

10.3 The Jordan Canonical Form

Chapter 4 asked how simple the matrix of an operator can be made by choosing the basis well, and Chapter 9 answered it only for operators with enough eigenvectors. The two previous sections have quietly finished the job: the space splits into generalized eigenspaces, and on each of them the operator is a scalar plus a nilpotent operator, which is a disjoint union of chains. Assembling these pieces gives the best matrix there is, and it is unique. This section states that theorem, proves it, reads the familiar invariants off it, and turns it into a similarity test.

10.3.1 The shape we are aiming for

Here is the goal, stated before its parts are named. **If p_T splits, then V has a basis in which the matrix of T is diagonal except for some ones immediately above the diagonal, and that matrix is unique up to the order of its diagonal blocks.** A diagonal matrix is the special case with no ones at all. So the dream of Chapter 4 comes true, with one extra ingredient beyond a diagonal matrix: the superdiagonal ones, one for each step of a Jordan chain.

The building block is the matrix of $\lambda \text{id} + N$ on a single chain. By Lemma 10.2.2 (b), the nilpotent part contributes $\mathbf{J}_k(0)$, so the block is $\lambda \mathbf{I}_k + \mathbf{J}_k(0)$.

Definition 10.3.1. Jordan Block

Let $k \geq 1$ and $\lambda \in F$. The **Jordan block** of size k for λ is the matrix

$$\mathbf{J}_k(\lambda) := \lambda \mathbf{I}_k + \mathbf{J}_k(0) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in M_k(F),$$

with λ in every diagonal entry, 1 in every **superdiagonal** entry $(j, j+1)$, and 0 elsewhere. In particular $\mathbf{J}_1(\lambda) = (\lambda)$.

Reading its columns: if $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{J}_k(\lambda)$ for a basis $\mathcal{B} = (b_1, \dots, b_k)$, then $Tb_1 = \lambda b_1$ and $Tb_j = b_{j-1} + \lambda b_j$ for $j \geq 2$; equivalently $(T - \lambda \text{id}_V)b_j = b_{j-1}$ and $(T - \lambda \text{id}_V)b_1 = 0$, so $\mathcal{B}$ is the Jordan chain generated by b_k (Definition 10.2.2). The single eigenvalue is λ , with $a(\lambda) = k$ and $g(\lambda) = 1$, since $\mathbf{J}_k(\lambda) - \lambda \mathbf{I}_k = \mathbf{J}_k(0)$ has rank $k-1$.

A Jordan matrix is a diagonal matrix with some chains glued in: blocks along the diagonal, each block one eigenvalue with ones above it.

Definition 10.3.2. Jordan Matrix, Jordan Basis, Jordan Form

A **Jordan matrix** is a matrix of the form

$$\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \mathbf{J}_{k_2}(\lambda_2) \oplus \cdots \oplus \mathbf{J}_{k_m}(\lambda_m)$$

for some $m \geq 1$, sizes $k_i \geq 1$, and scalars $\lambda_i \in F$, **not** assumed distinct. Let V be a finite-dimensional vector space over F with $\dim V \geq 1$ and $T \in \mathcal{L}(V)$. A **Jordan form** of T is

a Jordan matrix $\mathbf{J}$ with $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{J}$ for **some** basis $\mathcal{B}$ of V ; such a basis is called a **Jordan basis** for T . For $\mathbf{A} \in M_n(F)$, a **Jordan form** of $\mathbf{A}$ is a Jordan matrix $\mathbf{J}$ with $\mathbf{A} \sim \mathbf{J}$.

In words: the matrix is block diagonal (Definition 8.1.2), and each diagonal block is a single Jordan block. Different blocks may carry the same eigenvalue or different ones, and the sizes may repeat. A Jordan basis is a basis that is a concatenation of Jordan chains, one chain per block. The two descriptions match for $\mathbf{A} \in M_n(F)$ and $T = T_{\mathbf{A}}$, by Theorem 4.2.4.

The order of the blocks never matters. Reordering the blocks of $\mathbf{J}$ permutes the segments of the basis $\mathcal{B}$, so the reordered matrix is the matrix of the **same** operator in the permuted basis, hence similar to $\mathbf{J}$ (Theorem 4.2.4 (a)). This is why the uniqueness statement below can only be “up to the order of the blocks”, and why it costs nothing to list the blocks in any convenient order.

Examples.

- **Diagonal matrices.** $\text{diag}(d_1, \dots, d_n) = \mathbf{J}_1(d_1) \oplus \dots \oplus \mathbf{J}_1(d_n)$ is a Jordan matrix with all blocks of size 1. So every diagonalizable operator has a Jordan form, namely a diagonal one.
- **The standing counterexample.** $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \mathbf{J}_2(2)$ is its own Jordan form, and $(\mathbf{e}_1, \mathbf{e}_2)$ is a Jordan basis: $\mathbf{e}_1$ is an eigenvector and $(\mathbf{J} - 2\mathbf{I})\mathbf{e}_2 = \mathbf{e}_1$.
- **Nilpotent operators.** By Theorem 10.2.3, a nilpotent operator has the Jordan form $\mathbf{J}_{k_1}(0) \oplus \dots \oplus \mathbf{J}_{k_r}(0)$, all eigenvalues 0. Section 2 is exactly the case $\lambda = 0$ of this section.
- **Mixed blocks.** $\mathbf{J} = \mathbf{J}_2(3) \oplus \mathbf{J}_1(3) \oplus \mathbf{J}_2(1) \in M_5(F)$ has $p_{\mathbf{J}} = (x - 3)^3(x - 1)^2$, $g(3) = 2$, and $g(1) = 1$. It is the answer to the worked example at the end of this section.
- **The degenerate case.** On a space of dimension 1, every operator is $(\lambda) = \mathbf{J}_1(\lambda)$. One block, size one, no superdiagonal to speak of.

Nonexample by minimal change. The matrix $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ is upper triangular with a superdiagonal 1, but it is **not** a Jordan matrix: a 1 above the diagonal may only join two **equal** diagonal entries, because it records the step of a chain for a single eigenvalue. (Indeed this matrix has distinct eigenvalues, so it is diagonalizable, and its Jordan form is $\text{diag}(2, 3)$.) Changing the superdiagonal entry instead, $\begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix}$ is not a Jordan matrix either, since the entry above the diagonal must be 1; it is similar to $\mathbf{J}_2(2)$, by rescaling the second basis vector.

Why this definition. Two conventions are being fixed. First, the ones sit **above** the diagonal, which matches the chain order “eigenvector first” of Definition 10.2.2; putting them below the diagonal gives the transposed matrix, which is just as good and appears in some books. Second, the off-diagonal entries are 1 rather than arbitrary nonzero scalars: a nonzero entry c can be turned into a 1 by scaling one basis vector, so nothing is lost, and the normalization is what makes the form unique. Nothing else is allowed: if any further entry were permitted, the form would no longer be determined by the operator.

Warning. A Jordan form exists only when p_T splits, so it depends on the field.

The rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ has no Jordan form over $\mathbb{R}$, since $p_{\mathbf{R}} = x^2 + 1$ has no real root, and a Jordan matrix is triangular with its eigenvalues on the diagonal. Over $\mathbb{C}$ its Jordan form is $\text{diag}(i, -i)$. Over $\mathbb{Q}$ the matrix $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$ has no Jordan form either, and over

$\mathbb{R}$ it becomes $\text{diag}(\sqrt{2}, -\sqrt{2})$. Section 6 gives a canonical form valid over **every** field.

