(c) Decide if these are similar.

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

1.32 **(a)** Show that a matrix is invertible if and only if the constant term in its minimal polynomial is not 0.
(b) Show that if a square matrix T is not invertible then there is a nonzero matrix S such that ST and TS both equal the zero matrix.

✓ **1.33** **(a)** Finish the proof of Lemma 1.7.
(b) Give an example to show that the result does not hold if t is not linear.

5.IV.2   Jordan Canonical Form

This subsection moves from the canonical form for nilpotent matrices to the one for all matrices.

We have shown that if a map is nilpotent then all of its eigenvalues are zero. We can now prove the converse.

2.1 Lemma   A linear transformation whose only eigenvalue is zero is nilpotent.

PROOF. If a transformation t on an n-dimensional space has only the single eigenvalue of zero then its characteristic polynomial is x^n. The Cayley-Hamilton Theorem says that a map satisfies its characteristic polynimial so t^n is the zero map. Thus t is nilpotent.   QED

We have a canonical form for nilpotent matrices, that is, for each matrix whose single eigenvalue is zero: each such matrix is similar to one that is all zeroes except for blocks of subdiagonal ones. (To make this representation unique we can fix some arrangement of the blocks, say, from longest to shortest.) We next extend this to all single-eigenvalue matrices.

Observe that if t's only eigenvalue is λ then $t - \lambda$'s only eigenvalue is 0 because $t(\vec{v}) = \lambda\vec{v}$ if and only if $(t - \lambda)(\vec{v}) = 0 \cdot \vec{v}$. The natural way to extend the results for nilpotent matrices is to represent $t - \lambda$ in the canonical form N, and try to use that to get a simple representation T for t. The next result says that this try works.

2.2 Lemma   If the matrices $T - \lambda I$ and N are similar then T and $N + \lambda I$ are also similar, via the same change of basis matrices.

PROOF. With $N = P(T - \lambda I)P^{-1} = PTP^{-1} - P(\lambda I)P^{-1}$ we have $N = PTP^{-1} - PP^{-1}(\lambda I)$ since the diagonal matrix λI commutes with anything, and so $N = PTP^{-1} - \lambda I$. Therefore $N + \lambda I = PTP^{-1}$, as required.   QED

2.3 Example   The characteristic polynomial of

$$T = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$$

is $(x-3)^2$ and so T has only the single eigenvalue 3. Thus for

$$T - 3I = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$$

the only eigenvalue is 0, and $T - 3I$ is nilpotent. The null spaces are routine to find; to ease this computation we take T to represent the transformation $t: \mathbb{C}^2 \to \mathbb{C}^2$ with respect to the standard basis (we shall maintain this convention for the rest of the chapter).

$$\mathscr{N}(t-3) = \{\begin{pmatrix} -y \\ y \end{pmatrix} \mid y \in \mathbb{C}\} \qquad \mathscr{N}((t-3)^2) = \mathbb{C}^2$$

The dimensions of these null spaces show that the action of an associated map $t - 3$ on a string basis is $\vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{0}$. Thus, the canonical form for $t - 3$ with one choice for a string basis is

$$\mathrm{Rep}_{B,B}(t-3) = N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad B = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix} \rangle$$

and by Lemma 2.2, T is similar to this matrix.

$$\mathrm{Rep}_t(B, B) = N + 3I = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$$

We can produce the similarity computation. Recall from the Nilpotence section how to find the change of basis matrices P and P^{-1} to express N as $P(T - 3I)P^{-1}$. The similarity diagram

$$
\begin{array}{ccc}
\mathbb{C}^2_{\text{w.r.t. } \mathcal{E}_2} & \xrightarrow[T-3I]{t-3} & \mathbb{C}^2_{\text{w.r.t. } \mathcal{E}_2} \\
\text{id} \downarrow P & & \text{id} \downarrow P \\
\mathbb{C}^2_{\text{w.r.t. } B} & \xrightarrow[N]{t-3} & \mathbb{C}^2_{\text{w.r.t. } B}
\end{array}
$$

describes that to move from the lower left to the upper left we multiply by

$$P^{-1} = \left(\mathrm{Rep}_{\mathcal{E}_2, B}(\mathrm{id})\right)^{-1} = \mathrm{Rep}_{B, \mathcal{E}_2}(\mathrm{id}) = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}$$

and to move from the upper right to the lower right we multiply by this matrix.

$$P = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ -1/4 & 1/4 \end{pmatrix}$$

So the similarity is expressed by

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

which is easily checked.

