JORDAN CANONICAL FORMS

OCT. 2019

1. Characteristic polynomials, eigenvalues and eigenvectors

Suppose that $n \geq 1$ and that $T : \mathbb{C}^n \to \mathbb{C}^n$ is a linear transformation. Let

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} \right\}$$

be the standard basis for $\mathbb{C}^n$. Then T has a matrix

$$[T]_S^S = M = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{pmatrix}.$$

relative to the basis S.

Definition 1.1. The characteristic polynomial of T (and of M) is
(1.1)

$$p_T(z) = p_M(z) = \det(z \cdot I_n - M) = \det \begin{pmatrix} (z - a_{1,1}) & -a_{1,2} & -a_{1,3} & \cdots & -a_{1,n} \\ -a_{2,1} & (z - a_{2,2}) & -a_{2,3} & \cdots & -a_{2,n} \\ -a_{3,1} & -a_{3,2} & (z - a_{3,3}) & \cdots & -a_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_{n,1} & -a_{n,2} & -a_{n,3} & \cdots & (z - a_{n,n}) \end{pmatrix}$$

when I_n is the $n \times n$ identity matrix.

Keep in mind that this is $(-1)^n$ times the definition in the course text. The above definition is the usual one. It has the advantage of making the highest degree term of $p_T(z)$ equal to z^n rather than $(-1)^n z^n$.

Because $p_T(z)$ has coefficients in $\mathbb{C}$, we can factor it as

$$p_T(z) = \prod_{i=1}^{d} (z - \lambda_i)^{n(i)} \tag{1.2}$$

for some distinct roots $\lambda_1, \ldots, \lambda_d$ in $\mathbb{C}$ and some integers $n(i) \geq 1$ such that

$$\sum_{i=1}^{d} n(i) = n = \text{degree}(p_T(z)) \tag{1.3}$$

For each $i = 1, \ldots, d$, let

$$M_i = M - \lambda_i \cdot I_n = -(\lambda_i \cdot I_n - M)$$

The significance of the characteristic polynomial $p_T(z)$ in 1.1 is that substituting λ_i for z leads to

$$p_T(\lambda_i) = \det(\lambda_i \cdot I_n - M) = (-1)^n \det(M - \lambda_i \cdot I_n) = (-1)^n \det(M_i) = 0 \quad (1.4)$$

because the factor $(z - \lambda_i)^{n(i)}$ on the right in (1.2) is 0 when $z = \lambda_i$. The fact that $\det(M_i) = 0$ means that M_i has rank less than n, so the null space $\mathrm{Null}(M_i)$ is not 0. Here the following conditions are equivalent:

i. $\underline{0} \neq b(i) \in \mathrm{Null}(M_i)$ where $\underline{0}$ is the zero vector.
ii. $M_i b(i) = Mb(i) - \lambda_i b(i) = \underline{0} \neq b(i)$.
iii. $Mb(i) = \lambda_i b(i)$ and $b(i) \neq \underline{0}$.

The last condition is the basis for the following definition:

Definition 1.2. An eigenvector with eigenvalue $\lambda_i \in \mathbb{C}$ is a non-zero vector $b(i)$ such that $T(b(i)) = \lambda_i b(i) = Mb(i)$.

Theorem 1.3. *A complex number λ is an eigenvalue of some eigenvector for for M if and only if λ is one of the roots λ_i of $p_M(t)$. Suppose $\{b(i)\}_{i=1}^d$ is any set of (non-zero) eigenvectors with distinct eigenvalues $\{\lambda_i\}_{i=1}^d$. Then $\{b(i)\}_{i=1}^d$ is a set of independent vectors.*