10.3.2 The theorem

Every operator whose characteristic polynomial splits is a direct sum of Jordan blocks, in exactly one way.

Theorem 10.3.1. Jordan Canonical Form Theorem

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, and let $T \in \mathcal{L}(V)$.

- (a) **(Existence)** T has a Jordan form if and only if p_T splits over F .
- (b) **(Uniqueness)** Suppose $\mathbf{J}$ is a Jordan form of T . Then for every $\lambda \in F$ and every integer $j \geq 1$, the number of blocks $\mathbf{J}_k(\lambda)$ in $\mathbf{J}$ with $k \geq j$ equals

$$\text{rank}(T - \lambda \text{id}_V)^{j-1} - \text{rank}(T - \lambda \text{id}_V)^j,$$

a number depending only on T , λ , and j . Consequently any two Jordan forms of T have the same blocks, up to the order in which they are listed.

Equivalently, for $\mathbf{A} \in M_n(F)$ with $p_{\mathbf{A}}$ split: $\mathbf{A}$ is similar to a Jordan matrix, and that matrix is unique up to the order of its blocks. We call it **the** Jordan form of $\mathbf{A}$.

Idea. Existence is an assembly of the two previous sections, in three steps. ① Split $V = \bigoplus_i G_{\lambda_i}(T)$ (Theorem 10.1.3). ② On G_{λ_i} write $T = \lambda_i \text{id} + N_i$ with N_i nilpotent, and take a Jordan basis for N_i (Theorem 10.2.3); adding $\lambda_i \mathbf{I}$ turns each $\mathbf{J}_k(0)$ into $\mathbf{J}_k(\lambda_i)$. ③ Put the bases together: the matrix is block diagonal. For uniqueness, count ranks. A block $\mathbf{J}_k(\mu) - \lambda \mathbf{I}$ with $\mu \neq \lambda$ is invertible, so it contributes its full size to every power and drops out of differences; a block with $\mu = \lambda$ becomes $\mathbf{J}_k(0)$, whose j -th power has rank $\max(k-j, 0)$. So the difference of two consecutive ranks counts the λ -blocks that are still long enough.

Proof.

- (a) ($\Leftarrow$) Suppose p_T splits, with distinct eigenvalues $\lambda_1, \dots, \lambda_d$.

Step 1. Split the space. By Theorem 10.1.3, $V = G_1 \oplus \dots \oplus G_d$ with $G_i = G_{\lambda_i}(T)$ nonzero and T -invariant, $\dim G_i = a_i := a_T(\lambda_i)$, and

$$T|_{G_i} = \lambda_i \text{id}_{G_i} + N_i, \quad N_i = (T - \lambda_i \text{id}_V)|_{G_i} \text{ nilpotent.}$$

Step 2. One eigenvalue at a time. Fix i . By Theorem 10.2.3 applied to $N_i \in \mathcal{L}(G_i)$, there is a basis $\mathcal{B}_i$ of G_i , a concatenation of Jordan chains of N_i , with

$$[N_i]_{\mathcal{B}_i}^{\mathcal{B}_i} = \mathbf{J}_{k_{i1}}(0) \oplus \dots \oplus \mathbf{J}_{k_{ir_i}}(0), \quad \sum_l k_{il} = a_i.$$

Since $[\cdot]_{\mathcal{B}_i}^{\mathcal{B}_i}$ is linear (Theorem 4.1.4) and $[\text{id}_{G_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = \mathbf{I}_{a_i} = \mathbf{I}_{k_{i1}} \oplus \cdots \oplus \mathbf{I}_{k_{ir_i}}$,

$$\begin{aligned} [T|_{G_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} &= \lambda_i \mathbf{I}_{a_i} + [N_i]_{\mathcal{B}_i}^{\mathcal{B}_i} \\ &= (\lambda_i \mathbf{I}_{k_{i1}} + \mathbf{J}_{k_{i1}}(0)) \oplus \cdots \oplus (\lambda_i \mathbf{I}_{k_{ir_i}} + \mathbf{J}_{k_{ir_i}}(0)) \\ &= \mathbf{J}_{k_{i1}}(\lambda_i) \oplus \cdots \oplus \mathbf{J}_{k_{ir_i}}(\lambda_i), \end{aligned}$$

where the middle equality is the blockwise addition of Theorem 8.1.4 (a) and the last is Definition 10.3.1.

Step 3. Assemble. Let $\mathcal{B}$ be the basis of V obtained by writing $\mathcal{B}_1, \dots, \mathcal{B}_d$ one after another. By Theorem 9.1.5,

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [T|_{G_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} \oplus \cdots \oplus [T|_{G_d}]_{\mathcal{B}_d}^{\mathcal{B}_d},$$

which by Step 2 is a direct sum of Jordan blocks, that is, a Jordan matrix. So T has a Jordan form, and $\mathcal{B}$ is a Jordan basis.

($\Rightarrow$) Suppose $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{J}$ is a Jordan matrix. Each Jordan block is upper triangular, so $\mathbf{J}$ is upper triangular, and T is triangularizable (Definition 9.7.1). By Theorem 9.7.2, p_T splits over F .

- (b) Let $\mathbf{J} = \mathbf{J}_{k_1}(\mu_1) \oplus \cdots \oplus \mathbf{J}_{k_m}(\mu_m) = [T]_{\mathcal{B}}^{\mathcal{B}}$, with $\mu_1, \dots, \mu_m \in F$ not necessarily distinct; fix $\lambda \in F$ and let $j \geq 0$. Then $[T - \lambda \text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{J} - \lambda \mathbf{I}_n$, which is block diagonal with blocks $\mathbf{J}_{k_i}(\mu_i) - \lambda \mathbf{I}_{k_i}$, so by Theorem 8.1.4 (b),

$$[(T - \lambda \text{id}_V)^j]_{\mathcal{B}}^{\mathcal{B}} = (\mathbf{J}_{k_1}(\mu_1) - \lambda \mathbf{I}_{k_1})^j \oplus \cdots \oplus (\mathbf{J}_{k_m}(\mu_m) - \lambda \mathbf{I}_{k_m})^j,$$

and the rank of a block diagonal matrix is the sum of the ranks of its blocks (Theorem 8.2.1 (b), applied $m - 1$ times). Also $\text{rank}(T - \lambda \text{id}_V)^j$ equals the rank of this matrix (Theorem 4.1.5 (a)). Now compute block by block.

Blocks with $\mu_i \neq \lambda$. Here $\mathbf{J}_{k_i}(\mu_i) - \lambda \mathbf{I}_{k_i}$ is upper triangular with every diagonal entry equal to $\mu_i - \lambda \neq 0$, so its determinant is $(\mu_i - \lambda)^{k_i} \neq 0$ (Theorem 7.4.6) and it is invertible (Theorem 7.5.2). Hence its j -th power is invertible too, of rank k_i , for every $j \geq 0$.

Blocks with $\mu_i = \lambda$. Here $\mathbf{J}_{k_i}(\lambda) - \lambda \mathbf{I}_{k_i} = \mathbf{J}_{k_i}(0)$, whose standard basis is a single Jordan chain of length k_i , so $\text{rank } \mathbf{J}_{k_i}(0)^j = \max(k_i - j, 0)$ by Theorem 10.2.4 (a) with the one-element list of lengths (k_i) .