2.4 Example This matrix has characteristic polynomial $(x-4)^4$

$$T = \begin{pmatrix} 4 & 1 & 0 & -1 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 1 & 0 & 0 & 5 \end{pmatrix}$$

and so has the single eigenvalue 4. The nullities of $t-4$ are: the null space of $t-4$ has dimension two, the null space of $(t-4)^2$ has dimension three, and the null space of $(t-4)^3$ has dimension four. Thus, $t-4$ has the action on a string basis of $\vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{\beta}_3 \mapsto \vec{0}$ and $\vec{\beta}_4 \mapsto \vec{0}$. This gives the canonical form N for $t-4$, which in turn gives the form for t.

$$N + 4I = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

An array that is all zeroes, except for some number λ down the diagonal and blocks of subdiagonal ones, is a *Jordan block*. We have shown that Jordan block matrices are canonical representatives of the similarity classes of single-eigenvalue matrices.

2.5 Example The 3×3 matrices whose only eigenvalue is $1/2$ separate into three similarity classes. The three classes have these canonical representatives.

$$\begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \qquad \begin{pmatrix} 1/2 & 0 & 0 \\ 1 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \qquad \begin{pmatrix} 1/2 & 0 & 0 \\ 1 & 1/2 & 0 \\ 0 & 1 & 1/2 \end{pmatrix}$$

In particular, this matrix

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

belongs to the similarity class represented by the middle one, because we have adopted the convention of ordering the blocks of subdiagonal ones from the longest block to the shortest.

We will now finish the program of this chapter by extending this work to cover maps and matrices with multiple eigenvalues. The best possibility for general maps and matrices would be if we could break them into a part involving their first eigenvalue λ_1 (which we represent using its Jordan block), a part with λ_2, etc.

This ideal is in fact what happens. For any transformation $t : V \to V$, we shall break the space V into the direct sum of a part on which $t - \lambda_1$ is nilpotent, plus a part on which $t - \lambda_2$ is nilpotent, etc. More precisely, we shall take three

steps to get to this section's major theorem and the third step shows that $V = \mathscr{N}_\infty(t - \lambda_1) \oplus \cdots \oplus \mathscr{N}_\infty(t - \lambda_\ell)$ where $\lambda_1, \ldots, \lambda_\ell$ are t's eigenvalues.

Suppose that $t\colon V \to V$ is a linear transformation. Note that the restriction[*] of t to a subspace M need not be a linear transformation on M because there may be an $\vec{m} \in M$ with $t(\vec{m}) \notin M$. To ensure that the restriction of a transformation to a 'part' of a space is a transformation on the partwe need the next condition.

2.6 Definition Let $t\colon V \to V$ be a transformation. A subspace M is t *invariant* if whenever $\vec{m} \in M$ then $t(\vec{m}) \in M$ (shorter: $t(M) \subseteq M$).

Two examples are that the generalized null space $\mathscr{N}_\infty(t)$ and the generalized range space $\mathscr{R}_\infty(t)$ of any transformation t are invariant. For the generalized null space, if $\vec{v} \in \mathscr{N}_\infty(t)$ then $t^n(\vec{v}) = \vec{0}$ where n is the dimension of the underlying space and so $t(\vec{v}) \in \mathscr{N}_\infty(t)$ because $t^n(t(\vec{v}))$ is zero also. For the generalized range space, if $\vec{v} \in \mathscr{R}_\infty(t)$ then $\vec{v} = t^n(\vec{w})$ for some $\vec{w}$ and then $t(\vec{v}) = t^{n+1}(\vec{w}) = t^n(t(\vec{w}))$ shows that $t(\vec{v})$ is also a member of $\mathscr{R}_\infty(t)$.

Thus the spaces $\mathscr{N}_\infty(t - \lambda_i)$ and $\mathscr{R}_\infty(t - \lambda_i)$ are $t - \lambda_i$ invariant. Observe also that $t - \lambda_i$ is nilpotent on $\mathscr{N}_\infty(t - \lambda_i)$ because, simply, if $\vec{v}$ has the property that some power of $t - \lambda_i$ maps it to zero — that is, if it is in the generalized null space — then some power of $t - \lambda_i$ maps it to zero. The generalized null space $\mathscr{N}_\infty(t - \lambda_i)$ is a 'part' of the space on which the action of $t - \lambda_i$ is easy to understand.