Proof. The first statement follows from the fact that λ is an eigenvalue of M if and only if $\lambda \cdot I_n - M$ has a non-zero null space, and this is true if and only if $p_M(\lambda) = \det(\lambda \cdot I_n - M)$ equals 0. Now let the $b(i)$ be as in the Theorem. Suppose to the contrary that $\{b(i)\}_{i=1}^d$ is dependent. Then there is some non-trivial linear dependency relation among them. There will be such a dependency relation that involves a minimal number of elements of $\{b(i)\}_{i=1}^d$. After reordering the $b(i)$, we can assume that this dependency relation has the form

$$\sum_{i=1}^{\ell} c_i b(i) = \underline{0}$$

for some $\ell \leq d$ with all $c_1, \ldots, c_\ell$ not 0. We have to have $\ell > 1$ since $b(1) \neq \underline{0}$ and $c_1 \neq 0$. Now

$$M(\sum_{i=1}^{\ell} c_i b(i)) = \sum_{i=1}^{\ell} M(c_i b(i)) = \sum_{i=1}^{\ell} c_i \lambda_i b(i).$$

Therefore

$$\begin{aligned} \underline{0} &= \lambda_\ell (\sum_{i=1}^{\ell} c_i b(i)) - M(\sum_{i=1}^{\ell} c_i b(i)) \\ &= (\sum_{i=1}^{\ell} c_i \lambda_\ell b(i)) - \sum_{i=1}^{\ell} c_i \lambda_i b(i) \\ &= \sum_{i=1}^{\ell-1} c_i (\lambda_\ell - \lambda_i) b(i) \quad (1.5) \end{aligned}$$

since the last terms in the sums in the second row cancel one another. But $\lambda_\ell - \lambda_i \neq 0$ for $i < \ell$ since $\lambda_1, \ldots, \lambda_\ell$ are distinct. This shows (1.5) is a shorter non-trivial dependency relation, which is a contradiction. This means the original hypothesis that $\{b(i)\}_{i=1}^d$ is not a set of independent vectors must have been false. □

2. The "non-defective" case

The text defines M to be non-defective if $\mathbb{C}^n$ has a basis of eigenvectors for M. We can say exactly when this occurs:

Theorem 2.1. *Let M be an $n \times n$ matrix with complex entries.*

i. *The matrix M is non-defective if and only if for each eigenvalue λ_i, the dimension of the null space* $\mathrm{Null}(M_i)$ *of $M_i = M - \lambda_i \cdot I_n$ equals the multiplicity $n(i)$ of λ_i as a root of $p_M(t) = \prod_{i=1}^{d}(t-\lambda_i)^{n(i)}$.*
ii. *If M is not defective, a basis B of eigenvectors for M is given by $\cup_{i=1}^{d} B_i$ when B_i is a basis for* $\mathrm{Null}(M_i)$*. The transition matrix $P = P_{S \leftarrow B}$ has columns the vectors of B in the standard basis. The matrix $P^{-1}MP$ is diagonal, with diagonal entries the eigenvalues of B.*
iii. *If $d = n$, so that M has n distinct eigenvalues, then M is not defective, and* $\mathrm{Null}(M_i)$ *has dimension* 1 *for all i.*

Proof. The elements of $\mathrm{Null}(M_i)$ are just the eigenvectors of M with eigenvalue λ_i. Suppose first that $\mathrm{Null}(M_i)$ has dimension $n(i)$ for all i. Let B_i be a basis for $\mathrm{Null}(M_i)$. Let's check that $B = \cup_{i=1}^{d} B_i$ is independent. Recall that any linear combination of elements of B_i is an eigenvector for M with eigenvalue λ_i, and a set of non-zero eigenvectors with different eigenvalues is independent. Suppose some linear combination with non-zero coefficients of the elements of B is 0. We can write this combination in the form

$$\sum_{i=1}^{d}(\sum_{j=1}^{m_i} r_{i,j} b_{i,j}) = \sum_{i=1}^{d} b_i = 0$$

where $b_{i,j} \in B_i$ for all i, some of the $r_{i,j} \in \mathbb{C}$ are not zero, and $b_i = \sum_{j=1}^{m_i} r_{i,j} b_{i,j}$ is either 0 or an eigenvector with eigenvalue λ_i. Since eigenvectors associated to different eigenvalues are linearly independent, we conclude that all of the b_i must be 0. But then since some of the $r_{i,j}$ are not 0, this contradicts each B_i being independent sets of vectors. If $\mathrm{Null}(M_i)$ has dimension $n(i)$, then B_i has $n(i)$ elements and B has $\sum_{i=1}^{d} n(i) = n$ elements. This forces B to be a basis of eigenvectors for $\mathbb{C}^n$.