Adding up, with $c = \sum_{\mu_i \neq \lambda} k_i$ independent of j ,

$$\text{rank}(T - \lambda \text{id}_V)^j = c + \sum_{\mu_i = \lambda} \max(k_i - j, 0).$$

Subtracting the values at j and $j - 1$, the constant c cancels and each remaining bracket $\max(k_i - j + 1, 0) - \max(k_i - j, 0)$ equals 1 if $k_i \geq j$ and 0 otherwise, exactly as in the proof of Theorem 10.2.4 (b). Therefore

$$\text{rank}(T - \lambda \text{id}_V)^{j-1} - \text{rank}(T - \lambda \text{id}_V)^j = \#\{i : \mu_i = \lambda \text{ and } k_i \geq j\}.$$

The left-hand side is defined from T alone. So for each λ and each j the number of λ -blocks of size at least j , and hence by subtraction the number of size exactly j , is the same for every Jordan form of T . Two Jordan forms therefore consist of the same blocks with the same multiplicities, and differ only in the order in which the blocks are listed.

For a matrix $\mathbf{A} \in M_n(F)$ with $p_{\mathbf{A}}$ split, apply all of this to $T_{\mathbf{A}}$, using Theorem 4.2.4 to pass between bases and similarity. ■

Step 1 $\rightarrow$ Step 2 $\rightarrow$ Step 3, and the dream comes true: **for every complex square matrix there is a basis in which the matrix is diagonal apart from some ones on the superdiagonal**, since $p_{\mathbf{A}}$ always splits over $\mathbb{C}$ (Corollary 6.5.5). The proof also tells us where the basis comes from: one chain per block, found inside one generalized eigenspace at a time.

Quick check. How many Jordan forms, up to the order of the blocks, are possible for an operator with $p_T = (x - 2)^2(x - 3)$? Write them down, and say which invariant tells them apart.

Solution. The 3-blocks must have sizes summing to $a(3) = 1$, so there is exactly one, of size 1. The 2-blocks have sizes summing to $a(2) = 2$, so they form a partition of 2: either (2) or (1, 1). The two possibilities are

$$\mathbf{J}_2(2) \oplus \mathbf{J}_1(3) \quad \text{and} \quad \mathbf{J}_1(2) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(3) = \text{diag}(2, 2, 3).$$

They are told apart by $g(2)$, which is the number of 2-blocks: 1 in the first case, 2 in the second. Equivalently by $m_T = (x - 2)^2(x - 3)$ against $(x - 2)(x - 3)$, or by $\text{rank}(T - 2 \text{ id})$, which is 2 against 1.

10.3.3 Reading the invariants off the blocks

The Jordan form makes every invariant of Chapter 9 visible at a glance.

Corollary 10.3.2. Invariants from the Jordan Form

Let V be finite-dimensional with $\dim V = n \geq 1$, let $T \in \mathcal{L}(V)$ with p_T split, let $\mathbf{J}$ be a Jordan form of T , and let $\lambda \in F$. Then:

- (a) $a_T(\lambda)$ is the **sum of the sizes** of the λ -blocks of $\mathbf{J}$; in particular $\lambda \in \text{spec}(T)$ if and only if $\mathbf{J}$ has a λ -block;
- (b) $g_T(\lambda)$ is the **number** of λ -blocks of $\mathbf{J}$;
- (c) the exponent of $x - \lambda$ in m_T is the **largest size** of a λ -block of $\mathbf{J}$ (and 0 when $\mathbf{J}$ has no λ -block);
- (d) $\dim \ker(T - \lambda \text{ id}_V)^j = \sum \min(k_i, j)$, summed over the sizes k_i of the λ -blocks, for every $j \geq 0$;
- (e) T is diagonalizable if and only if every block of $\mathbf{J}$ has size 1, that is, if and only if $\mathbf{J}$ is diagonal.

Proof. Write $\mathbf{J} = \mathbf{J}_{k_1}(\mu_1) \oplus \cdots \oplus \mathbf{J}_{k_m}(\mu_m)$, with $\mu_i \in F$ not necessarily distinct, and $\mathcal{B}$ for a Jordan basis, so $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{J}$.

- (a) $\mathbf{J}$ is upper triangular with diagonal entries μ_i repeated k_i times, so $p_T = \prod_{i=1}^m (x - \mu_i)^{k_i}$

(Theorem 7.4.6, Definition 7.10.2). Grouping equal μ_i and using that the multiplicity of λ as a root of a product of powers of distinct linear factors is the exponent it carries (Lemma 6.4.1), $a_T(\lambda) = \sum_{\mu_i=\lambda} k_i$. The last claim follows since $\text{spec}(T)$ is the set of roots of p_T (Theorem 9.2.1).

- (b) Take $j = 1$ in [Theorem 10.3.1](#) (b): the number of λ -blocks of size ≥ 1 , which is all of them, equals

$$\begin{aligned} \text{rank}(T - \lambda \text{id}_V)^0 - \text{rank}(T - \lambda \text{id}_V) &= n - \text{rank}(T - \lambda \text{id}_V) \\ &= \dim \ker(T - \lambda \text{id}_V) = g_T(\lambda), \end{aligned}$$

by Theorem 2.3.1 and Definition 9.3.2.

- (c) First, $m_{\mathbf{J}_k(\nu)} = (x - \nu)^k$ for every $\nu \in F$: the polynomial $(x - \nu)^k$ annihilates $\mathbf{J}_k(\nu)$ because $\mathbf{J}_k(\nu) - \nu \mathbf{I} = \mathbf{J}_k(0)$ and $\mathbf{J}_k(0)^k = \mathbf{O}$, while $\mathbf{J}_k(0)^{k-1} \neq \mathbf{O}$, so no smaller power of $x - \nu$ annihilates it; and $m_{\mathbf{J}_k(\nu)}$ divides $p_{\mathbf{J}_k(\nu)} = (x - \nu)^k$ (Corollary 9.6.2), hence is a power of $x - \nu$ (Lemma 9.5.2). Applying Proposition 9.5.6 $m - 1$ times,

$$m_T = m_{\mathbf{J}} = \text{lcm}((x - \mu_1)^{k_1}, \dots, (x - \mu_m)^{k_m}) = \prod_{\nu \in \text{spec}(T)} (x - \nu)^{\max\{k_i : \mu_i = \nu\}},$$

the last equality because the least common multiple of powers of linear factors takes, for each factor, the largest exponent occurring (Definition 6.2.4, Theorem 6.3.5).

- (d) The proof of [Theorem 10.3.1](#) (b) computed

$$\text{rank}(T - \lambda \text{id}_V)^j = c + \sum_{\mu_i=\lambda} \max(k_i - j, 0), \quad c := \sum_{\mu_i \neq \lambda} k_i.$$

Since $\sum_{i=1}^m k_i = n$, we have $n - c = \sum_{\mu_i=\lambda} k_i$, so by Theorem 2.3.1

$$\begin{aligned} \dim \ker(T - \lambda \text{id}_V)^j &= n - c - \sum_{\mu_i=\lambda} \max(k_i - j, 0) \\ &= \sum_{\mu_i=\lambda} (k_i - \max(k_i - j, 0)) \\ &= \sum_{\mu_i=\lambda} \min(k_i, j). \end{aligned}$$

- (e) If every block has size 1, then $\mathbf{J}$ is diagonal and T is diagonalizable (Definition 9.4.1). Conversely, if T is diagonalizable, then m_T is a product of distinct linear factors (Theorem 9.8.2), so every exponent in (c) is 1, which says that every block has size 1. ■

Parts (a) and (b) give the practical reading: **the total size of the λ -blocks is $a(\lambda)$, and their number is $g(\lambda)$** . The gap $a(\lambda) - g(\lambda)$, which Chapter 9 could only measure, is now explained: it counts the superdiagonal ones belonging to λ .