The next result is the first of our three steps. It establishes that $t - \lambda_j$ leaves $t - \lambda_i$'s part unchanged.

2.7 Lemma A subspace is t invariant if and only if it is $t - \lambda$ invariant for any scalar λ. In particular, where λ_i is an eigenvalue of a linear transformation t, then for any other eigenvalue λ_j, the spaces $\mathscr{N}_\infty(t - \lambda_i)$ and $\mathscr{R}_\infty(t - \lambda_i)$ are $t - \lambda_j$ invariant.

PROOF. For the first sentence we check the two implications of the 'if and only if' separately. One of them is easy: if the subspace is $t - \lambda$ invariant for any λ then taking $\lambda = 0$ shows that it is t invariant. For the other implication suppose that the subspace is t invariant, so that if $\vec{m} \in M$ then $t(\vec{m}) \in M$, and let λ be any scalar. The subspace M is closed under linear combinations and so if $t(\vec{m}) \in M$ then $t(\vec{m}) - \lambda\vec{m} \in M$. Thus if $\vec{m} \in M$ then $(t - \lambda)(\vec{m}) \in M$, as required.

The second sentence follows straight from the first. Because the two spaces are $t - \lambda_i$ invariant, they are therefore t invariant. From this, applying the first sentence again, we conclude that they are also $t - \lambda_j$ invariant. QED

The second step of the three that we will take to prove this section's major result makes use of an additional property of $\mathscr{N}_\infty(t - \lambda_i)$ and $\mathscr{R}_\infty(t - \lambda_i)$, that they are complementary. Recall that if a space is the direct sum of two others $V = \mathscr{N} \oplus \mathscr{R}$ then any vector $\vec{v}$ in the space breaks into two parts $\vec{v} = \vec{n} + \vec{r}$ where $\vec{n} \in \mathscr{N}$ and $\vec{r} \in \mathscr{R}$, and recall also that if $B_\mathscr{N}$ and $B_\mathscr{R}$ are bases for $\mathscr{N}$

---

[*] More information on restrictions of functions is in the appendix.

and $\mathscr{R}$ then the concatenation $B_{\mathscr{N}} \frown B_{\mathscr{R}}$ is linearly independent (and so the two parts of $\vec{v}$ do not "overlap"). The next result says that for any subspaces $\mathscr{N}$ and $\mathscr{R}$ that are complementary as well as t invariant, the action of t on $\vec{v}$ breaks into the "non-overlapping" actions of t on $\vec{n}$ and on $\vec{r}$.

2.8 Lemma Let $t\colon V \to V$ be a transformation and let $\mathscr{N}$ and $\mathscr{R}$ be t invariant complementary subspaces of V. Then t can be represented by a matrix with blocks of square submatrices T_1 and T_2

$$\left(\begin{array}{c|c} T_1 & Z_2 \\ \hline Z_1 & T_2 \end{array}\right) \quad \begin{array}{l} \}\dim(\mathscr{N})\text{-many rows} \\ \}\dim(\mathscr{R})\text{-many rows} \end{array}$$

where Z_1 and Z_2 are blocks of zeroes.

PROOF. Since the two subspaces are complementary, the concatenation of a basis for $\mathscr{N}$ and a basis for $\mathscr{R}$ makes a basis $B = \langle \vec{\nu}_1, \ldots, \vec{\nu}_p, \vec{\mu}_1, \ldots, \vec{\mu}_q \rangle$ for V. We shall show that the matrix

$$\mathrm{Rep}_{B,B}(t) = \left(\begin{array}{c|c|c|c} & \vdots & & \vdots \\ \mathrm{Rep}_B(t(\vec{\nu}_1)) & & \cdots & \mathrm{Rep}_B(t(\vec{\mu}_q)) \\ & \vdots & & \vdots \end{array}\right)$$

has the desired form.

Any vector $\vec{v} \in V$ is in $\mathscr{N}$ if and only if its final q components are zeroes when it is represented with respect to B. As $\mathscr{N}$ is t invariant, each of the vectors $\mathrm{Rep}_B(t(\vec{\nu}_1))$, $\ldots$, $\mathrm{Rep}_B(t(\vec{\nu}_p))$ has that form. Hence the lower left of $\mathrm{Rep}_{B,B}(t)$ is all zeroes.