Conversely, suppose there is a basis B of $\mathbb{C}^n$ of eigenvectors for M. We have shown that the eigenvalue of any eigenvector must be one of $\lambda_1, \ldots, \lambda_d$. Let $m(i)$ be the number of elements of B which have eigenvalue λ_i. The matrix of the linear tranformation T defined by M in the standard basis has the property that $M' = [T]_B^B$ is a diagonal matrix, with diagonal entries given by the eigenvalues of B. Therefore

$$\prod_{i=1}^{d}(t-\lambda_i)^{n(i)} = p_M(t) = p_{M'}(t) = \prod_{i=1}^{d}(t-\lambda_i)^{m(i)}.$$

This forces $m(i) = n(i)$, so $B \cap \mathrm{Null}(M_i) = B_i$ has $m(i) = n(i)$ elements for all i. Here

$$n(i) = \#B_i \le \dim(\mathrm{Null}(M_i))$$

since the B_i are independent. Because eigenvectors associated to different eigenvalues are independent, we get

$$\sum_{i=1}^{d} \dim(\mathrm{Null}(M_i)) \le \dim(\mathbb{C}^n) = n.$$

Putting all of this together shows

$$n = \sum_{i=1}^{d} n(i) \leq \sum_{i=1}^{d} \dim(\mathrm{Null}(M_i)) \leq n.$$

The only say this is possible is for $m(i) = n(i) = \dim(\mathrm{Null}(M_i))$ for all i

Finally, suppose $n = d$, so that M has distinct eigenvalues. Since M_i has determinant 0, $\dim(\mathrm{Null}(M_i)) \geq 1$. Let B be a set consisting of one non-zero element of $\mathrm{Null}(M_i)$ for $i = 1, \ldots, d$. Since $d = n$, B is an independent set of size n, so B is a basis of $\mathbb{C}^n$ and M is not defective. □

3. A consequence to computing powers of M

Suppose that as in Theorem 2.1, M is not defective. Let $B = \{b(1), \ldots, b(n)\}$ be a basis of $\mathbb{C}^n$ consisting of eigenvectors for M, and let $\lambda(i)$ be the eigenvalue of $b(i)$. Note that we could have $\lambda(i) = \lambda(j)$ for some $i \neq j$ since some eigenvalues may be repeated roots of $p_M(t)$. We can easily compute the powers M^ℓ of M for integers $\ell \geq 1$. One has

$$[T^\ell]_B^B = ([T]_B^B)^\ell = \begin{pmatrix} \lambda(1)^\ell & 0 & 0 & \cdots & 0 \\ 0 & \lambda(2)^\ell & 0 & \cdots & 0 \\ 0 & 0 & \lambda(3)^\ell & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda(n)^\ell \end{pmatrix}.$$

so

$$M^\ell = ([T]_S^S)^\ell = [T^\ell]_S^S = P \cdot \begin{pmatrix} \lambda(1)^\ell & 0 & 0 & \cdots & 0 \\ 0 & \lambda(2)^\ell & 0 & \cdots & 0 \\ 0 & 0 & \lambda(3)^\ell & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda(n)^\ell \end{pmatrix} \cdot P^{-1} \quad \text{when} \quad P = P_{S \leftarrow B}.$$

One can interpret this result in the following way. Each $v \in \mathbb{C}^n$ is a unique linear combination

$$v = c_1 b(1) + c_2 b(2) + \cdots + c_n b(n)$$

of the elements of B. One has

$$T^\ell(v) = c_1 \lambda(1)^\ell b(1) + c_2 \lambda(2)^\ell b(2) + \cdots + c_n \lambda(n)^\ell b(n).$$

4. The general case

Suppose now that T has characteristic polynomial

$$p_T(z) = \prod_{i=1}^{d} (z - \lambda_i)^{n(i)}$$

with $d \leq n$ and $n(i) \geq 1$. We first need to define generalized eigenvectors.

Definition 4.1. A vector $b(i) \in \mathbb{C}^n$ is a generalized eigenvector with eigenvalue λ_i and multiplicity $m \geq 1$ if $(M - \lambda_i I_n)^m b(i) = \underline{0}$ but $(M - \lambda_i I_n)^{m-1} b(i) \neq \underline{0}$.