Warning. The multiplicities $a(\lambda)$ and $g(\lambda)$ do not always determine the blocks. For $n = 7$ and one eigenvalue λ with $a(\lambda) = 7$ and $g(\lambda) = 3$, the block sizes form a partition of 7 into 3 parts, and there are four of those: $(5, 1, 1)$, $(4, 2, 1)$, $(3, 3, 1)$, $(3, 2, 2)$. Knowing $a(\lambda)$ and $g(\lambda)$ means knowing $\dim \ker(T - \lambda \text{id}_V)^j$ for $j = 1$ and for j large; the intermediate values $j = 2, 3, \dots$ are genuinely new information, and [Theorem 10.3.1](#) (b) needs them all.

10.3.4 Deciding similarity

We can now settle the question Chapter 4 raised and Chapter 7 sharpened: which data decide whether two matrices are similar?

Corollary 10.3.3. Similarity via the Jordan Form

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$, $n \geq 1$, and suppose $p_{\mathbf{A}}$ splits over F . The following are equivalent:

- (a) $\mathbf{A} \sim \mathbf{B}$;
- (b) $p_{\mathbf{B}}$ splits over F , and $\mathbf{A}$ and $\mathbf{B}$ have the same Jordan form, up to the order of the blocks;
- (c) $p_{\mathbf{A}} = p_{\mathbf{B}}$ and $\text{rank}(\mathbf{A} - \lambda \mathbf{I})^j = \text{rank}(\mathbf{B} - \lambda \mathbf{I})^j$ for every $\lambda \in \text{spec}(\mathbf{A})$ and every $j = 1, \dots, n$.

Proof.

- (a) $\Rightarrow$ (c). Similar matrices have the same characteristic polynomial (Theorem 7.10.3). If $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, then $(\mathbf{B} - \lambda \mathbf{I})^j = \mathbf{P}^{-1}(\mathbf{A} - \lambda \mathbf{I})^j\mathbf{P}$ for every λ and j (Proposition 4.2.6 (c) applied to the polynomial $(x - \lambda)^j$), so the ranks agree (Proposition 4.2.6 (a)).
- (b) $\Rightarrow$ (b). From $p_{\mathbf{B}} = p_{\mathbf{A}}$ we get that $p_{\mathbf{B}}$ splits and that $\text{spec}(\mathbf{B}) = \text{spec}(\mathbf{A})$ (Theorem 9.2.1). By Theorem 10.3.1 (a) both matrices have Jordan forms, and by (b) of that theorem the number of λ -blocks of each size in each of them is determined by the ranks $\text{rank}(\cdot - \lambda \mathbf{I})^j$, with $j \leq n$ sufficing because $(\mathbf{A} - \lambda \mathbf{I})^n$ and all higher powers have the same rank (Theorem 10.1.1 (d), Theorem 2.3.1). These numbers agree by hypothesis for $\lambda \in \text{spec}(\mathbf{A}) = \text{spec}(\mathbf{B})$, and for $\lambda \notin \text{spec}(\mathbf{A})$ neither form has a λ -block (Corollary 10.3.2 (a)). So the two Jordan forms have the same blocks.
- (c) $\Rightarrow$ (a). Let $\mathbf{J}$ be a Jordan form of $\mathbf{A}$ and $\mathbf{J}'$ one of $\mathbf{B}$, with the same blocks in possibly different orders. Then $\mathbf{J} \sim \mathbf{J}'$, because reordering blocks is a change of basis, as noted after Definition 10.3.2. Hence $\mathbf{A} \sim \mathbf{J} \sim \mathbf{J}' \sim \mathbf{B}$, and similarity is an equivalence relation (Example (Similarity)).

■

This is the promise of Chapter 4 kept, and of Chapter 7: when $p_{\mathbf{A}}$ splits — always, over $\mathbb{C}$ — the characteristic polynomial together with the ranks of the powers $(\mathbf{A} - \lambda \mathbf{I})^j$ is a **complete** set of similarity invariants. Chapter 4's warning pair is decided instantly: $\mathbf{I}_2$ has Jordan form $\text{diag}(1, 1)$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \mathbf{J}_2(1)$ is its own, so they are not similar. And the sharper failure of Chapter 9 becomes transparent.

Remark. The matrices $\mathbf{J}_2(1) \oplus \mathbf{J}_2(1)$ and $\mathbf{J}_2(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(1)$ in $M_4(F)$ have the same characteristic polynomial $(x - 1)^4$ **and** the same minimal polynomial $(x - 1)^2$, yet they are not similar: their Jordan forms differ, and $\text{rank}(\mathbf{A} - \mathbf{I})$ is 2 against 1. So no pair of polynomials can decide similarity; the block sizes carry strictly more information.

When $p_{\mathbf{A}}$ does not split, the Jordan form does not exist and this test says nothing. Section 6 replaces it by the rational canonical form, which works over every field.

Warning. **The Jordan form is not stable under small changes of the matrix.** For $\varepsilon \neq 0$, $\mathbf{A}_\varepsilon = \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix}$ over $\mathbb{C}$ has $p = x^2 - \varepsilon$ with two distinct roots $\pm\sqrt{\varepsilon}$, so $\mathbf{A}_\varepsilon$ is diagonalizable and its Jordan form is $\text{diag}(\sqrt{\varepsilon}, -\sqrt{\varepsilon})$, with **no** superdiagonal one. At $\varepsilon = 0$ the Jordan form jumps to $\mathbf{J}_2(0)$. So arbitrarily small perturbations change the form completely, which is why numerical work avoids it; Chapter 20 returns to this point.

10.3.5 Computing a Jordan form

The proof is an algorithm: it says to work one eigenvalue at a time and to find chains inside each generalized eigenspace. The rank data tell us the shape before we look for a single vector.

Algorithm. **Jordan form and a Jordan basis of $\mathbf{A} \in M_n(F)$.**

1. Factor $p_{\mathbf{A}}$. If it does not split over F , stop: there is no Jordan form.
2. For each eigenvalue λ , compute $\text{rank}(\mathbf{A} - \lambda\mathbf{I})^j$ for $j = 0, 1, 2, \dots$ until the rank stops changing, and get the λ -block sizes from [Theorem 10.3.1](#) (b), subtracting twice for the exact counts.
3. For each λ , find a Jordan basis of the nilpotent $(\mathbf{A} - \lambda\mathbf{I})$ restricted to $G_\lambda(\mathbf{A})$, as in Section 2: tops of long chains first, then vectors of $\ker(\mathbf{A} - \lambda\mathbf{I})$ to complete.
4. Let $\mathbf{P}$ be the matrix whose columns are all the chains, in order; then $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$. Check the answer as $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{J}$, column by column.

Example. The Jordan Form of a 5×5 Matrix

Let

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 0 & 0 & -1 \\ -2 & 2 & 2 & -2 & -1 \\ 1 & 0 & 2 & -1 & 1 \\ 1 & 0 & -1 & 2 & 1 \\ -2 & -1 & 2 & -2 & 2 \end{pmatrix} \in M_5(\mathbb{Q}), \quad p_{\mathbf{A}} = (x-3)^3(x-1)^2.$$

Find the Jordan form $\mathbf{J}$ of $\mathbf{A}$, a Jordan basis, and an invertible $\mathbf{P}$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$. Verify the answer by hand.