The argument for the upper right is similar. QED

To see that t has been decomposed into its action on the parts, observe that the restrictions of t to the subspaces $\mathscr{N}$ and $\mathscr{R}$ are represented, with respect to the obvious bases, by the matrices T_1 and T_2. So, with subspaces that are invariant and complementary, we can split the problem of examining a linear transformation into two lower-dimensional subproblems. The next result illustrates this decomposition into blocks.

2.9 Lemma If T is a matrices with square submatrices T_1 and T_2

$$T = \left(\begin{array}{c|c} T_1 & Z_2 \\ \hline Z_1 & T_2 \end{array}\right)$$

where the Z's are blocks of zeroes, then $|T| = |T_1| \cdot |T_2|$.

PROOF. Suppose that T is $n \times n$, that T_1 is $p \times p$, and that T_2 is $q \times q$. In the permutation formula for the determinant

$$|T| = \sum_{\text{permutations } \phi} t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} \, \mathrm{sgn}(\phi)$$

each term comes from a rearrangement of the column numbers $1, \ldots, n$ into a new order $\phi(1), \ldots, \phi(n)$. The upper right block Z_2 is all zeroes, so if a ϕ has at least one of $p + 1, \ldots, n$ among its first p column numbers $\phi(1), \ldots, \phi(p)$ then the term arising from ϕ is zero, e.g., if $\phi(1) = n$ then $t_{1,\phi(1)} t_{2,\phi(2)} \ldots t_{n,\phi(n)} = 0 \cdot t_{2,\phi(2)} \ldots t_{n,\phi(n)} = 0$.

So the above formula reduces to a sum over all permutations with two halves: any significant ϕ is the composition of a ϕ_1 that rearranges only $1, \ldots, p$ and a ϕ_2 that rearranges only $p + 1, \ldots, p + q$. Now, the distributive law (and the fact that the signum of a composition is the product of the signums) gives that this

$$|T_1| \cdot |T_2| = \left(\sum_{\substack{\text{perms } \phi_1 \\ \text{of } 1, \ldots, p}} t_{1,\phi_1(1)} \cdots t_{p,\phi_1(p)} \, \text{sgn}(\phi_1) \right)$$

$$\cdot \left(\sum_{\substack{\text{perms } \phi_2 \\ \text{of } p+1, \ldots, p+q}} t_{p+1,\phi_2(p+1)} \cdots t_{p+q,\phi_2(p+q)} \, \text{sgn}(\phi_2) \right)$$

equals $|T| = \sum_{\text{significant } \phi} t_{1,\phi(1)} t_{2,\phi(2)} \cdots t_{n,\phi(n)} \, \text{sgn}(\phi)$. \hfill QED

2.10 Example

$$\begin{vmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 1 & 2 \end{vmatrix} \cdot \begin{vmatrix} 3 & 0 \\ 0 & 3 \end{vmatrix} = 36$$

From Lemma 2.9 we conclude that if two subspaces are complementary and t invariant then t is nonsingular if and only if its restrictions to both subspaces are nonsingular.

Now for the promised third, final, step to the main result.

2.11 Lemma If a linear transformation $t \colon V \to V$ has the characteristic polynomial $(x - \lambda_1)^{p_1} \ldots (x - \lambda_\ell)^{p_\ell}$ then (1) $V = \mathscr{N}_\infty(t - \lambda_1) \oplus \cdots \oplus \mathscr{N}_\infty(t - \lambda_\ell)$ and (2) $\dim(\mathscr{N}_\infty(t - \lambda_i)) = p_i$.

PROOF. Because $\dim(V)$ is the degree $p_1 + \cdots + p_\ell$ of the characteristic polynomial, to establish statement (1) we need only show that statement (2) holds and that $\mathscr{N}_\infty(t - \lambda_i) \cap \mathscr{N}_\infty(t - \lambda_j)$ is trivial whenever $i \neq j$.

For the latter, by Lemma 2.7, both $\mathscr{N}_\infty(t - \lambda_i)$ and $\mathscr{N}_\infty(t - \lambda_j)$ are t invariant. Notice that an intersection of t invariant subspaces is t invariant and so the restriction of t to $\mathscr{N}_\infty(t - \lambda_i) \cap \mathscr{N}_\infty(t - \lambda_j)$ is a linear transformation. But both $t - \lambda_i$ and $t - \lambda_j$ are nilpotent on this subspace and so if t has any eigenvalues on the intersection then its "only" eigenvalue is both λ_i and λ_j. That cannot be, so this restriction has no eigenvalues: $\mathscr{N}_\infty(t - \lambda_i) \cap \mathscr{N}_\infty(t - \lambda_j)$ is trivial (Lemma 3.10 shows that the only transformation without any eigenvalues is on the trivial space).