Lemma 4.2. *Suppose $b(i)$ is a generalized eigenvector with eigenvalue λ_i and multiplicity $m \geq 1$. Write $M_i = M - \lambda_i I_n$ as before. Then the $\mathbb{C}$ vector space W spanned by*

$$\{M_i^{m-1} b(i), M_i^{m-2} b(i), \cdots, b(i)\} = \{d_1, d_2, \ldots, d_m\} = D$$

has dimension m. The action of T takes W back itself. The matrix of T acting on W relative to the basis D is the $m \times m$ *Jordan block matrix*

$$[T_{|W}]_D^D = J(m,\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \tag{4.6}$$

which has λ *down the diagonal and* 1 *just above the diagonal.*

Proof. Suppose D is not independent. Then there is some linear dependency relation

$$c_0 b(i) + c_1 M_i b(i) + \cdots c_{m-1} M_i^{m-1} b(i) = \underline{0} \tag{4.7}$$

in which not all of the c_j are 0. Suppose that J is the smallest integer such that $c_J \neq 0$. Then $M_i^{m-J-1} M_i^J b(i) = M_i^{m-1} b(i) \neq \underline{0}$. However, if $j > J$, then $m - J - 1 + j \geq m$, so $M_i^{m-J-1} M_i^j b(i) = \underline{0}$ because $M_i^m b(i) = \underline{0}$. If we apply M_i^{m-J-1} to both sides of (4.7) we therefore get

$$M_i^{m-J-1}(c_0 b(i) + c_1 M_i b(i) + \cdots c_{m-1} M_i^{m-1} b(i)) = c_J M_i^{m-1} b(i) = \underline{0}.$$

But this contradicts $c_J \neq 0$ and $M_i^{m-1} b(i) \neq \underline{0}$. So D must have been an independent set of vectors.

To compute the action of M on W, note that

$$M = M_i + \lambda_i I_n.$$

and as above

$$M_i^m b(i) = \underline{0}.$$

Hence for $d_1 = M_i^{m-1} b(i)$ we get

$$M \cdot d_1 = (M_i + \lambda_i I_n) \cdot M_i^{m-1} b(i) = M_i^m b(i) + \lambda_i M_i^{m-1} b(i) = \lambda_i d_1$$

since $M_i^m b(i) = \underline{0}$ and $M_i^{m-1} b(i) = d_1$. This is consistent with the matrix (4.6) since d_1 corresponds to the column vector

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

when we use the coordinates associated to the basis $D = \{d_1, d_2, \ldots, d_m\}$ of W. We find in a similar way that if $1 < j \leq m$ then

$$M \cdot d_j = (M_i + \lambda_i I_n) \cdot M_i^{m-j} b(i) = M_i^{m-(j-1)} b(i) + \lambda_i M_i^{m-j} b(i) = d_{j-1} + \lambda_i d_j$$

and this is consistent with the matrix (4.6). □

The main Theorem about Jordan canonical forms is:

Theorem 4.3. *Suppose the characteristic polynomial of T is*

$$p_T(z) = \prod_{i=1}^{d} (z - \lambda_i)^{n(i)}$$

in which $\lambda_1, \cdots, \lambda_d$ *are distinct complex numbers* $n(i) \geq 1$ *for all i. There is a basis* $B = \cup_{i=1}^d B_i$ *for* $\mathbb{C}^n$ *with the following properties.*

i. *For each i there is an integer $j(i) \geq 1$ so that*
$$B_i = \cup_{j=1}^{j(i)} B_{i,j}$$
with
$$B_{i,j} = \{M_i^{n(i,j)-1} b(i,j), M^{n(i,j)-2} b(i,j), \ldots, b(i,j)\}$$
for some generalized eigenvector $b(i,j)$ with eigenvalue λ_i and multiplicity $n(i,j) \geq 1$.

