18.700 JORDAN NORMAL FORM NOTES

These are some supplementary notes on how to find the Jordan normal form of a small matrix. First we recall some of the facts from lecture, next we give the general algorithm for finding the Jordan normal form of a linear operator, and then we will see how this works for small matrices.

1. Facts

Throughout we will work over the field \( \mathbb{C} \) of complex numbers, but if you like you may replace this with any other algebraically closed field. Suppose that \( V \) is a \( \mathbb{C} \)-vector space of dimension \( n \) and suppose that \( T : V \to V \) is a \( \mathbb{C} \)-linear operator. Then the characteristic polynomial of \( T \) factors into a product of linear terms, and the irreducible factorization has the form

\[
c_T(X) = (X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \cdots (X - \lambda_r)^{m_r},
\]

for some distinct numbers \( \lambda_1, \ldots, \lambda_r \in \mathbb{C} \) and with each \( m_i \) an integer \( m_i \geq 1 \) such that \( m_1 + \cdots + m_r = n \).

Recall that for each eigenvalue \( \lambda_i \), the eigenspace \( E_{\lambda_i} \) is the kernel of \( T - \lambda_i I_V \). We generalized this by defining for each integer \( k = 1, 2, \ldots \) the vector subspace

\[
E_{(X-\lambda_i)^k} = \ker(T - \lambda_i I_V)^k.
\]

It is clear that we have inclusions

\[
E_{\lambda_i} = E_{X-\lambda_i} \subset E_{(X-\lambda_i)^2} \subset \cdots \subset E_{(X-\lambda_i)^e} \subset \cdots .
\]

Since \( \dim(V) = n \), it cannot happen that each \( \dim(E_{(X-\lambda_i)^k}) < \dim(E_{(X-\lambda_i)^{k+1}}) \), for each \( k = 1, \ldots, n \). Therefore there is some least integer \( e_i \leq n \) such that \( E_{(X-\lambda_i)^{e_i}} = E_{(X-\lambda_i)^{e_i+1}} \).

As was proved in class, for each \( k \geq e_i \) we have \( E_{(X-\lambda_i)^k} = E_{(X-\lambda_i)^{e_i}} \), and we defined the \emph{generalized eigenspace} \( E_{\lambda_i}^{\text{gen}} \) to be \( E_{(X-\lambda_i)^{e_i}} \).

It was proved in lecture that the subspaces \( E_{\lambda_1}^{\text{gen}}, \ldots, E_{\lambda_r}^{\text{gen}} \) give a direct sum decomposition of \( V \). From this our criterion for diagonalizability of follows: \( T \) is diagonalizable iff for each \( i = 1, \ldots, r \), we have \( E_{\lambda_i}^{\text{gen}} = E_{\lambda_i} \). Notice that in this case \( T \) acts on each \( E_{\lambda_i}^{\text{gen}} \) as \( \lambda_i \) times the identity. This motivates the definition of the \emph{semisimple part} of \( T \) as the unique \( \mathbb{C} \)-linear operator \( S : V \to V \) such that for each \( i = 1, \ldots, r \) and for each \( v \in E_{\lambda_i}^{\text{gen}} \), we have \( S(v) = \lambda_i v \).

We defined \( N = T - S \) and observed that \( N \) preserves each \( E_{\lambda_i}^{\text{gen}} \) and is \emph{nilpotent}, i.e. there exists an integer \( e \geq 1 \) (really just the maximum of \( e_1, \ldots, e_r \)) such that \( N^e \) is the zero linear operator. To summarize:

(A) The \emph{generalized eigenspaces} \( E_{\lambda_1}^{\text{gen}}, \ldots, E_{\lambda_r}^{\text{gen}} \) defined by

\[
E_{\lambda_i}^{\text{gen}} = \{ v \in V | \exists e, (T - \lambda_i I_V)^e(v) = 0 \},
\]

Date: Fall 2001.
give a direct sum decomposition of $V$. Moreover, we have $\dim(E_{\lambda_i}^{\text{gen}})$ equals the algebraic multiplicity of $\lambda_i$, $m_i$.