To prove statement (2), fix the index i. Decompose V as $\mathcal{N}_\infty(t - \lambda_i) \oplus \mathcal{R}_\infty(t - \lambda_i)$ and apply Lemma 2.8.

$$T = \left(\begin{array}{c|c} T_1 & Z_2 \\ \hline Z_1 & T_2 \end{array}\right) \quad \begin{array}{l} \}\dim(\mathcal{N}_\infty(t - \lambda_i)\text{)-many rows} \\ \}\dim(\mathcal{R}_\infty(t - \lambda_i)\text{)-many rows} \end{array}$$

By Lemma 2.9, $|T - xI| = |T_1 - xI| \cdot |T_2 - xI|$. By the uniqueness clause of the Fundamental Theorem of Arithmetic, the determinants of the blocks have the same factors as the characteristic polynomial $|T_1 - xI| = (x - \lambda_1)^{q_1} \ldots (x - \lambda_\ell)^{q_\ell}$ and $|T_2 - xI| = (x - \lambda_1)^{r_1} \ldots (x - \lambda_\ell)^{r_\ell}$, and the sum of the powers of these factors is the power of the factor in the characteristic polynomial: $q_1 + r_1 = p_1$, $\ldots$, $q_\ell + r_\ell = p_\ell$. Statement (2) will be proved if we will show that $q_i = p_i$ and that $q_j = 0$ for all $j \neq i$, because then the degree of the polynomial $|T_1 - xI|$ — which equals the dimension of the generalized null space — is as required.

For that, first, as the restriction of $t - \lambda_i$ to $\mathcal{N}_\infty(t - \lambda_i)$ is nilpotent on that space, the only eigenvalue of t on it is λ_i. Thus the characteristic equation of t on $\mathcal{N}_\infty(t - \lambda_i)$ is $|T_1 - xI| = (x - \lambda_i)^{q_i}$. And thus $q_j = 0$ for all $j \neq i$.

Now consider the restriction of t to $\mathcal{R}_\infty(t - \lambda_i)$. By Note II.2.2, the map $t - \lambda_i$ is nonsingular on $\mathcal{R}_\infty(t - \lambda_i)$ and so λ_i is not an eigenvalue of t on that subspace. Therefore, $x - \lambda_i$ is not a factor of $|T_2 - xI|$, and so $q_i = p_i$.   QED

Our major result just translates those steps into matrix terms.

2.12 Theorem Any square matrix is similar to one in *Jordan form*

$$\begin{pmatrix} J_{\lambda_1} & & & -zeroes- & & \\ & J_{\lambda_2} & & & & \\ & & \ddots & & & \\ & & & & J_{\lambda_{\ell-1}} & \\ & -zeroes- & & & & J_{\lambda_\ell} \end{pmatrix}$$

where each J_λ is the Jordan block associated with the eigenvalue λ of the original matrix (that is, is all zeroes except for λ's down the diagonal and some subdiagonal ones).

PROOF. Given an $n \times n$ matrix T, consider the linear map $t \colon \mathbb{C}^n \to \mathbb{C}^n$ that it represents with respect to the standard bases. Use the prior lemma to write $\mathbb{C}^n = \mathcal{N}_\infty(t - \lambda_1) \oplus \cdots \oplus \mathcal{N}_\infty(t - \lambda_\ell)$ where $\lambda_1, \ldots, \lambda_\ell$ are the eigenvalues of t. Because each $\mathcal{N}_\infty(t - \lambda_i)$ is t invariant, Lemma 2.8 and the prior lemma show that t is represented by a matrix that is all zeroes except for square blocks along the diagonal. To make those blocks into Jordan blocks, pick each B_{λ_i} to be a string basis for the action of $t - \lambda_i$ on $\mathcal{N}_\infty(t - \lambda_i)$.   QED

Jordan form is a canonical form for similarity classes of square matrices, provided that we make it unique by arranging the Jordan blocks from least eigenvalue to greatest and then arranging the subdiagonal 1 blocks inside each Jordan block from longest to shortest.