Solution. The eigenvalue 3. Put $\mathbf{N} = \mathbf{A} - 3\mathbf{I}$, so

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

Rows 3 and 4 are equal, and so are rows 2 and 5, so the row space is spanned by rows 1, 2, 3. Those three are independent: their entries in columns 1, 2, 4 form the matrix

$\begin{pmatrix} 0 & 1 & 0 \\ -2 & -1 & -2 \\ 1 & 0 & -1 \end{pmatrix}$, of determinant $-4 \neq 0$. Hence $\text{rank } \mathbf{N} = 3$ and nullity $\mathbf{N} = 2$ (Theorem 3.5.7). Solving $\mathbf{N}\mathbf{x} = \mathbf{0}$ means solving

$$x_2 - x_5 = 0, \quad -2x_1 - x_2 + 2x_3 - 2x_4 - x_5 = 0, \quad x_1 - x_3 - x_4 + x_5 = 0.$$

The first gives $x_2 = x_5$; substituting into the second gives $-2x_1 + 2x_3 - 2x_4 - 2x_5 = 0$, that is, $x_1 = x_3 - x_4 - x_5$; and the third then reads $(x_3 - x_4 - x_5) - x_3 - x_4 + x_5 = -2x_4 = 0$, so $x_4 = 0$. Hence

$$E_3(\mathbf{A}) = \ker \mathbf{N} = \text{span}((1, 0, 1, 0, 0), (-1, 1, 0, 0, 1)), \quad g_{\mathbf{A}}(3) = 2.$$

Squaring,

$$\mathbf{N}^2 = 4 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 & 1 \end{pmatrix},$$

whose nonzero rows span a 2-dimensional space, so $\text{rank } \mathbf{N}^2 = 2$. Since $\dim G_3(\mathbf{A}) = a_{\mathbf{A}}(3) = 3$ and $\dim \ker \mathbf{N}^2 = 3$ (Theorem 3.5.7), we have $G_3(\mathbf{A}) = \ker \mathbf{N}^2$, so $\text{rank } \mathbf{N}^3 = \text{rank } \mathbf{N}^2 = 2$. The block sizes for $\lambda = 3$ follow from Theorem 10.3.1 (b) with $r_0, r_1, r_2, r_3 = 5, 3, 2, 2$: blocks of size ≥ 1 : $5 - 3 = 2$; of size ≥ 2 : $3 - 2 = 1$; of size ≥ 3 : $2 - 2 = 0$. So the 3-blocks have sizes 2 and 1.

The chain for 3. We need a top $\mathbf{u} \in \ker \mathbf{N}^2 \setminus \ker \mathbf{N}$. The second column of $\mathbf{N}^2$ is $\mathbf{0}$, so $\mathbf{e}_2 \in \ker \mathbf{N}^2$, while the second column of $\mathbf{N}$ is $(1, -1, 0, 0, -1) \neq \mathbf{0}$, so $\mathbf{e}_2 \notin \ker \mathbf{N}$. Take $\mathbf{u} = \mathbf{e}_2$; its chain is

$$\mathbf{N}\mathbf{u} = (1, -1, 0, 0, -1), \quad \mathbf{u} = \mathbf{e}_2.$$

For the block of size 1 we need a vector of $\ker \mathbf{N}$ completing $\{\mathbf{N}\mathbf{u}\}$ to a basis of $\ker \mathbf{N}$. The vectors $\mathbf{N}\mathbf{u} = (1, -1, 0, 0, -1)$ and $\mathbf{v} = (1, 0, 1, 0, 0)$ are not proportional (their second entries are -1 and 0 , their first entries both 1), so, $\ker \mathbf{N}$ being 2-dimensional, they form a basis of it (Theorem 1.6.4 (a)) and $\mathbf{v}$ does it.

The eigenvalue 1. Put $\mathbf{M} = \mathbf{A} - \mathbf{I}$ and solve $\mathbf{M}\mathbf{x} = \mathbf{0}$. Rows 3 and 4 of $\mathbf{M}$ read $x_1 + x_3 - x_4 + x_5 = 0$ and $x_1 - x_3 + x_4 + x_5 = 0$; adding and subtracting them gives $x_1 + x_5 = 0$ and $x_3 = x_4$. Row 1 reads $2x_1 + x_2 - x_5 = 0$, and row 2, after using $x_3 = x_4$, reads $-2x_1 + x_2 - x_5 = 0$; adding and subtracting these two gives $x_2 = x_5$ and $x_1 = 0$, hence $x_5 = 0$ and $x_2 = 0$. So

$$E_1(\mathbf{A}) = \ker \mathbf{M} = \text{span}((0, 0, 1, 1, 0)), \quad g_{\mathbf{A}}(1) = 1.$$

Since $a_{\mathbf{A}}(1) = 2$, the only partition of 2 with one part is (2): there is a single 1-block, of size 2. Its chain needs a top $\mathbf{w} \in \ker \mathbf{M}^2 \setminus \ker \mathbf{M}$. Computing,

$$\mathbf{M}^2 = 4 \begin{pmatrix} 1 & 1 & 0 & 0 & -1 \\ -1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & -1 & 1 \end{pmatrix},$$

whose kernel is cut out by $x_1 + x_2 - x_5 = 0$, $-x_1 + x_3 - x_4 = 0$, and $x_3 - x_4 = 0$, giving $x_1 = 0$, $x_3 = x_4$, $x_2 = x_5$; so $\ker \mathbf{M}^2 = \text{span}((0, 0, 1, 1, 0), (0, 1, 0, 0, 1))$, of dimension $2 = a_{\mathbf{A}}(1)$ as expected. Take $\mathbf{w} = (0, 1, 0, 0, 1)$; then

$$\begin{aligned}\mathbf{M}\mathbf{w} &= (\text{column 2} + \text{column 5 of } \mathbf{M}) \\ &= (1, 1, 0, 0, -1) + (-1, -1, 1, 1, 1) = (0, 0, 1, 1, 0),\end{aligned}$$

which is the eigenvector, as it must be.

The answer. Writing the three chains one after another,

$$\mathcal{B} = ((1, -1, 0, 0, -1), \mathbf{e}_2, (1, 0, 1, 0, 0), (0, 0, 1, 1, 0), (0, 1, 0, 0, 1)),$$

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

$$\mathbf{J} = \mathbf{J}_2(3) \oplus \mathbf{J}_1(3) \oplus \mathbf{J}_2(1) = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\mathbf{P}$ is invertible: its fourth row is $(0, 0, 0, 1, 0)$, so expanding $\det \mathbf{P}$ along that row (Theorem 7.6.1) leaves

$$\det \mathbf{P} = \det \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} = 1,$$

expanding along the third row and then along the first row.