**(B)** The semisimple part $S$ of $T$ and the nilpotent part $N$ of $T$ defined to be the unique $\mathbb{C}$-linear operators $V \to V$ such that for each $i = 1, \ldots, r$ and each $v \in E_{\lambda_i}^{\text{gen}}$ we have

$$S(v) = S^{(i)}(v) = \lambda_i v, \quad N(v) = N^{(i)}(v) = T(v) - \lambda_i v,$$

satisfy the properties:

1. $S$ is diagonalizable with $c_S(X) = c_T(X)$, and the $\lambda_i$-eigenspace of $S$ is $E_{\lambda_i}^{\text{gen}}$ (for $T$).
2. $N$ is nilpotent, $N$ preserves each $E_{\lambda_i}^{\text{gen}}$ and if $N^{(i)} : E_{\lambda_i}^{\text{gen}} \to E_{\lambda_i}^{\text{gen}}$ is the unique linear operator with $N^{(i)}(v) = N(v)$, then $[N^{(i)}]_{e_i}^{-1}$ is nonzero but $[N^{(i)}]_{e_i}^{-1} = 0$.
3. $T = S + N$.
4. $SN = NS$.
5. For any other $\mathbb{C}$-linear operator $T' : V \to V$, $T'$ commutes with $T$ ($T'T = TT'$) iff $T'$ commutes with both $S$ and $N$. Moreover $T'$ commutes with $S$ iff for each $i = 1, \ldots, r$, we have $(E_{\lambda_i}^{\text{gen}}) \subset E_{\lambda_i}^{\text{gen}}$.
6. If $(S', N')$ is any pair of a diagonalizable operator $S'$ and a nilpotent operator $N'$ such that $T = S' + N'$ and $S'N' = N'S'$, then $S' = S$ and $N' = N$. We call the unique pair $(S, N)$ the semisimple-nilpotent decomposition of $T$.

**(C)** For each $i = 1, \ldots, r$, choose an ordered basis $B^{(i)} = (v^{(i)}_1, \ldots, v^{(i)}_{m_i})$ of $E_{\lambda_i}^{\text{gen}}$ and let $B = (B^{(1)}, \ldots, B^{(r)})$ be the concatenation, i.e.

$$B = \left(v^{(1)}_1, \ldots, v^{(1)}_{m_1}, v^{(2)}_1, \ldots, v^{(2)}_{m_2}, \ldots, v^{(r)}_1, \ldots, v^{(r)}_{m_r}\right).$$

For each $i$ let $S^{(i)}$, $N^{(i)}$ be as above and define the $m_i \times m_i$ matrices

$$D^{(i)} = [S^{(i)}]_{B^{(i)}, B^{(i)}}, \quad C^{(i)} = [N^{(i)}]_{B^{(i)}, B^{(i)}}.$$  

Then we have $D^{(i)} = \lambda_i I_{m_i}$ and $C^{(i)}$ is a nilpotent matrix of exponent $e_i$. Moreover we have the block forms of $S$ and $N$:

$$[S]_{B,B} = \begin{pmatrix} \lambda_1 I_{m_1} & 0_{m_1 \times m_2} & \ldots & 0_{m_1 \times m_r} \\ 0_{m_2 \times m_1} & \lambda_2 I_{m_2} & \ldots & 0_{m_2 \times m_r} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{m_r \times m_1} & 0_{m_r \times m_2} & \ldots & \lambda_r I_{m_r} \end{pmatrix},$$

$$[N]_{B,B} = \begin{pmatrix} C^{(1)} & 0_{m_1 \times m_2} & \ldots & 0_{m_1 \times m_r} \\ 0_{m_2 \times m_1} & C^{(2)} & \ldots & 0_{m_2 \times m_r} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{m_r \times m_1} & 0_{m_r \times m_2} & \ldots & C^{(r)} \end{pmatrix}.$$  

Notice that $D^{(i)}$ has a nice form with respect to ANY basis $B^{(i)}$ for $E_{\lambda_i}^{\text{gen}}$. But we might hope to improve $C^{(i)}$ by choosing a better basis.
A very simple kind of nilpotent linear transformation is the nilpotent Jordan block, i.e. $T_{J_a} : \mathbb{C}^a \to \mathbb{C}^a$ where $J_a$ is the matrix

$$J_a = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \quad (10)$$

In other words,

$$J_a e_1 = e_2, J_a e_2 = e_3, \ldots, J_a e_{a-1} = e_a, J_a e_a = 0. \quad (11)$$

Notice that the powers of $J_a$ are very easy to compute. In fact $J_a^a = 0_{a,a}$ and for $d = 1, \ldots, a - 1$, we have

$$J_a^d e_1 = e_{d+1}, J_a^d e_2 = e_{d+2}, \ldots, J_a^d e_{a-d} = e_a, J_a^d e_{a+1-d} = 0, \ldots, J_a^d e_a = 0. \quad (12)$$

Notice that we have $\ker(J_a^d) = \text{span}(e_{a+1-d}, e_{a+2-d}, \ldots, e_a)$.