2.13 Example This matrix has the characteristic polynomial $(x-2)^2(x-6)$.

$$T = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 6 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

We will handle the eigenvalues 2 and 6 separately.

Computation of the powers, and the null spaces and nullities, of $T - 2I$ is routine. (Recall from Example 2.3 the convention of taking T to represent a transformation, here $t\colon \mathbb{C}^3 \to \mathbb{C}^3$, with respect to the standard basis.)

power p	$(T-2I)^p$	$\mathscr{N}((t-2)^p)$	nullity
1	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{pmatrix}$	$\{\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \mid x \in \mathbb{C}\}$	1
2	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{pmatrix}$	$\{\begin{pmatrix} x \\ -z/2 \\ z \end{pmatrix} \mid x, z \in \mathbb{C}\}$	2
3	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 64 & 32 \\ 0 & 0 & 0 \end{pmatrix}$	*–same–*	—

So the generalized null space $\mathscr{N}_\infty(t-2)$ has dimension two. We've noted that the restriction of $t - 2$ is nilpotent on this subspace. From the way that the nullities grow we know that the action of $t - 2$ on a string basis $\vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{0}$. Thus the restriction can be represented in the canonical form

$$N_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \text{Rep}_{B,B}(t-2) \qquad B_2 = \langle \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \rangle$$

where many choices of basis are possible. Consequently, the action of the restriction of t to $\mathscr{N}_\infty(t-2)$ is represented by this matrix.

$$J_2 = N_2 + 2I = \text{Rep}_{B_2,B_2}(t) = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$$

The second eigenvalue's computations are easier. Because the power of $x-6$ in the characteristic polynomial is one, the restriction of $t-6$ to $\mathscr{N}_\infty(t-6)$ must be nilpotent of index one. Its action on a string basis must be $\vec{\beta}_3 \mapsto \vec{0}$ and since it is the zero map, its canonical form N_6 is the 1×1 zero matrix. Consequently, the canonical form J_6 for the action of t on $\mathscr{N}_\infty(t-6)$ is the 1×1 matrix with the single entry 6. For the basis we can use any nonzero vector from the generalized null space.

$$B_6 = \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle$$

Taken together, these two give that the Jordan form of T is

$$\text{Rep}_{B,B}(t) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

where B is the concatenation of B_2 and B_6.

2.14 Example Contrast the prior example with

$$T = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

which has the same characteristic polynomial $(x-2)^2(x-6)$.

While the characteristic polynomial is the same,

power p	$(T-2I)^p$	$\mathscr{N}((t-2)^p)$	nullity
1	$\begin{pmatrix} 0 & 2 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{pmatrix}$	$\{ \begin{pmatrix} x \\ -z/2 \\ z \end{pmatrix} \mid x, z \in \mathbb{C}\}$	2
2	$\begin{pmatrix} 0 & 8 & 4 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{pmatrix}$	$-same-$	—

here the action of $t-2$ is stable after only one application — the restriction of of $t-2$ to $\mathscr{N}_\infty(t-2)$ is nilpotent of index only one. (So the contrast with the prior example is that while the characteristic polynomial tells us to look at the action of the $t-2$ on its generalized null space, the characteristic polynomial does not describe completely its action and we must do some computations to find, in this example, that the minimal polynomial is $(x-2)(x-6)$.) The restriction of $t-2$ to the generalized null space acts on a string basis as $\vec{\beta}_1 \mapsto \vec{0}$ and $\vec{\beta}_2 \mapsto \vec{0}$, and we get this Jordan block associated with the eigenvalue 2.

$$J_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

For the other eigenvalue, the arguments for the second eigenvalue of the prior example apply again. The restriction of $t-6$ to $\mathscr{N}_\infty(t-6)$ is nilpotent of index one (it can't be of index less than one, and since $x-6$ is a factor of the characteristic polynomial to the power one it can't be of index more than one either). Thus $t-6$'s canonical form N_6 is the 1×1 zero matrix, and the associated Jordan block J_6 is the 1×1 matrix with entry 6.

Therefore, T is diagonalizable.

$$\text{Rep}_{B,B}(t) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix} \qquad B = B_2 {}^\frown B_6 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \rangle$$

(Checking that the third vector in B is in the nullspace of $t-6$ is routine.)