The check, column by column. The claim $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$ is the same as $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{J}$. Writing $\mathbf{p}_1, \dots, \mathbf{p}_5$ for the columns of $\mathbf{P}$, the five columns of $\mathbf{P}\mathbf{J}$ are

$$3\mathbf{p}_1, \quad \mathbf{p}_1 + 3\mathbf{p}_2, \quad 3\mathbf{p}_3, \quad \mathbf{p}_4, \quad \mathbf{p}_4 + \mathbf{p}_5,$$

read straight off the columns of $\mathbf{J}$. So the identity $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{J}$ is the list of five statements

$$\mathbf{A}\mathbf{p}_1 = 3\mathbf{p}_1, \quad \mathbf{A}\mathbf{p}_2 = \mathbf{p}_1 + 3\mathbf{p}_2, \quad \mathbf{A}\mathbf{p}_3 = 3\mathbf{p}_3, \quad \mathbf{A}\mathbf{p}_4 = \mathbf{p}_4, \quad \mathbf{A}\mathbf{p}_5 = \mathbf{p}_4 + \mathbf{p}_5,$$

which are exactly how the chains were constructed: $\mathbf{p}_1, \mathbf{p}_3 \in \ker(\mathbf{A} - 3\mathbf{I})$, $(\mathbf{A} - 3\mathbf{I})\mathbf{p}_2 = \mathbf{p}_1$, $\mathbf{p}_4 \in \ker(\mathbf{A} - \mathbf{I})$, and $(\mathbf{A} - \mathbf{I})\mathbf{p}_5 = \mathbf{p}_4$. Two of them verified in full:

$$\mathbf{A}\mathbf{p}_2 = \mathbf{A}\mathbf{e}_2 = (1, 2, 0, 0, -1) = (1, -1, 0, 0, -1) + 3(0, 1, 0, 0, 0) = \mathbf{p}_1 + 3\mathbf{p}_2,$$

reading the second column of $\mathbf{A}$; and

$$\begin{aligned}\mathbf{A}\mathbf{p}_4 &= \mathbf{A}(0, 0, 1, 1, 0) = (0 + 0, 2 - 2, 2 - 1, -1 + 2, 2 - 2) \\ &= (0, 0, 1, 1, 0) = \mathbf{p}_4,\end{aligned}$$

adding the third and fourth columns of $\mathbf{A}$. The remaining three are the same kind of one-line computation.

Cross-checks. The diagonal of $\mathbf{J}$ gives $\text{tr } \mathbf{A} = 3 + 3 + 3 + 1 + 1 = 11$, and adding the diagonal of $\mathbf{A}$ gives $3 + 2 + 2 + 2 + 2 = 11$. By [Corollary 10.3.2](#), $a(3) = 2 + 1 = 3$, $g(3) = 2$, $a(1) = 2$, $g(1) = 1$, and $m_{\mathbf{A}} = (x - 3)^2(x - 1)^2$, all matching the computations above.

10.3.6 Exercises

10.3.6.1 A. Check your understanding

Exercise. A1

- Define a Jordan block and a Jordan matrix, and state what a Jordan form of T is.
- State the Jordan Canonical Form Theorem, including the uniqueness clause.
- How do you read $a(\lambda)$, $g(\lambda)$, and the exponent of $x - \lambda$ in m_T off a Jordan form?
- True or false: every matrix in $M_n(\mathbb{R})$ has a Jordan form over $\mathbb{R}$. Justify your answer.
- True or false: if $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic and the same minimal polynomial, they are similar. Justify your answer.
- Is $\begin{pmatrix} 5 & 1 \\ 0 & 4 \end{pmatrix}$ a Jordan matrix? What is its Jordan form?

Solution.

- $\mathbf{J}_k(\lambda) = \lambda \mathbf{I}_k + \mathbf{J}_k(0)$ has λ on the diagonal and 1 on the superdiagonal ([Definition 10.3.1](#)); a Jordan matrix is a direct sum of Jordan blocks; a Jordan form of T is a Jordan matrix equal to $[T]_{\mathcal{B}}^{\mathcal{B}}$ for some basis $\mathcal{B}$ ([Definition 10.3.2](#)).
- T has a Jordan form if and only if p_T splits, and then the number of λ -blocks of size at least j is $\text{rank}(T - \lambda \text{id}_V)^{j-1} - \text{rank}(T - \lambda \text{id}_V)^j$, so the blocks are unique up to their order ([Theorem 10.3.1](#)).
- $a(\lambda)$ is the total size of the λ -blocks, $g(\lambda)$ their number, and the exponent in m_T the largest of their sizes ([Corollary 10.3.2](#)).
- False. A Jordan matrix is triangular, so its existence forces p_T to split; the rotation by a right angle has $p = x^2 + 1$, which does not split over $\mathbb{R}$.
- False. $\mathbf{J}_2(1) \oplus \mathbf{J}_2(1)$ and $\mathbf{J}_2(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(1)$ share $p = (x - 1)^4$ and $m = (x - 1)^2$ but have different Jordan forms, hence are not similar ([Corollary 10.3.3](#)).
- No: a superdiagonal 1 may only stand between two equal diagonal entries. The eigenvalues 5 and 4 are distinct, so the matrix is diagonalizable ([Corollary 9.4.2](#)) and its Jordan form is $\text{diag}(5, 4)$.

10.3.6.2 B. Practice

Exercise. B1: Jordan forms from polynomial data

In each case T is an operator on a complex vector space with the given characteristic and minimal polynomials. List all possible Jordan forms, up to the order of the blocks, and say

what extra data would pin the answer down.

- (a) $p_T = (x - 2)^4$, $m_T = (x - 2)^2$.
- (b) $p_T = (x - 1)^3(x + 1)^2$, $m_T = (x - 1)^3(x + 1)$.
- (c) $p_T = (x - 5)^5$, $m_T = (x - 5)^3$.

Solution. By [Corollary 10.3.2](#), the λ -block sizes form a partition of $a(\lambda)$ whose largest part is the exponent of $x - \lambda$ in m_T .

- (a) Partitions of 4 with largest part 2: $(2, 2)$ and $(2, 1, 1)$. So $\mathbf{J}_2(2) \oplus \mathbf{J}_2(2)$ or $\mathbf{J}_2(2) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(2)$. The value of $g(2)$ — that is, 2 against 3 — decides.
- (b) For 1: partitions of 3 with largest part 3, only (3) . For -1 : partitions of 2 with largest part 1, only $(1, 1)$. So the Jordan form is $\mathbf{J}_3(1) \oplus \mathbf{J}_1(-1) \oplus \mathbf{J}_1(-1)$, and no extra data is needed.
- (c) Partitions of 5 with largest part 3: $(3, 2)$ and $(3, 1, 1)$. So $\mathbf{J}_3(5) \oplus \mathbf{J}_2(5)$ or $\mathbf{J}_3(5) \oplus \mathbf{J}_1(5) \oplus \mathbf{J}_1(5)$, told apart by $g(5) = 2$ against 3, or by $\dim \ker(T - 5 \text{ id})^2 = 4$ against 5.

Exercise. B2: Jordan forms and Jordan bases

For each matrix, find the Jordan form $\mathbf{J}$ and an invertible $\mathbf{P}$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$. Verify your $\mathbf{P}$ by checking $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{J}$ column by column.

- (a) $\mathbf{A}_1 = \begin{pmatrix} -1 & -1 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \in M_3(\mathbb{Q})$, with $p_{\mathbf{A}_1} = (x - 1)^3$.
- (b) $\mathbf{A}_2 = \begin{pmatrix} 2 & -1 & -1 \\ -3 & 4 & -1 \\ 3 & -2 & 3 \end{pmatrix} \in M_3(\mathbb{Q})$, with $p_{\mathbf{A}_2} = (x - 2)^2(x - 5)$.
- (c) $\mathbf{A}_3 = \begin{pmatrix} 2 & 1 & 1 & -2 \\ -3 & 1 & -1 & -1 \\ -3 & -1 & 1 & -1 \\ -3 & -1 & -1 & 1 \end{pmatrix} \in M_4(\mathbb{Q})$, with $p_{\mathbf{A}_3} = (x - 2)^3(x + 1)$.