ii. *One has*
$$\sum_{j=1}^{j(i)} \# B_{i,j} = \sum_{j=1}^{j(i)} n(i,j) = n(i)$$

iii. *The matrix $[T]_B^B$ of T relative to the basis in part (i) is a block diagonal matrix with the Jordan blocks $J(n(i,j), \lambda_i)$ for $1 \leq i \leq d$ and $1 \leq j \leq j(i)$ running down the diagonal and all entries not in one of these blocks equal to 0.*

iv. *Each Jordan block $J(n(i,j), \lambda_i)$ produces a one dimensional space of eigenvectors with eigenvalue λ_i. The null space $Null(M_i)$ is the space spanned by these subspaces and has dimension equal to the number $j(i)$ of Jordan blocks associated to the eigenvalue λ_i.*

v. *From the dimensions of the Null spaces $\mathrm{Null}(M_i^\ell)$ for $1 \leq i \leq d$ and $1 \leq \ell \leq n(i)$ one can determine the sizes of all the Jordan blocks of $[T]_B^B = P^{-1}MP$ when $P = P_{S \leftarrow B}$.*

I will not prove this, or the fact that the Jordan blocks appearing in part (ii) are unique up to permutation. Instead I'll first discuss how to find all possible Jordan forms associated to a given characteristic polynomial. I'll then discuss how to find the Jordan canonical form of two-by-two and three-by-three matrices. In a final (optional!) section I'll give a general algorithm for finding a basis B of the kind in Theorem 4.3.

5. Finding all the Jordan forms which can arise from a given characteristic polynomial.

Suppose we want to find all Jordan forms of $n \times n$ matrices which have a given characteristic polynomial $p_M(z) = \prod_{i=1}^{d} (z - \lambda_i)^{n(i)}$ with $\lambda_1, \ldots, \lambda_d$ a given set of distinct complex numbers and $n(1), \ldots, n(d)$ a given set of positive integers whose sum is n.

If all the $n(i) = 1$, then the only Jordan form (up to permuting the Jordan blocks) is the diagonal matrix with $\lambda_1, \ldots, \lambda_n$ down the diagonal, where $n = d$.

For those $n(i)$ larger than 1, we need to consider all ways of writing
$$n(i) = \sum_{j=1}^{j(i)} n(i,j)$$
with $j(i) \geq 1$ and all $n(i,j) \geq 1$. Suppose we make such a choice of the $j(i)$ and the $n(i,j)$. Then there will be a matrix with the right characteristic polynomial which has Jordan blocks $J(n(i,j), \lambda_i)$ for all i and j. Reordering the blocks corresponds to permuting the basis.

As an example, suppose $n = 3$ and that $p_T(z) = (z-1)^2(z-2)$. Then $d = 2$, $\lambda_1 = 1$ and $\lambda_2 = 2$. There is a unique 1×1 Jordan block $J(1, \lambda_2) = J(1, 2)$ associated to $\lambda_2 = 2$, since $(z - \lambda_2) = (z - 2)$ occurs to the first power in $p_T(z)$, meaning $n(2) = 1$. There are two possibilities for the Jordan blocks associated to the eigenvalue $\lambda_1 = 1$.

1. We write $n(1) = 2 = 1 + 1$ so $j(1) = 2$ and $n(1,1) = n(1,2) = 1$. We get two 1×1 Jordan blocks $J(1,1)$ associated to $\lambda_1 = 1$.
2. We write $n(1) = 2$, so $j(1) = 1$ and $n(1,1) = 2$. Then there is one 2×2 Jordan block $J(2,1)$ associated to $\lambda_1 = 1$.

The final conclusion is that up to permuting the blocks, there are only two possible Jordan canonical forms:

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

and

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

6. Finding the Jordan canonical form of a two-by-two matrix.

In this section we fix $n = 2$ and describe how to find the Jordan canonical form of a two by two matrix

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

The characteristic polynomial is

$$p_M(t) = \det(\begin{pmatrix} t - a_{1,1} & -a_{1,2} \\ -a_{2,1} & t - a_{2,2} \end{pmatrix}) = t^2 - \text{Trace}(M)\ t + \det(M)$$

where $\text{Trace}(M) = a_{1,1} + a_{2,2}$. There are two ways a quadratic polynomial can factor