A nilpotent matrix $C \in M_{m \times m}(\mathbb{C})$ is said to be in Jordan normal form if it is of the form

$$C = \begin{pmatrix} J_{a_1} & 0_{a_1 \times a_2} & \cdots & 0_{a_1 \times a_t} & 0_{a_1 \times b} \\ 0_{a_2 \times a_1} & J_{a_2} & \cdots & 0_{a_2 \times a_t} & 0_{a_2 \times b} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{a_t \times a_1} & 0_{a_t \times a_2} & \cdots & J_{a_t} & 0_{a_t \times b} \\ 0_{b \times a_1} & 0_{b \times a_2} & \cdots & 0_{b \times a_t} & 0_{b \times b} \end{pmatrix}, \quad (13)$$

where $a_1 \geq a_2 \geq \cdots \geq a_t \geq 2$ and $a_1 + \cdots + a_t + b = m$.

We say that a basis $\mathcal{B}^{(i)}$ puts $T^{(i)}$ in Jordan normal form if $C^{(i)}$ is in Jordan normal form. We say that a basis $\mathcal{B} = (\mathcal{B}^{(1)}, \ldots, \mathcal{B}^{(r)})$ puts $T$ in Jordan normal form if each $\mathcal{B}^{(i)}$ puts $T^{(i)}$ in Jordan normal form.

**WARNING:** Usually such a basis is not unique. For example, if $T$ is diagonalizable, then ANY basis $\mathcal{B}^{(i)}$ puts $T^{(i)}$ in Jordan normal form.

## 2. Algorithm

In this section we present the general algorithm for finding bases $\mathcal{B}^{(i)}$ which put $T$ in Jordan normal form.

Suppose that we already had such bases. How could we describe the basis vectors? One observation is that for each Jordan block $J_a$, we have that $e_{d+1} = J_a^d(e_1)$ and also that $\text{span}_1$ and $\ker(J_a^{a-1})$ give a direct sum decomposition of $\mathbb{C}^a$.

What if we have two Jordan blocks, say

$$J = \begin{pmatrix} J_{a_1} & 0_{a_1 \times a_2} \\ 0_{a_2 \times a_1} & J_{a_2} \end{pmatrix}, a_1 \geq a_2. \quad (14)$$
This is the matrix such that
\[ Je_1 = e_2, \ldots, Je_{a_1 - 1} = e_{a_1}, Je_{a_1} = 0, Je_{a_1 + 1} = e_{a_1 + 2}, \ldots, Je_{a_1 + a_2 - 1} = e_{a_1 + a_2}, Je_{a_1 + a_2} = 0. \] (15)

Again we have that \( e_{d+1} = J^d e_1 \) and \( e_{d+a_1+1} = J^d e_{a_1+1}. \) So if we wanted to reconstruct this basis, what we really need is just \( e_1 \) and \( e_{a_1+1}. \) We have already seen that a distinguishing feature of \( e_1 \) is that it is an element of \( \ker(J^a) \) which is not in \( \ker(J^{a_1-1}). \) If \( a_2 = a_1, \) then this is also a distinguishing feature of \( e_{a_1+1}. \) But if \( a_2 < a_1, \) this doesn’t work. In this case it turns out that the distinguishing feature is that \( e_{a_1+1} \) is in \( \ker(J^{a_2}) \) but is not in \( \ker(J^{a_2-1}) + J(\ker(J^{a_2+1})). \) This motivates the following definition:

**Definition 1.** Suppose that \( B \in M_{n \times n}(\mathbb{C}) \) is a matrix such that \( \ker(B^e) = \ker(B^{e+1}). \) For each \( k = 1, \ldots, e, \) we say that a subspace \( G_k \subset \ker(B^k) \) is primitive (for \( k \)) if

1. \( G_k + \ker(B^{k-1}) + B(\ker(B^{k+1})) = \ker(B^k), \) and
2. \( G_k \cap (\ker(B^{k-1}) + B(\ker(B^{k+1}))) = \{0\} \).

