - I_4)^2\underline{v}_2 = \begin{pmatrix} -2 & 0 & -2 & 0 \\ 2 & 0 & 2 & 0 \\ 6 & 0 & 6 & 0 \\ -6 & 0 & -6 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \underline{0} \implies \begin{pmatrix} v_1 + v_3 = 0 \\ v_2 = a \\ v_4 = b \end{pmatrix}$$

So letting $v_1 = s$, we have that $v_3 = -s$ so $ker((A - I_4)^2)$ is spanned by:

$$\begin{pmatrix} s \\ a \\ -s \\ b \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + a \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Notice, this time $ker((A - I_4)^2)$ has dimension 3, which is what we expected.
All the above work tells us that the resulting Jordan Normal Form will be composed of 3 Jordan Blocks:

- 1 corresponding to $\lambda_1 = -1$, which will be 1×1
- 2 corresponding to $\lambda_2 = 1$:

– 1 corresponds to $ker(A - I_4)$, which will be 2×2

– 1 corresponds to $ker((A - I_4)^2)$, which will be 1×1

To compute the Jordan Normal Form, we need a basis of 4 elements.

The first element corresponds to the first block (associated with eigenvalue $\lambda_1 = 1$). Since $dim(ker(A + I_4)) = 1$, the basis vector:

$$\underline{u} = \begin{pmatrix} -1 \\ 1 \\ 3 \\ -3 \end{pmatrix}$$

does the trick as a basis for $E^{gen}(-1, A)$.

We now focus on the more complicated case of $E^{gen}(1, A)$. To do so, we follow the algorithm. We start with basis $\beta = \emptyset$ and $d_2 = 1$

This tells us that we need to find a vector $\underline{v}$ such that:

$$(A - I_4)^2 \underline{v} = \underline{0} \qquad (A - I_4)\underline{v} \neq \underline{0}$$

This is equivalent to taking an element in $ker((A - I_4)^2)$ which isn't in $ker(A - I_4)$. We have 2 choices for this, since a general element of $ker((A - I_4)^2)$ is:

$$s \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + a \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

We can pick:

$$\underline{v}_{2,1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

The algorithm then tells us to compute:

$$(A - I_4)\underline{v}_{2,1} = \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ -2 & 1 & -2 & 1 \\ 2 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

We let:

$$\underline{v}_{1,1} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

and update β:

$$\beta = \emptyset \cup \{\underline{v}_{2,1}, \underline{v}_{1,1}\} = \{\underline{v}_{2,1}, \underline{v}_{1,1}\}$$

Now, we need to pick a vector $\underline{v}$ in $ker(A - I_4)$ which isn't in β and is linearly independent to β. Looking again at a general term in $ker(A - I_4)$:

$$t \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

and:

$$\beta = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

so for example we can pick:

$$\underline{v}_{1,2} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

This is the last step, since this gives us 3 vectors for $\lambda_2 = -1$, with basis:

$$\beta = \{\underline{v}_{2,1}, \underline{v}_{1,1}\} \cup \{\underline{v}_{1,2}\}$$

Overall, our final basis for the JNF is given by:

$$\{\underline{u}\} \cup \{\underline{v}_{2,1}, \underline{v}_{1,1}\} \cup \{\underline{v}_{1,2}\} = \{\underline{u}, \underline{v}_{2,1}, \underline{v}_{1,1}, \underline{v}_{1,2}\}$$

However, this basis needs to be ordered, for the theorem to work, according to the rules provided:

> *This lemma gives me a basis of W which I will order via the ordering on subscripts $(j, k) < (j', k')$ if and only if $k < k'$ or $k = k'$ and $j < j'$. So for instance $(3, 2) < (1, 3)$ and $(1, 3) < (2, 3)$ so that $\underline{v}_{1,3}$ would appear in the list after $\underline{v}_{3,2}$ but before $\underline{v}_{2,3}$.*