Solution.

- (a) $\mathbf{A}_1 - \mathbf{I} = \begin{pmatrix} -2 & -1 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ has rank 2, so $g(1) = 1$ and there is one block, of size 3:

$\mathbf{J} = \mathbf{J}_3(1)$. A top must satisfy $(\mathbf{A}_1 - \mathbf{I})^2 \mathbf{u} \neq \mathbf{0}$, and $(\mathbf{A}_1 - \mathbf{I})^2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$ has

nonzero first column, so $\mathbf{u} = \mathbf{e}_1$ works. The chain is

$$(\mathbf{A}_1 - \mathbf{I})^2 \mathbf{e}_1 = (1, 0, -1), \quad (\mathbf{A}_1 - \mathbf{I}) \mathbf{e}_1 = (-2, 1, 1), \quad \mathbf{e}_1,$$

so $\mathbf{P} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$, with $\det \mathbf{P} = 1$. The check $\mathbf{A}_1 \mathbf{P} = \mathbf{P} \mathbf{J}_3(1)$ reads $\mathbf{A}_1 \mathbf{p}_1 = \mathbf{p}_1$,

$\mathbf{A}_1 \mathbf{p}_2 = \mathbf{p}_1 + \mathbf{p}_2$, $\mathbf{A}_1 \mathbf{p}_3 = \mathbf{p}_2 + \mathbf{p}_3$, which are the three chain relations.

(b) $\mathbf{A}_2 - 2\mathbf{I} = \begin{pmatrix} 0 & -1 & -1 \\ -3 & 2 & -1 \\ 3 & -2 & 1 \end{pmatrix}$ has rank 2, so $g(2) = 1$ and the single 2-block has size

$a(2) = 2$; the eigenvalue 5 gives one block of size 1. So $\mathbf{J} = \mathbf{J}_2(2) \oplus \mathbf{J}_1(5)$. For the chain, the third column of $\mathbf{A}_2 - 2\mathbf{I}$ is $(-1, -1, 1) \neq \mathbf{0}$ and $(\mathbf{A}_2 - 2\mathbf{I})(-1, -1, 1) = \mathbf{0}$, so $\mathbf{u} = \mathbf{e}_3$ is a top with $(\mathbf{A}_2 - 2\mathbf{I})\mathbf{e}_3 = (-1, -1, 1)$. Solving $(\mathbf{A}_2 - 5\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\text{span}((0, -1, 1))$. Hence

$$\mathbf{P} = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \det \mathbf{P} = -1,$$

and $\mathbf{A}_2 \mathbf{P} = \mathbf{P} \mathbf{J}$ reads $\mathbf{A}_2 \mathbf{p}_1 = 2\mathbf{p}_1$, $\mathbf{A}_2 \mathbf{p}_2 = \mathbf{p}_1 + 2\mathbf{p}_2$, $\mathbf{A}_2 \mathbf{p}_3 = 5\mathbf{p}_3$.

(c) $\mathbf{A}_3 - 2\mathbf{I} = \begin{pmatrix} 0 & 1 & 1 & -2 \\ -3 & -1 & -1 & -1 \\ -3 & -1 & -1 & -1 \\ -3 & -1 & -1 & -1 \end{pmatrix}$ has rank 2 (rows 2, 3, 4 are equal), so $g(2) = 2$:

with $a(2) = 3$ the 2-blocks have sizes 2 and 1. The eigenvalue -1 has $a = 1$, so $\mathbf{J} = \mathbf{J}_2(2) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(-1)$. Now $\ker(\mathbf{A}_3 - 2\mathbf{I}) = \text{span}((0, -1, 1, 0), (-1, 2, 0, 1))$, and the second column of $\mathbf{A}_3 - 2\mathbf{I}$ is $(1, -1, -1, -1)$, which is nonzero and lies in that kernel, so $\mathbf{u} = \mathbf{e}_2$ is the top of the 2-chain. Completing $\{(1, -1, -1, -1)\}$ to a basis of $\ker(\mathbf{A}_3 - 2\mathbf{I})$ with $(0, -1, 1, 0)$, and solving $(\mathbf{A}_3 + \mathbf{I})\mathbf{x} = \mathbf{0}$ to get $\text{span}((0, 1, 1, 1))$, we obtain

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad \det \mathbf{P} = 1,$$

and $\mathbf{A}_3 \mathbf{P} = \mathbf{P} \mathbf{J}$ reads $\mathbf{A}_3 \mathbf{p}_1 = 2\mathbf{p}_1$, $\mathbf{A}_3 \mathbf{p}_2 = \mathbf{p}_1 + 2\mathbf{p}_2$, $\mathbf{A}_3 \mathbf{p}_3 = 2\mathbf{p}_3$, $\mathbf{A}_3 \mathbf{p}_4 = -\mathbf{p}_4$. For instance the second of these is $\mathbf{A}_3 \mathbf{e}_2 = (1, 1, -1, -1) = (1, -1, -1, -1) + 2(0, 1, 0, 0)$, read off the second column of $\mathbf{A}_3$.

Exercise. B3: Similar or not?

Determine whether each pair of matrices is similar. Justify your answer.

- (a) $\mathbf{J}_2(4) \oplus \mathbf{J}_1(4)$ and $\mathbf{J}_1(4) \oplus \mathbf{J}_2(4)$.
- (b) $\mathbf{J}_3(0) \oplus \mathbf{J}_1(0)$ and $\mathbf{J}_2(0) \oplus \mathbf{J}_2(0)$.
- (c) $\mathbf{J}_2(1) \oplus \mathbf{J}_2(1)$ and $\mathbf{J}_2(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(1)$.
- (d) $\text{diag}(2, 2, 5)$ and the matrix $\mathbf{A}_2$ of [Exercise \(B2: Jordan forms and Jordan bases\)](#) (b).

Solution.

- (a) Similar: the blocks are the same and only their order differs, which is a change of basis, as noted after [Definition 10.3.2](#).

- (b) Not similar. Both are nilpotent of rank 2, but $\text{rank}(\mathbf{J}_3(0) \oplus \mathbf{J}_1(0))^2 = 1$ while $(\mathbf{J}_2(0) \oplus \mathbf{J}_2(0))^2 = \mathbf{O}$ (Theorem 8.1.4 (b)). By [Corollary 10.3.3](#) (or Corollary 10.2.5) they are not similar; their partitions are $(3, 1)$ and $(2, 2)$.
- (c) Not similar, although p and m agree, as in the remark after [Corollary 10.3.3](#): $\text{rank}(\mathbf{A} - \mathbf{I}) = 2$ for the first and 1 for the second, so $g(1)$ is 2 against 3.
- (d) Not similar. $\text{diag}(2, 2, 5)$ is diagonal, so $g(2) = 2$, while $\mathbf{A}_2$ has $g(2) = 1$ by [Exercise \(B2: Jordan forms and Jordan bases\)](#) (b); its Jordan form is $\mathbf{J}_2(2) \oplus \mathbf{J}_1(5)$, which is not diagonal. The characteristic polynomials agree, so this pair is another witness that $p_{\mathbf{A}}$ alone decides nothing.