Here we make the convention that \( B^0 = I_n. \)

It is clear that for each \( k \) we can find a primitive \( G_k: \) simply find a basis for \( \ker(B^{k-1}) + B(\ker(B^{k+1})) \) and then extend it to a basis for all of \( \ker(B^k). \) The new basis vectors will span a primitive \( G_k. \)

Now we are ready to state the algorithm. Suppose that \( T \) is as in the previous section. For each eigenvalue \( \lambda_i, \) choose any basis \( C \) for \( V \) and let \( A = [T]_c.c. \) Define \( B = A - \lambda_i I_n. \) Let \( 1 \leq k_1 < \cdots < k_u \leq n \) be the distinct integers such that there exists a nontrivial primitive subspace \( G_{k_j}. \) For each \( j = 1, \ldots, u, \) choose a basis \( (v[j]_1, \ldots, v[j]_{p_j}) \) for \( G_{k_j}. \) Then the desired basis is simply

\[
B^{(j)} = (v[u]_1, Bv[u]_1, \ldots, B^{k_{u-1}}v[u]_1, v[u]_2, Bv[u]_2, \ldots, B^{k_{u-1}}v[u]_2, \ldots, v[u]_{p_u}, B^{k_{u-1}}v[u]_{p_u}, \ldots, v[j]_1, Bv[j]_1, \ldots, B^{k_{j-1}}v[j]_1, \ldots, v[j]_{p_j}, B^{k_{j-1}}v[j]_{p_j}).
\]

When we perform this for each \( i = 1, \ldots, r, \) we get the desired basis for \( V. \)

### 3. Small cases

The algorithm above sounds more complicated than it is. To illustrate this, we will present a step-by-step algorithm in the \( 2 \times 2 \) and \( 3 \times 3 \) cases and illustrate with some examples.

#### 3.1. Two-by-two matrices.

First we consider the two-by-two case. If \( A \in M_{2 \times 2}(\mathbb{C}) \) is a matrix, its characteristic polynomial \( c_A(X) \) is a quadratic polynomial. The first dichotomy is whether \( c_A(X) \) has two distinct roots or one repeated root.

**Two distinct roots** Suppose that \( c_A(X) = (X - \lambda_1)(X - \lambda_2) \) with \( \lambda_1 \neq \lambda_2. \) Then for each \( i = 1, 2 \) we form the matrix \( B_i = A - \lambda_i I_2. \) By performing Gauss-Jordan elimination we may find a basis for \( \ker(B_i). \) In fact each kernel will be one-dimensional, so let \( v_1 \) be a basis
for \( \text{ker}(B_1) \) and let \( v_2 \) be a basis for \( \text{ker}(B_2) \). Then with respect to the basis \( B = (v_1, v_2) \), we will have

\[
[A]_B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.
\]  

(16)

Said a different way, if we form the matrix \( P = (v_1|v_2) \) whose first column is \( v_1 \) and whose second column is \( v_2 \), then we have

\[
A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.
\]  

(17)

To summarize:

\[
\text{span}(v_1) = E_{\lambda_1} = \text{ker}(A - \lambda_1 I_2) = \text{ker}(A - \lambda_1 I_2)^2 = \ldots = E_{\lambda_1}^{\text{gen}},
\]  

(18)

\[
\text{span}(v_2) = E_{\lambda_2} = \text{ker}(A - \lambda_2 I_1) = \text{ker}(A - \lambda_2 I_2)^2 = \ldots = E_{\lambda_2}^{\text{gen}}.
\]  

(19)

Setting \( B = (v_1, v_2) \) and \( P = (v_1|v_2) \), We also have

\[
[A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, A = P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.
\]  

(20)

Also \( S = A \) and \( N = 0_{2\times2} \).

Now we consider an example. Consider the matrix

\[
A = \begin{pmatrix} 38 & -70 \\ 21 & -39 \end{pmatrix}.
\]  

(21)

The characteristic polynomial is \( X^2 - \text{trace}(A)X + \det(A) \), which is \( X^2 + X - 12 \). This factors as \( (X + 4)(X - 3) \), so we are in the case discussed above. The two eigenvalues are \(-4\) and \(3\).

First we consider the eigenvalue \( \lambda_1 = -4 \). Then we have

\[
B_1 = A + 4I_2 = \begin{pmatrix} 42 & -70 \\ 21 & -35 \end{pmatrix}.
\]  

(22)

Performing Gauss-Jordan elimination on this matrix gives a basis of the kernel: \( v_1 = (5,3)^\dagger \).

Next we consider the eigenvalue \( \lambda_2 = 3 \). Then we have

\[
B_2 = A - 3I_2 = \begin{pmatrix} 35 & -70 \\ 21 & -42 \end{pmatrix}.
\]  

(23)

Performing Gauss-Jordan elimination on this matrix gives a basis of the kernel: \( v_2 = (2,1)^\dagger \).