So our ordered basis will be:

$$\mathcal{B} = \{\underline{u}, \underline{v}_{1,1}, \underline{v}_{2,1}, \underline{v}_{1,2}\}$$

In particular, we claim that with the matrix $P = (\underline{u}, \underline{v}_{1,1}, \underline{v}_{2,1}, \underline{v}_{1,2})$ is such that:

$$P^{-1}AP = diag(J(-1, 1), J(2, 1), J(1, 1))$$

Indeed:

$$P = \begin{pmatrix} -1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 1 & 0 & -1 \\ -3 & -1 & 0 & 0 \end{pmatrix}$$

$$P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ -3 & 0 & -3 & -2 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 \end{pmatrix}$$

$$\begin{aligned}
P^{-1}AP &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ -3 & 0 & -3 & -2 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ -2 & 1 & -1 & 1 \\ 2 & -1 & 2 & 0 \end{pmatrix}\begin{pmatrix} -1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 1 & 0 & -1 \\ -3 & -1 & 0 & 0 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} -1 & 0 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} -1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 1 & 0 & -1 \\ -3 & -1 & 0 & 0 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

Just as expected.

3.3 Trinity College Dublin Example

Consider the matrix:

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$$

We begin by computing its characteristic polynomial:

$$\begin{aligned}
\mathcal{X}_A(x) &= 5(4-(4-x)) - x[(-2-x)(4-x)+14] \\
&= 5x - x[-(2+x)(4-x)+14] \\
&= x[5+(8+2x-x^2)-14] \\
&= -x[x^2-2x+1] \\
&= -x(x-1)^2
\end{aligned}$$

Thus, $\lambda_1 = 0$ has algebraic multiplicity 1, whilst $\lambda_2 = 1$ has algebraic multiplicity 2.

The next step is to compute the eigenvectors which span any necessary (generalised) eigenspace.

With $\lambda_1 = 1$ we have:

$$A\underline{v} = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \implies \begin{pmatrix} v_1 = 0 \\ 2v_1 + v_2 = 0 \end{pmatrix}$$

Letting $s = v_1$, we thus get that $ker(A - 0I_3)$ is spanned by:

$$\begin{pmatrix} 0 \\ s \\ -2s \end{pmatrix} = s \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

The dimension of the kernel is 1, which is the algebraic multiplicity of $\lambda_1 = 0$, so we are done.

Moving on to $\lambda_2 = 1$:

$$(A - I_3)\underline{v} = \begin{pmatrix} -3 & 2 & 1 \\ -7 & 3 & 2 \\ 5 & 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \underline{0} \implies \begin{pmatrix} 5v_1 - v_3 = 0 \\ -3v_1 + 2v_2 + v_3 = 0 \\ -7v_1 + 3v_2 + 2v_3 = 0 \end{pmatrix}$$

Letting $v_1 = s$, the first equation tells us that $v_3 = 5s$. The second equation then tells us that:

$$-3s + 2v_2 + 5s = 0 \implies v_2 = -s$$

Thus, $ker(A - I_3)$ is spanned by:

$$\begin{pmatrix} s \\ -s \\ 5s \end{pmatrix} = s \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}$$

This kernel has dimension one, but we have algebraic multiplicity 2. Thus, we need to compute $ker((A-I_3)^2)$.

We begin by squaring:

$$(A - I_3)^2 = \begin{pmatrix} -3 & 2 & 1 \\ -7 & 3 & 2 \\ 5 & 0 & -1 \end{pmatrix} \begin{pmatrix} -3 & 2 & 1 \\ -7 & 3 & 2 \\ 5 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix}$$

So we seek to satisfy:

$$(A - I_3)^2\underline{v} = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \underline{0}$$

Notice, letting $v_1 = s, v_2 = t$, we get that $ker((A - I_3)^2)$ is spanned by:

$$\begin{pmatrix} s \\ t \\ \frac{-10s+5t}{-3} \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ \frac{10}{3} \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ -\frac{5}{3} \end{pmatrix}$$

We now apply the algorithm. For $\lambda_1 = 0$, we can just choose:

$$\underline{u} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

as a basis for $E^{gen}(0, A)$.