A. Suppose $p_M(t) = (t - \lambda_1) \cdot (t - \lambda_2)$ for distinct roots $\lambda_1, \lambda_2 \in \mathbb{C}$. In this case, M is not defective. The space $\text{Null}(M_i) = \text{Null}(M - \lambda_i I_2)$ is one dimensional for $i = 1, 2$. If we let $b(i)$ be a basis for $\text{Null}(M_i)$ then $B = \{b(1), b(2)\}$ is a basis of eigenvectors. If $P = P_{S \leftarrow B}$ is the two by two matrix with columns given by the vectors $b(1)$ and $b(2)$ in the standard basis, then

$$P^{-1}MP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{and} \quad M = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.$$

B. Suppose $p_M(t) = (t - \lambda_1)^2$ for some λ_1. Then the multiplicity of λ_1 as a root is $n(1) = 2$ and $d = 1$. There are two subcases:

B1. The matrix M is non-defective if and only if $\dim(\text{Null}(M_1)) = n(1) = 2$ when $M_1 = M - \lambda_1 I_2$. This occurs if and only if M_1 is the zero matrix, in which case

$$M = \lambda_1 I_2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$$

B2. Suppose M is defective. Then it has to be the case that $1 \leq \dim(\text{Null}(M_1)) < n(1) = 2$, so $\text{Null}(M_1)$ is one-dimensional. Recall that each Jordan block in the Jordan normal form of M produces exactly one new independent eigenvector. Since we have just one eigenvalue, and there is only a one dimensional space of eigenvectors for this eigenvalue, there can be just one Jordan block. The sum of the sizes of all the Jordan blocks must be $n = 2$, so we see that there is just one two-by-two Jordan block. The Jordan normal form of M is thus

$$M' = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$$

Notice that

$$M' - \lambda_1 I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

has the property that $(M' - \lambda_1 I_2)^2$ is the zero matrix. Changing bases back from B to the standard basis S, we see that $(M - \lambda_1 I_2)^2$ is also the zero matrix. To find a B explicitly, let v_1 is any element of $\mathbb{C}^2$ which is not in $\text{Null}(M_1)$. Define $v_2 = M_1 v_1 = (M - \lambda_1 I_2)v_1$. Then v_2 is an eigenvector for M with eigenvalue λ_1, since

$$(M - \lambda_1 I_2)v_2 = (M - \lambda_1 I_2)^2 v_1 = 0.$$

Thus

$$Mv_2 = \lambda_1 v_2.$$

We have

$$Mv_1 = \lambda_1 v_1 + (M - \lambda_1 I_2)v_1 = \lambda_1 v_1 + v_2.$$

So if we change bases from the standard basis S to $B = \{v_2, v_1\}$, M becomes

$$M' = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$$

To check this recall the the vectors corresponding to v_2 and v_1 relative to B are

$$[v_2]_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad [v_1]_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Multiplying these vectors on the left by M' sends them to the images of v_2 and v_1, respectively, under the linear transformation T defined by M in the standard basis.

7. Finding the Jordan canonical form of a three-by-three matrix.

In this section we fix $n = 3$ and describe how to find the Jordan canonical form of a three by three matrix

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

The characteristic polynomial is

$$p_M(t) = \det(tI_3 - M)$$

This is a cubic in t. It can factor in three ways.

A. Suppose $p_M(t) = (t - \lambda_1) \cdot (t - \lambda_2) \cdot (t - \lambda_3)$ for distinct roots $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$. In this case, M is not defective. The space $\text{Null}(M_i) = \text{Null}(M - \lambda_i I_2)$ is one dimensional for $i = 1, 2, 3$. If we let $b(i)$ be a basis for $\text{Null}(M_i)$ then $B = \{b(1), b(2), b(3)\}$ is a basis of eigenvectors. If $P = P_{S\leftarrow B}$ is the three by three matrix with columns given by the vectors $b(1)$, $b(2)$, $b(3)$ in the standard basis, then

$$P^{-1}MP = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \text{and} \quad M = P\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}P^{-1}.$$

B. Suppose $p_M(t) = (t-\lambda_1)^2 \cdot (t-\lambda_2)$ for some $\lambda_1 \neq \lambda_2$. Then the multiplicity of λ_1 as a root is $n(1) = 2$, the multiplicity of λ_2 is $n(2) = 1$ and $d = 2$. There are two subcases:

B1. The matrix M is non-defective if and only if $\dim(\text{Null}(M_1)) = n(1) = 2$ when $M_1 = M - \lambda_1 I_3$. In this case, we can form a basis B of eigenvectors by taking the union of bases for $\text{Null}(M_1)$ and $\text{Null}(M_2)$ when $M_2 = M - \lambda_2 I_3$. Here $\text{Null}(M_2)$ is one dimensional. The resulting Jordan canonical form is

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

B2. Suppose M is defective. This occurs precisely when $1 = \dim(\text{Null}(M_1)) < n(1) = 2$, We have only a one-dimensional space of eigenvectors for each of the eigenvalues λ_1 and λ_2. So there are exactly two Jordan blocks. These must have sizes 2 and 1 since the sum of their sizes must be $n = 3$. The block of size 2 is associated to λ_1 since $p_M(t) = (z-\lambda_1)^2 \cdot (z-\lambda_2)$ and the diagonal entries down the Jordan normal form determine $p_M(t)$. So the Jordan normal form is

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

A basis B which leads to this Jordan normal form for M can be found using the algorithm in the optional section below. This specializes in this case to the following. Let v_1 be any element of the $\text{Null}(M_1^2)$ which is not in $\text{Null}(M_1)$. Let $v_2 = M_1 v_1$. We can take $B = \{v_2, v_1, v_3\}$ when v_3 is any eigenvector for λ_2.

C. Suppose $p_M(t) = (t-\lambda_1)^2$. Then $n(1) = 3$ and $d = 1$. There are these subcases:

C1. M is non-defective if and only if $\dim(\text{Null}(M_1)) = 3$ when $M_1 = M - \lambda_1 I_3$. This happens if and only if $M = \lambda_1 I_3$. The Jordan normal form is just

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

C2. Suppose M is defective, so that $1 = \dim(\text{Null}(M_1)) < 3 = n(1)$. The number of independent eigenvectors for M is just $\dim(\text{Null}(M_1))$, and this is the number of Jordan blocks. So if $\dim(\text{Null}(M_1)) = 1$, there is just one Jordan block, and the Jordan normal form is

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

If $\dim(\text{Null}(M_1)) = 1$, the algorithm in the next section shows that we can arrive at the Jordan normal form by picking a basis of the form $B = \{M_1^2 v_1, M_1 v_1, v_1\}$

when v_1 is any vector such that $M_1^2 v_1 \neq 0$. If dim(Null(M_1)) = 2 the Jordan normal form is

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

When dim(Null(M_1)) = 2 we can take $B = \{M_1 v_1, v_1, v_3\}$ when v_1 is any vector not in Null(M_1) and v_3 is chosen so that $\{M_1 v_1, v_3\}$ is a basis for the two-dimensional space Null(M_1).

8. **Optional Section**: A general algorithm for finding a basis B of the kind in Theorem 4.3.

1. For $1 \le i \le d$ recall $M_i = M - \lambda_i I_n$. Fix i. Choose successive finite subsets $D(i, n(i)), D(i, n(i)-1), \dots D(i,1)$ so that for all integers $N = n(i), n(i)-1, \dots, 1$ one has

$$\left(\cup_{m=N}^{n(i)} M_i^{m-N} D(i,m)\right) \cup \text{Basis}(Null(M_i^{N-1})) = \text{Basis}(Null(M_i^N)) \tag{8.8}$$

   Here $Null(M_i^t)$ is the null space of the power M_i^t of M_i. On the left side of (8.8), Basis($Null(M_i^{N-1})$) means any choice of basis of $Null(M_i^{N-1})$, and $M_i^{m-N} D(i,m)$ means the finite set of vectors which results from applying M_i^{m-N} to the elements of $D(i,m)$. The right side of (8.8) means that the left side should form a basis for $Null(M_i^N)$.
2. Each $b \in D(i,m)$ is now a generalized eigenvector with eigenvalue λ_i and multiplicity m. The transformation T sends $Span(\{M_i^{m-1} b, M_i^{m-2} b, \dots, b\})$ back to itself, and the matrix of the action of T on this span is the Jordan block matrix $J(m, \lambda_i)$.
3. The union $B_i = \cup_{m=1}^{n(i)} \cup_{b \in D(i,m)} \{M_i^{m-1} b, M_i^{m-2} b, \dots, b\}$ is a set of the kind in Theorem 4.3(i). This gives a basis $B = \cup_{i=1}^{d} B_i$ for which $[T]_B^B$ is a Jordan block matrix of the kind in Theorem 4.3.