We conclude that:

\[
E_{-4} = \text{span}\left( \begin{pmatrix} 5 \\ 3 \end{pmatrix} \right), E_3 = \text{span}\left( \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right).
\]  

(24)

and that

\[
A = P \begin{pmatrix} -4 & 0 \\ 0 & 3 \end{pmatrix} P^{-1}, P = \begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}.
\]  

(25)
One repeated root: Next suppose that $c_A(X)$ has one repeated root: $c_A(X) = (X - \lambda_1)^2$. Again we form the matrix $B_1 = A - \lambda_1 I_2$. There are two cases depending on the dimension of $E_{\lambda_1} = \ker(B_1)$. The first case is that $\dim(E_{\lambda_1}) = 2$. In this case $A$ is diagonalizable. In fact, with respect to some basis $\mathcal{B}$ we have

$$[A]_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}. \quad (26)$$

But, if you think about it, this means that $A$ has the above form with respect to ANY basis. In other words, our original matrix, expressed with respect to any basis, is simply $\lambda_1 I_2$. This case is readily identified, so if $A$ is not already in diagonal form at the beginning of the problem, we are in the second case.

In the second case $E_{\lambda_1}$ has dimension 1. According to our algorithm, we must find a primitive subspace $G_2 \subset \ker(B_1^2) = \mathbb{C}^2$. Such a subspace necessarily has dimension 1, i.e. it is of the form $\text{span}(v_1)$ for some $v_1$. And the condition that $G_2$ be primitive is precisely that $v_1 \notin \ker(B_1)$. In other words, we begin by choosing ANY vector $v_1 \notin \ker(B_1)$. Then we define $v_2 = B(v_1)$. We form the basis $\mathcal{B} = (v_1, v_2)$, and the transition matrix $P = (v_1|v_2)$. Then we have $E_{\lambda_1} = \text{span}(v_2)$ and also

$$[A]_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}, \quad A = P \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} P^{-1}. \quad (27)$$

This is the one case where we have nontrivial nilpotent part:

$$S = \lambda_1 I_2 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \quad N = A - \lambda_1 I_2 = B_1 = P \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} P^{-1}. \quad (28)$$

Let’s see how this works in an example. Consider the matrix from the practice problems:

$$A = \begin{pmatrix} -5 & -4 \\ 1 & -1 \end{pmatrix}. \quad (29)$$

The trace of $A$ is $-6$ and the determinant is $(-5)(-1) - (-4)(1) = 9$. So $c_A(X) = X^2 + 6X + 9 = (X + 3)^2$. So the characteristic polynomial has a repeated root of $\lambda_1 = -3$. We form the matrix $B_1 = A + 3I_2$,

$$B_1 = A + 3I_2 = \begin{pmatrix} -2 & -4 \\ 1 & 2 \end{pmatrix}. \quad (30)$$

Performing Gauss-Jordan elimination (or just by inspection) a basis for the kernel is given by $(2, -1)^\dagger$. So for $v_1$ we choose ANY vector which is not a multiple of this vector, for example $v_1 = e_1 = (1, 0)^\dagger$. Then we find that $v_2 = B_1 v_1 = (-2, 1)^\dagger$. So we define

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}\right), \quad P = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}. \quad (31)$$

Then we have

$$[A]_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} -3 & 0 \\ 1 & -3 \end{pmatrix}, \quad A = P \begin{pmatrix} -3 & 0 \\ 1 & -3 \end{pmatrix} P^{-1}. \quad (32)$$

The semisimple part is just $S = -3I_2$, and the nilpotent part is:

$$N = B_1 = P \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} P^{-1}. \quad (33)$$
3.2. Three-by-three matrices. This is basically as in the last subsection, except now there are more possible types of $A$. The first question to answer is: what is the characteristic polynomial of $A$. Of course we know this is $c_A(X) = \det(XI_3 - A)$. But a faster way of calculating this is as follows. We know that the characteristic polynomial has the form
\[ c_A(X) = X^3 - \text{trace}(A)X^2 + tX - \det(A), \]
for some complex number $t \in \mathbb{C}$. Usually trace($A$) and det($A$) are not hard to find. So it only remains to determine $t$. This can be done by choosing any convenient number $c \in \mathbb{C}$ other than $c = 0$, computing det($cI_2 - A$) (here it is often useful to choose $c$ equal to one of the diagonal entries to reduce the number of computations), and then solving the one linear equation
\[ ct + (c^3 - \text{trace}(A)c^2 - \det(A)) = \det(cI_2 - A), \]
to find $t$. Let’s see an example of this:
\[ D = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 1 & 2 \\ 0 & -1 & 3 \end{pmatrix}. \]
Here we easily compute trace($D$) = 6 and det($D$) = 8. Finally to compute the coefficient $t$, we set $c = 2$ and we get
\[ \det(2I_2 - A) = \det \begin{pmatrix} 0 & -1 & 1 \\ -1 & 1 & -2 \\ 0 & 1 & -1 \end{pmatrix} = 0. \]
Plugging this in, we get
\[ (2)^3 - 6(2)^2 + t(2) - 8 = 0 \]
or $t = 12$, i.e. $c_A(X) = X^3 - 6X^2 + 12X - 8$. Notice from above that 2 is a root of this polynomial (since det($2I_3 - A$) = 0). In fact it is easy to see that $c_A(X) = (X - 2)^3$.

Now that we know how to compute $c_A(X)$ in a more efficient way, we can begin our analysis. There are three cases depending on whether $c_A(X)$ has three distinct roots, two distinct roots, or only one root.

Three roots: Suppose that $c_A(X) = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$ where $\lambda_1, \lambda_2, \lambda_3$ are distinct. For each $i = 1, 2, 3$ define $B_i = \lambda_i I_3 - A$. By Gauss-Jordan elimination, for each $B_i$ we can compute a basis for ker($B_i$). In fact each ker($B_i$) has dimension 1, so we can find a vector $v_i$ such that $E_{\lambda_i} = \ker(B_i) = \text{span}(v_i)$. We form a basis $B = (v_1, v_2, v_3)$ and the transition matrix $P = (v_1|v_2|v_3)$. Then we have
\[ [A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}, A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} P^{-1}. \]
We also have $S = A$ and $N = 0$. 
Let’s see how this works in an example. Consider the matrix

\[
A = \begin{pmatrix}
7 & -7 & 2 \\
8 & -8 & 2 \\
4 & -4 & 1
\end{pmatrix}.
\]  

(40)

It is easy to see that \(\text{trace}(A) = 0\) and also \(\det(A) = 0\). Finally we consider the determinant of \(I_3 - A\). Using cofactor expansion along the third column, this is:

\[
\det \begin{pmatrix}
-6 & 7 & -2 \\
-8 & 9 & -2 \\
-4 & 4 & 0
\end{pmatrix} = -2((-8)4 - 9(-4)) - (-2)((-6)4 - 7(-4)) = -2(4) + 2(4) = 0.
\]  

(41)

So we have the linear equation

\[
1^3 - 0 \cdot 1^2 + t \cdot 1 - 0 = 0, t = -1.
\]  

(42)

Thus \(c_A(X) = X^3 - X = (X + 1)X(X - 1)\). So \(A\) has the three eigenvalues \(\lambda_1 = -1, \lambda_2 = 0, \lambda_3 = 1\). We define \(B_1 = A - (-1)I_3, B_2 = A, B_3 = A - I_3\). By Gauss-Jordan elimination we find

\[
E_{-1} = \ker(B_1) = \text{span} \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, E_0 = \ker(B_2) = \text{span} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},
\]

\[
E_1 = \ker(B_3) = \text{span} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.
\]

We define

\[
\mathcal{B} = \begin{pmatrix} 3 & 1 & 2 \\
4 & 1 & 2 \\
2 & 0 & 1
\end{pmatrix}, P = \begin{pmatrix} 3 & 1 & 2 \\
4 & 1 & 2 \\
2 & 0 & 1
\end{pmatrix}.
\]  

(43)

Then we have

\[
[A]_{\mathcal{B}, \mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}, A = P \begin{pmatrix} -1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix} P^{-1}.
\]  

(44)

**Two roots:** Suppose that \(c_A(X)\) has two distinct roots, say \(c_A(X) = (X - \lambda_1)^2(X - \lambda_2)\). Then we form \(B_1 = A - \lambda_1 I_3\) and \(B_2 = A - \lambda_2 I_3\). By performing Gauss-Jordan elimination, we find bases for \(E_{\lambda_1} = \ker(B_1)\) and for \(E_{\lambda_2} = \ker(B_2)\). There are two cases depending on the dimension of \(E_{\lambda_1}\).