For $\lambda_2 = 1$, we pick $\underline{v}$ such that:

$$\underline{v} \in ker((A - I_3)^2) \qquad \underline{v} \notin ker(A - I_3)$$

We can just pick:

$$\underline{v}_{2,1} = \begin{pmatrix} 1 \\ 0 \\ \frac{10}{3} \end{pmatrix}$$

We now just have to apply:

$$\underline{v}_{1,1} = (A - I_3)\underline{v}_{2,1} = \begin{pmatrix} -3 & 2 & 1 \\ -7 & 3 & 2 \\ 5 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \frac{10}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{5}{3} \end{pmatrix}$$

As ordered basis we then pick:

$$\mathcal{B} = \{\underline{u}, \underline{v}_{1,1}, \underline{v}_{2,1}\}$$

Such that:

$$P = \begin{pmatrix} 0 & \frac{1}{3} & 1 \\ 1 & -\frac{1}{3} & 0 \\ -2 & \frac{5}{3} & \frac{10}{3} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 0 & 1 & 3 \\ 3 & -1 & 0 \\ -6 & 5 & 10 \end{pmatrix}$$

$$P^{-1} = \begin{pmatrix} 10 & -5 & -3 \\ 30 & -18 & -9 \\ -9 & 6 & 3 \end{pmatrix}$$

Then, we predict a JNF with 2 Jordan Blocks: 1 corresponding to $\lambda_1 = 0$ (size 1×1), and 1 corresponding to $\lambda_2 = 1$ of size 2×2 (corresponding to $ker((A - I_3)^2)$. We compute:

$$P^1AP = \frac{1}{3}\begin{pmatrix} 10 & -5 & -3 \\ 30 & -18 & -9 \\ -9 & 6 & 3 \end{pmatrix}\begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 3 \\ 3 & -1 & 0 \\ -6 & 5 & 10 \end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 \\ 21 & -12 & -6 \\ -9 & 6 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 & 3 \\ 3 & -1 & 0 \\ -6 & 5 & 10 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 7 & -4 & -2 \\ -3 & 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 3 \\ 3 & -1 & 0 \\ -6 & 5 & 10 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

as expected.

For more examples, check Trinity College Dublin - Jordan Normal Form (Some Examples)

4 Workshop

1. **True or false. There exists a matrix in $Mat(3;\mathbb{C})$ with an eigenvalue of geometric multiplicity 2 and algebraic multiplicity 1.**

   This is false. By Remark 6.3.3, the algebraic multiplicity of any eigenvalue is always greater than or equal to the geometric multiplicity.

2. **The matrix with entries in $\mathbb{C}$:**

   $$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix}$$

   has characteristic polynomial:

   $$p_A(x) = x^2(-1-x)$$

   Find by hand an invertible matrix P such that $P^{-1}AP$ is in Jordan Normal Form.

   We have 2 eigenvalues:

   $$\lambda = 0 \qquad \lambda = -1$$

   $\lambda = 0$ has algebraic multiplicity 2, so to compute $E^{gen}(0, A)$, we require 2 basis vectors. We begin by computing the eigenvector corresponding to $ker(A - 0I_3)$:

   $$(A - 0I_3)\underline{v} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \underline{0}$$

which implies that:

$$v_1 + v_2 - v_3 = 0 \qquad 2v_1 + 4v_2 - 3v_3 = 0$$

The first equality tells us that $v_3 = v_1 + v_2$ so:

$$2v_1 + 4v_2 - 3(v_1 + v_2) = 0 \implies v_2 - v_1 = 0 \implies v_1 = v_2$$

Hence, any eigenvector corresponding to A with eigenvalue $\lambda = 0$ is spanned by:

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

However, this doesn't give us a basis for $E^{gen}(0, A)$. We thus compute the vectors associated to $ker((A - 0I_3)^2)$. We first compute:

$$(A - 0I_3)^2 = A^2 = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix} = \begin{pmatrix} -1 & -3 & 2 \\ -2 & -6 & 4 \\ -4 & -12 & 8 \end{pmatrix}$$

Thus:

$$(A - 0I_3)^2 \underline{v} = \begin{pmatrix} -1 & -3 & 2 \\ -2 & -6 & 4 \\ -4 & -12 & 8 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \underline{0}$$

which implies that:

$$-v_1 - 3v_2 + 2v_3 = 0 \implies v_3 = \frac{v_1 + 3v_2}{2}$$

Hence, the vectors corresponding to $ker((A - 0I_3)^2)$ are given by the span of:

$$\begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \frac{3}{2} \end{pmatrix}$$

or alternatively:

$$\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$$

Notice, this span does include the original eigenvector $(1, 1, 2)^T$, since $(1, 1, 2)^T = (1, 0, 1/2)^T + (0, 1, 3/2)^T$, so we could've perfectly constructed the basis by using $(2, 0, 1)^T, (1, 1, 2)^T$ as is done in the solutions.

Notice, this basis contains to vectors, so this is the basis we were seeking for $E^{gen}(0, A)$.

We now compute a basis for $E(-1, A)$. Since $\lambda = -1$ has algebraic multiplicity 1, we seek a single vector for this. We compute a basis for $ker(A + I_3)$:

$$(A + I_3)\underline{v} = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 5 & -3 \\ 4 & 8 & -5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \underline{0}$$

which implies that:

$$2v_1 + v_2 - v_3 = 0 \qquad 2v_1 + 5v_2 - 3v_3 = 0 \qquad 4v_1 + 8v_2 - 5v_3 = 0$$

From the first equality, $v_3 = 2v_1 + v_2$ so:

$$2v_1 + 5v_2 - 3(2v_1 + v_2) = 0 \implies -4v_1 + 2v_2 = 0$$

$$4v_1 + 8v_2 - 5(2v_1 + v_2) = 0 \implies -6v_1 + 3v_2 = 0$$

So we must have:

$$v_2 = 2v_1$$

Thus, the vector spanning $ker(A + I_3)$ is:

$$\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

For the block corresponding to $\lambda = -1$, we need a single vector, so we just pick:

$$\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

For the block(s) corresponding to $\lambda = 0$, we first seek a vector in $ker((A - 0I_3)^2)$ which isn't in $ker(A - 0I_3)$. For example:

$$\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

We then have to compute:

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Hence, we have found 3 vectors, so our (ordered)basis will be:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

> *The order of the basis is important; for each eigenvector corresponding to an eigenvalue, we order from left to right according to the power of the operator by which it is multiplied. In this case, for $\lambda = 0$, we use $A^0\underline{v}, A^1\underline{v}$ in the ordering.*

We construct the transformation matrix:

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{pmatrix}$$

We need to invert. Just for fun, and since I want to practice, I'll use 2 methods:

- **Adjugate**:

  We begin by computing the cofactors:

  $$C_{1,1} = 1 - 0 = 1 \qquad C_{1,2} = -(2-0) = -2 \qquad C_{1,3} = 4 - 4 = 0$$
  $$C_{2,1} = -(1-4) = 3 \qquad C_{2,2} = 1 - 8 = -7 \qquad C_{2,3} = -(2-4) = 2$$
  $$C_{3,1} = 0 - 2 = -2 \qquad C_{3,2} = -(0-4) = 4 \qquad C_{3,3} = 1 - 2 = -1$$