Example 8.1. Suppose $n(i) = 1$ for all i, as in the previous section, so that $d = n$. Condition (1) of the algorithm only applies with $N = n(i) = 1$. It says that we should have

$$\left(\cup_{m=1}^{1} M_i^{m-1} D(i,m)\right) \cup \text{Basis}(Null(M_i^{1-1})) = \text{Basis}(Null(M_i^1))$$

The only integer m in the union on the left is $m = 1$. So this simplifies to

$$D(i,1) \cup \text{Basis}(Null(M_i^0)) = \text{Basis}(Null(M_i))$$

Now M_i^0 should be interpreted as the identity matrix, so Basis($Null(M_i^0)$) is the empty set, and we just want

$$D(i,1) = \text{Basis}(Null(M_i))$$

In other words, $D(i,1)$ should consist of a basis for the nullspace of M_i, and we know that when all $n(i) = 1$, this nullspace has dimension 1. Do $D(i,1)$ is a one element set. The set B_i in part (3) of the algorithm will just be $D(i,1)$, and now the basis $B = \cup_{i=1}^{d} B_i = \cup_{i=1}^{n} D(i,1)$ consists of one eigenvector for each λ_i. This agrees with the Theorem 2.1 of the previous section.

Example 8.2. Here is more interesting example. Suppose $n = 2$ and that $T : \mathbb{C}^2 \to \mathbb{C}^2$ has standard matrix

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

The characteristic polynomial is

$$p_M(z) = \det \begin{pmatrix} (z-1) & -1 \\ 1 & (z-3) \end{pmatrix} = (z-1)(z-3)+1 = z^2-4z+3+1 = z^2-4z+4 = (z-2)^2$$

So $d = 1$, $\lambda_1 = 2$ and $n(1) = 2$.

In step (1) of the algorithm, we let $i = 1$ and

$$M_1 = M - \lambda_1 I_2 = M - 2I_2 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$$

Step (1) says we first want to find a set $D(1, n(1)) = D(1, 2)$ with the following properties. Setting $N = n(i) = n(1) = 2$, we want

$$D(1,2) \cup \text{Basis}(Null(M_1^{2-1})) = \text{Basis}(Null(M_1^2)) \tag{8.9}$$

Here

$$Null(M_1^{2-1}) = Null(M_1) = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} = \mathbb{C} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

The matrix M_1^2 is the zero matrix, so $Null(M_1^2) = \mathbb{C}^2$. Thus (8.9) says that adding $D(1, 2)$ to the basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for $Null(M_1)$ should give a basis for $\mathbb{C}^2$. One choice is to let

$$D(1,2) = \{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$$

Now steps 2 and 3 of the algorithm says that the set

(8.10)

$$\cup_{b \in D(1,2)} \{M_1^{2-1}b, M_1^{2-2}b\} = \{\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\} = \{\begin{pmatrix} -2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$$

will a subset of the basis of $\mathbb{C}^2$ we want to construct, and this subset will produce $\#D(1, 2) = 1$ Jordan block of the form

$$J(2, \lambda_1) = J(2, 2) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

We know that any basis has to have two elements, so (8.10) is in fact all of the basis B we want. Relative to this B, $[T]_B^B$ is the Jordan block $J(2, \lambda_1) = J(2, 2)$. The transition matrix

$$P = P_{S \leftarrow B}$$

should have as its columns the vectors

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

because of (8.10). So

$$P = \begin{pmatrix} -2 & 1 \\ -2 & -1 \end{pmatrix} \quad \text{and} \quad P^{-1} = \frac{1}{4} \begin{pmatrix} -1 & -1 \\ 2 & -2 \end{pmatrix}$$

We should have

$$[T]_B^B = J(2,2) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = P^{-1}MP = P^{-1} \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} P$$

and some calculations show this is indeed the case.