The first case is when \(E_{\lambda_1}\) has dimension 2. Then we have a basis \((v_1, v_2)\) for \(E_{\lambda_1}\) and a basis \(v_3\) for \(E_{\lambda_2}\). With respect to the basis \(\mathcal{B} = (v_1, v_2, v_3)\) and defining \(P = (v_1|v_2|v_3)\), we have

\[
[A]_{\mathcal{B}, \mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & 0 \\
0 & \lambda_1 & 0 \\
0 & 0 & \lambda_2
\end{pmatrix}, A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\
0 & \lambda_1 & 0 \\
0 & 0 & \lambda_2
\end{pmatrix} P^{-1}.
\]  

(45)

In this case \(S = A\) and \(N = 0\).
The second case is when $E_{\lambda_1}$ has dimension 2. Using Gauss-Jordan elimination we find a basis for $E_{\lambda_1}^{\text{gen}} = \ker(B_1^*).$ Choose any vector $v_1 \in E_{\lambda_1}^{\text{gen}}$ which is not in $E_{\lambda_1}$ and define $v_2 = B_1 v_1.$ Also using Gauss-Jordan elimination we may find a vector $v_3$ which forms a basis for $E_{\lambda_2}.$ Then with respect to the basis $B = (v_1, v_2, v_3)$ and forming the transition matrix 

$$P = (v_1|v_2|v_3),$$

we have

$$[A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}, A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} P^{-1}. \tag{46}$$

Also we have

$$[S]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}, S = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} P^{-1}, \tag{47}$$

and

$$[N]_{B,B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A = P \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}. \tag{48}$$

Let’s see how this works in an example. Consider the matrix

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & -1 \\ -1 & 0 & 2 \end{pmatrix}. \tag{49}$$

It isn’t hard to show that $c_A(X) = (X - 3)^2(X - 2).$ So the two eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 2.$ We define the two matrices

$$B_1 = A - 3I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & -1 \end{pmatrix}, B_2 = A - 2I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}. \tag{50}$$

By Gauss-Jordan elimination we calculate that $E_2 = \ker(B_2)$ has a basis consisting of $v_3 = (0, 1, 1)\dagger.$ By Gauss-Jordan elimination, we find that $E_3 = \ker(B_1)$ has a basis consisting of $(0, 1, 0)\dagger.$ In particular it has dimension 1, so we have to keep going. We have

$$B_1^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}. \tag{51}$$

By Gauss-Jordan elimination (or inspection), we conclude that a basis consists of $(1, 0, -1)\dagger,$ $(0, 1, 0)\dagger.$ A vector in $E_3^{\text{gen}} = \ker(B_1^2)$ which isn’t in $E_3$ is $v_1 = (1, 0, -1)\dagger.$ We define $v_2 = B_1 v_1 = (0, 1, 0)\dagger.$ Then with respect to the basis

$$B = \begin{pmatrix} (1) & (0) & (0) \\ (0) & (1) & (1) \\ (-1) & (0) & (1) \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}. \tag{52}$$

we have

$$[A]_{B,B} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}, A = P \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} P^{-1}. \tag{53}$$
We also have that
\[
[S]_{B,B} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad S = P \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} P^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 3 & -1 \\ -1 & 0 & 2 \end{pmatrix}, \quad (54)
\]
\[
[N]_{B,B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad N = P \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \quad (55)
\]

**One root:** The final case is when there is only a single root of \(c_A(X)\), say \(c_A(X) = (X - \lambda_1)^3\). Again we form \(B_1 = A_1 - \lambda_1 I_3\). This case breaks up further depending on the dimension of \(E_{\lambda_1} = \ker(B_1)\). The simplest case is when \(E_{\lambda_1}\) is three-dimensional, because in this case \(A\) is diagonal with respect to ANY basis and there is nothing more to do.

**Dimension 2** Suppose that \(E_{\lambda_1}\) is two-dimensional. This is a case in which both \(G_1\) and \(G_2\) are nontrivial. We begin by finding a basis \((v_1, v_2)\) for \(E_{\lambda_1}\). Choose any vector \(v_1\) which is not in \(E_{\lambda_1}\) and define \(v_2 = B_1 v_1\). Then find a vector \(v_3\) in \(E_{\lambda_1}\) which is NOT in the span of \(v_2\). Define the basis \(B = (v_1, v_2, v_3)\) and the transition matrix \(P = (v_1|v_2|v_3)\). Then we have
\[
[A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \quad A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} P^{-1}. \quad (56)
\]
Notice that there is a Jordan block of size 2 and a Jordan block of size 1. Also, \(S = \lambda_1 I_3\) and we have \(N = B_1\).