10.3.6.3 C. Going deeper

Exercise. C1: A matrix is similar to its transpose

Let $\mathbf{A} \in M_n(F)$ with $p_{\mathbf{A}}$ split over F .

- (a) Let $\mathbf{R} \in M_k(F)$ be the reversal permutation matrix, with 1 in each position $(j, k+1-j)$ and 0 elsewhere. Prove that $\mathbf{R}^{-1} = \mathbf{R}$ and $\mathbf{R}^{-1}\mathbf{J}_k(\lambda)\mathbf{R} = \mathbf{J}_k(\lambda)^{\top}$.
- (b) Deduce that every Jordan matrix $\mathbf{J}$ satisfies $\mathbf{J} \sim \mathbf{J}^{\top}$.
- (c) Deduce that $\mathbf{A} \sim \mathbf{A}^{\top}$.

Hint: for (c), transpose a similarity $\mathbf{A} = \mathbf{PJP}^{-1}$. The conclusion holds over every field; Section 7 removes the hypothesis that $p_{\mathbf{A}}$ splits.

Solution.

- (a) $\mathbf{R}\mathbf{e}_j = \mathbf{e}_{k+1-j}$, so $\mathbf{R}^2\mathbf{e}_j = \mathbf{e}_j$ for every j and $\mathbf{R}^2 = \mathbf{I}_k$, that is, $\mathbf{R}^{-1} = \mathbf{R}$. Since $\mathbf{J}_k(\lambda) = \lambda\mathbf{I}_k + \mathbf{J}_k(0)$ and $\mathbf{R}^{-1}(\lambda\mathbf{I}_k)\mathbf{R} = \lambda\mathbf{I}_k$, it suffices to prove $\mathbf{R}^{-1}\mathbf{J}_k(0)\mathbf{R} = \mathbf{J}_k(0)^{\top}$. Both sides are computed on basis vectors: $\mathbf{J}_k(0)\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{J}_k(0)\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \geq 2$, so

$$\mathbf{R}\mathbf{J}_k(0)\mathbf{R}\mathbf{e}_j = \mathbf{R}\mathbf{J}_k(0)\mathbf{e}_{k+1-j} = \begin{cases} \mathbf{R}\mathbf{e}_{k-j} = \mathbf{e}_{j+1}, & j \leq k-1, \\ \mathbf{R}\mathbf{0} = \mathbf{0}, & j = k, \end{cases}$$

which is exactly the action of $\mathbf{J}_k(0)^{\top}$, whose columns satisfy $\mathbf{J}_k(0)^{\top}\mathbf{e}_j = \mathbf{e}_{j+1}$ for $j \leq k-1$ and $\mathbf{J}_k(0)^{\top}\mathbf{e}_k = \mathbf{0}$. Hence $\mathbf{R}^{-1}\mathbf{J}_k(\lambda)\mathbf{R} = \mathbf{J}_k(\lambda)^{\top}$.

- (b) Let $\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \cdots \oplus \mathbf{J}_{k_m}(\lambda_m)$ and let $\mathbf{Q} = \mathbf{R}_{k_1} \oplus \cdots \oplus \mathbf{R}_{k_m}$, the direct sum of the reversal matrices of the block sizes. Then $\mathbf{Q}$ is invertible with $\mathbf{Q}^{-1} = \mathbf{R}_{k_1}^{-1} \oplus \cdots \oplus \mathbf{R}_{k_m}^{-1}$ (Theorem 8.1.4 (c)), and multiplying blockwise (Theorem 8.1.4 (a)),

$$\mathbf{Q}^{-1}\mathbf{JQ} = (\mathbf{R}_{k_1}^{-1}\mathbf{J}_{k_1}(\lambda_1)\mathbf{R}_{k_1}) \oplus \cdots = \mathbf{J}_{k_1}(\lambda_1)^{\top} \oplus \cdots \oplus \mathbf{J}_{k_m}(\lambda_m)^{\top} = \mathbf{J}^{\top},$$

the last equality because transposing a block diagonal matrix transposes each block (Proposition 8.1.2).

- (c) By [Theorem 10.3.1](#) there are a Jordan matrix $\mathbf{J}$ and an invertible $\mathbf{P}$ with $\mathbf{A} = \mathbf{PJP}^{-1}$. Transposing and using $(\mathbf{XY})^{\top} = \mathbf{Y}^{\top}\mathbf{X}^{\top}$ and $(\mathbf{P}^{-1})^{\top} = (\mathbf{P}^{\top})^{-1}$ (Theorem 0.9.5, Theorem 0.9.7),

$$\mathbf{A}^{\top} = (\mathbf{P}^{-1})^{\top}\mathbf{J}^{\top}\mathbf{P}^{\top} = (\mathbf{P}^{\top})^{-1}\mathbf{J}^{\top}\mathbf{P}^{\top},$$

so $\mathbf{A}^\top \sim \mathbf{J}^\top$. By (b), $\mathbf{J}^\top \sim \mathbf{J} \sim \mathbf{A}$, and similarity is an equivalence relation (Example (Similarity)), so $\mathbf{A} \sim \mathbf{A}^\top$.

Exercise. C2: How much do p and m decide?

Work over $\mathbb{C}$.

- (a) Find all Jordan forms with $p_T = (x-1)^3(x-2)^2$ and $m_T = (x-1)^2(x-2)$.
- (b) Find all Jordan forms with $p_T = (x-1)^4(x-2)^2$ and $m_T = (x-1)^2(x-2)$.
- (c) In (b), what is the least amount of extra information that determines the Jordan form?
- (d) Prove that p_T and m_T determine the Jordan form whenever $a(\lambda) \leq 3$ for every eigenvalue λ .

Solution.

- (a) For $\lambda = 1$: partitions of 3 with largest part 2, so $(2, 1)$ only. For $\lambda = 2$: partitions of 2 with largest part 1, so $(1, 1)$ only. There is exactly one Jordan form,

$$\mathbf{J}_2(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(2).$$

- (b) For $\lambda = 1$: partitions of 4 with largest part 2, namely $(2, 2)$ and $(2, 1, 1)$. For $\lambda = 2$: $(1, 1)$ as before. So there are two Jordan forms,

$$\begin{aligned} & \mathbf{J}_2(1) \oplus \mathbf{J}_2(1) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(2) \\ \text{and } & \mathbf{J}_2(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(1) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(2). \end{aligned}$$

- (c) The single number $g(1)$, which is 2 for the first and 3 for the second; equivalently $\text{rank}(T - \text{id}_V)$ or $\dim \ker(T - \text{id}_V)$ ([Corollary 10.3.2](#) (b)).
- (d) Fix λ with $a = a(\lambda) \leq 3$ and let s be the exponent of $x - \lambda$ in m_T , which is the largest block size ([Corollary 10.3.2](#) (c)). The λ -blocks form a partition of a with largest part s . For $a = 1$ the only partition is (1) . For $a = 2$: $s = 2$ gives (2) , and $s = 1$ gives $(1, 1)$. For $a = 3$: $s = 3$ gives (3) , $s = 2$ gives $(2, 1)$, and $s = 1$ gives $(1, 1, 1)$. In every case the pair (a, s) determines the partition, so the blocks for each eigenvalue, and hence the whole Jordan form, are determined by p_T and m_T . The first ambiguity appears at $a = 4$, $s = 2$, which is part (a) of [Exercise \(B1: Jordan forms from polynomial data\)](#).