2.15 Example A bit of computing with

$$T = \begin{pmatrix} -1 & 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & -4 & -1 & 0 & 0 \\ 3 & -9 & -4 & 2 & -1 \\ 1 & 5 & 4 & 1 & 4 \end{pmatrix}$$

shows that its characteristic polynomial is $(x-3)^3(x+1)^2$. This table

power p	$(T-3I)^p$	$\mathcal{N}((t-3)^p)$	nullity
1	$\begin{pmatrix} -4 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -4 & -4 & 0 & 0 \\ 3 & -9 & -4 & -1 & -1 \\ 1 & 5 & 4 & 1 & 1 \end{pmatrix}$	$\left\{ \begin{pmatrix} -(u+v)/2 \\ -(u+v)/2 \\ (u+v)/2 \\ u \\ v \end{pmatrix} \,\middle|\, u,v \in \mathbb{C} \right\}$	2
2	$\begin{pmatrix} 16 & -16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 16 & 16 & 0 & 0 \\ -16 & 32 & 16 & 0 & 0 \\ 0 & -16 & -16 & 0 & 0 \end{pmatrix}$	$\left\{ \begin{pmatrix} -z \\ -z \\ z \\ u \\ v \end{pmatrix} \,\middle|\, z,u,v \in \mathbb{C} \right\}$	3
3	$\begin{pmatrix} -64 & 64 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -64 & -64 & 0 & 0 \\ 64 & -128 & -64 & 0 & 0 \\ 0 & 64 & 64 & 0 & 0 \end{pmatrix}$	*–same–*	—

shows that the restriction of $t-3$ to $\mathcal{N}_\infty(t-3)$ acts on a string basis via the two strings $\vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{0}$ and $\vec{\beta}_3 \mapsto \vec{0}$.

A similar calculation for the other eigenvalue

power p	$(T+1I)^p$	$\mathcal{N}((t+1)^p)$	nullity
1	$\begin{pmatrix} 0 & 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & 0 \\ 3 & -9 & -4 & 3 & -1 \\ 1 & 5 & 4 & 1 & 5 \end{pmatrix}$	$\left\{ \begin{pmatrix} -(u+v) \\ 0 \\ -v \\ u \\ v \end{pmatrix} \,\middle|\, u,v \in \mathbb{C} \right\}$	2
2	$\begin{pmatrix} 0 & 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 0 & -16 & 0 & 0 & 0 \\ 8 & -40 & -16 & 8 & -8 \\ 8 & 24 & 16 & 8 & 24 \end{pmatrix}$	*–same–*	—

shows that the restriction of $t + 1$ to its generalized null space acts on a string basis via the two separate strings $\vec{\beta}_4 \mapsto \vec{0}$ and $\vec{\beta}_5 \mapsto \vec{0}$.

Therefore T is similar to this Jordan form matrix.

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

We close with the statement that the subjects considered earlier in this Chpater are indeed, in this sense, exhaustive.

2.16 Corollary Every square matrix is similar to the sum of a diagonal matrix and a nilpotent matrix.

Exercises

2.17 Do the check for Example 2.3.

2.18 Each matrix is in Jordan form. State its characteristic polynomial and its minimal polynomial.

(a) $\begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$

(e) $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$ (f) $\begin{pmatrix} 4 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 1 & -4 \end{pmatrix}$ (g) $\begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

(h) $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ (i) $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$

✓ **2.19** Find the Jordan form from the given data.

(a) The matrix T is 5×5 with the single eigenvalue 3. The nullities of the powers are: $T - 3I$ has nullity two, $(T - 3I)^2$ has nullity three, $(T - 3I)^3$ has nullity four, and $(T - 3I)^4$ has nullity five.

(b) The matrix S is 5×5 with two eigenvalues. For the eigenvalue 2 the nullities are: $S - 2I$ has nullity two, and $(S - 2I)^2$ has nullity four. For the eigenvalue -1 the nullities are: $S + 1I$ has nullity one.

2.20 Find the change of basis matrices for each example.

(a) Example 2.13   (b) Example 2.14   (c) Example 2.15

✓ **2.21** Find the Jordan form and a Jordan basis for each matrix.

(a) $\begin{pmatrix} -10 & 4 \\ -25 & 10 \end{pmatrix}$

(b) $\begin{pmatrix} 5 & -4 \\ 9 & -7 \end{pmatrix}$

(c) $\begin{pmatrix} 4 & 0 & 0 \\ 2 & 1 & 3 \\ 5 & 0 & 4 \end{pmatrix}$