Let’s see how this works in an example. Consider the matrix
\[
A = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (57)
\]
It is easy to compute \(c_A(X) = (X+2)^3\). So the only eigenvalue of \(A\) is \(\lambda_1 = -2\). We define \(B_1 = A - (-2)I_3\), and we have
\[
B_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (58)
\]
By Gauss-Jordan elimination, or by inspection, we see that \(E_{-2} = \ker(B_1)\) has a basis \(((1,1,0)^\dagger, (0,0,1)^\dagger)\). Since this is 2-dimensional, we are in the case above. So we choose any vector not in \(E_{-2}\), say \(v_1 = (1, 0, 0)^\dagger\). We define \(v_2 = B_1 v_1 = (1, 1, 0)^\dagger\). Finally, we choose a vector in \(E_{\lambda_1}\) which is not in the span of \(v_2\), say \(v_3 = (0,0,1)^\dagger\). Then we define
\[
B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (59)
\]
We have

$$[A]_{B,B} = \begin{pmatrix} -2 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, A = P \begin{pmatrix} -2 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} P^{-1}. \quad (60)$$

We also have $S = -2I_3$ and $N = B_1$.

**Dimension One** In the final case for three by three matrices, we could have that $c_A(X) = (X - \lambda_1)^3$ and $E_{\lambda_1} = \ker(B_1)$ is one-dimensional. In this case we must also have $\ker(B_1^2)$ is two-dimensional. By Gauss-Jordan we compute a basis for $\ker(B_1^2)$ and then choose ANY vector $v_1$ which is not contained in $\ker(B_1^2)$. We define $v_2 = B_1v_1$ and $v_3 = B_1v_2 = B_1^2v_1$. Then with respect to the basis $B = (v_1, v_2, v_3)$ and the transition matrix $P = (v_1|v_2|v_3)$, we have

$$[A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 1 & \lambda_1 \end{pmatrix}, A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 1 & \lambda_1 \end{pmatrix} P^{-1}. \quad (61)$$

We also have $S = \lambda_1I_3$ and $N = B_1$.

Let’s see how this works in an example. Consider the matrix

$$A = \begin{pmatrix} 5 & -4 & 0 \\ 1 & 1 & 0 \\ 2 & -3 & 3 \end{pmatrix}. \quad (62)$$

The trace is visibly 9. Using cofactor expansion along the third column, the determinant is $+3(5 \times 1 - 1(-4)) = 27$. Finally, we compute $\det(3I_3 - A) = 0$ since $3I_3 - A$ has the zero vector for its third column. Plugging in this gives the linear relation

$$(3)^3 - 9(3)^2 + t(3) - 27 = 0, t = 27. \quad (63)$$

So we have $c_A(X) = X^3 - 9X^2 + 27X - 27$. Also we see from the above that $X = 3$ is a root. In fact it is easy to see that $c_A(X) = (X - 3)^3$. So $A$ has the single eigenvalue $\lambda_1 = 3$.

We define $B_1 = A_1 - 3I_3$, which is

$$B_1 = \begin{pmatrix} 2 & -4 & 0 \\ 1 & -2 & 0 \\ 2 & -3 & 0 \end{pmatrix}. \quad (64)$$

By Gauss-Jordan elimination we see that $E_{3} = \ker(B_1)$ has basis $(0, 0, 1)^\dagger$. Thus we are in the case above. Now we compute

$$B_1^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -2 & 0 \end{pmatrix}. \quad (65)$$

Either by Gauss-Jordan elimination or by inspection, we see that $\ker(B_1^2)$ has basis $((2, 1, 0)^\dagger, (0, 0, 1)^\dagger)$. So for $v_1$ we choose any vector not in the span of these vectors, say $v_1 = (1, 0, 0)^\dagger$. Then we define $v_2 = B_1v_1 = (2, 1, 2)^\dagger$ and we define $v_3 = B_1v_2 = B_1^2v_1 = (0, 0, 1)^\dagger$. So with respect to
the basis and transition matrix
\[ \mathcal{B} = \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \] (66)
we have
\[ [A]_{\mathcal{B}, \mathcal{B}} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}, \quad A = P \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix} P^{-1}. \] (67)
We also have \( S = 3I_3 \) and \( N = B_1 \).