  Thus, the matrix of minors is:

  $$\begin{pmatrix} 1 & -2 & 0 \\ 3 & -7 & 2 \\ -2 & 4 & -1 \end{pmatrix}$$

  so transposing this gives us:

  $$\begin{pmatrix} 1 & 3 & -2 \\ -2 & -7 & 4 \\ 0 & 2 & -1 \end{pmatrix}$$

  Finally, we need to divide by the determinant of the matrix. If we expand along the second row:

  $$det(P) = 2C_{2,1} + C_{2,2} + 0C_{2,3} = 6 + (-7) = -1$$

  Hence, dividing by -1:

  $$P^{-1} = \begin{pmatrix} -1 & -3 & 2 \\ 2 & 7 & -4 \\ 0 & -2 & 1 \end{pmatrix}$$

- **Identity**:

$$
\begin{aligned}
& \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 4 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \\
R_2 - 2R_1, R_3 - 4R_1 \implies & \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -4 & -2 & 1 & 0 \\ 0 & -2 & -7 & -4 & 0 & 1 \end{array}\right) \\
-R_2 \implies & \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 & -1 & 0 \\ 0 & -2 & -7 & -4 & 0 & 1 \end{array}\right) \\
R_1 - R_2, R_3 + 2R_2 \implies & \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 1 & 0 \\ 0 & 1 & 4 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right) \\
R_1 + 2R_3, R_2 - 4R_3 \implies & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -3 & 2 \\ 0 & 1 & 0 & 2 & 7 & -4 \\ 0 & 0 & 1 & 0 & -2 & 1 \end{array}\right)
\end{aligned}
$$

So we have that:

$$
P^{-1} = \begin{pmatrix} -1 & -3 & 2 \\ 2 & 7 & -4 \\ 0 & -2 & 1 \end{pmatrix}
$$

Now, we expect the Jordan matrix to be composed of 2 blocks: the first block will be 1×1 and correspond to $\lambda = -1$; the second block will be 2×2, and correspond to $\lambda = 0$.

We can confirm this:

$$
\begin{aligned}
P^{-1}AP &= \begin{pmatrix} -1 & -3 & 2 \\ 2 & 7 & -4 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 2 & 4 & -3 \\ 4 & 8 & -6 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 3 & -2 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

3. **A (5×5) matrix with entries in $\mathbb{C}$ has 2 eigenvalues: 0 and 1. How many possible JNFs (up to re-ordering) does the matrix have?**

   Notice, the diagonal will always have at least one 1 or 0. In particular, the diagonal can be in one of 4 forms (up to reordering of the blocks):

   - $(1, 0, 0, 0, 0)$
   - $(1, 1, 0, 0, 0)$
   - $(1, 1, 1, 0, 0)$
   - $(1, 1, 1, 1, 0)$

   There isn't any sophisticated mathematical formulae which yields an answer; we need to be careful with how we count.

   Notice, the first 2 cases are analogous to the last 2 cases, with the 0s and 1s "flipped", so counting the possibilities for $(1, 0, 0, 0, 0)$ and $(1, 1, 0, 0, 0)$, and then multiplying by 2 gives us our answer.

   For $(1, 0, 0, 0, 0)$ the following block distributions are possible:

   - 1|0000 (block of 4 zeroes)
   - 1|0|000 (block of 1 zero, block of 3 zeroes)
   - 1|0|0|00 (2 blocks of 1 zero, block of 2 zeros)
   - 1|0|0|0|0 (4 blocks of 1 zero)
   - 1|00|00 (2 blocks of 2 zeroes)

   For $(1, 1, 0, 0, 0)$ the following block distributions are possible:

   - 1|1|000 (2 blocks of ones, block of 3 zeroes)
   - 1|1|0|00 (2 blocks of ones, block of 1 zero, block of 2 zeroes)
   - 1|1|0|0|0 (2 blocks of ones, 3 blocks of zero)
   - 11|000 (block of 2 ones, block of 3 zeroes)
   - 11|0|00 (block of 2 ones, block of 1 zero, block
   - 11|0|0|0 (block of 2 ones, 3 blocks of zero)

   Hence, we have 11 possibilities for the first 2 options. Thus, in total, we have $2 \times 11 = 22$ possible JNFs.