Hence the canonical form of N is given by

6.7 Jordan Canonical Form

We shall now briefly look at the Jordan cannonical form of a general matrix $A \in \mathbb{F}^{n \times n}$ whose characteristic polynomial and minimal polynomial are given as usual by

$$c_{A}(\lambda) = (\lambda - \lambda_{1})^{a_{1}}(\lambda - \lambda_{2})^{a_{2}} \cdots (\lambda - \lambda_{k})^{a_{k}}$$

$$(6.7.1)$$

$$m_A(\lambda) = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$
 (6.7.2)

In Theorem 6.5.3 we have seen that any such A can be written as the sum D+N of a diagonalizable matrix D and a nilpotent matrix N and that this decomposition is unique if D and N commute. The simplest such sum is one where D is a diagonal matrix and N is a canonical nilpotent matrix. The simplest diagonal D is the diagonal matrix all of whose diagonal entries are the same, say a. Thus the simplest such general matrix D+N we can think of is a matrix whose diagonal entries are all a and the entries to the right of the diagonal are 1 and all other entries are zero. Such a matrix is called a "canonical Jordan matrix". We have

Definition 6.7.1 The $m \times m$ matrix given below is called the "CANON-ICAL $m \times m$ JORDAN MATRIX with diagonal a" and is denoted by $J_m(a)$:

$$J_m(a) \; = \; \left(egin{array}{cccc} a & 1 & & & & \ & a & 1 & & & \ & & \ddots & \ddots & & \ & & & a & 1 \ & & & & a \end{array}
ight)_{m imes m}$$

It turns out that all matrices $A \in \mathbb{F}^{n \times n}$ whose characteristic polynomial and minimal polynomial are given by (6.7.1) and (6.7.2) can be built by such canonical Jordan matrices as follows:

We can find an invertible matrix $P \in \mathbb{F}^{n \times n}$ such that

$$P^{-1}AP = A_{can}$$

where A_{can} has the structure which we describe in detail below: $A_{can} \in \mathbb{F}n \times n$ is divided into k diagonal blocks $A^{(1)}, A^{(2)}, \cdots, A^{(j)}, \cdots, A^{(k)}$, one block corresponding to each of the k distinct eignevalues $\lambda_1, \lambda_2, \cdots, \lambda_j, \cdots, \lambda_k$ respectively. Thus we have

$$P^{-1}AP = \begin{pmatrix} A^{(1)} & & & & & & \\ & A^{(2)} & & & & & \\ & & \ddots & & & & \\ & & & A^{(j)} & & & \\ & & & & A^{(k)} \end{pmatrix}$$

How is each of these blocks structured?

We shall describe, in general, how $A^{(j)}$ is structured.

 $A^{(j)}$ is further subdivided into diagonal subblocks, each subblock being a canonical Jordan matrix with diagonal λ_j . What remains to be answered are the following questions:

- 1. What is the size of $A^{(j)}$?,
- 2. How many subblocks are there in $A^{(j)}$?,
- 3. How are these subblocks arranged?, and
- 4. What are the sizes of these blocks?

We now proceed to answer these questions.

Question 1:

What is the size of $A^{(j)}$?

 $A^{(j)}$ will be an $a_j \times a_j$ matrix

Question 2:

How many subblocks are there in $A^{(j)}$?

The number of subblocks in $A^{(j)}$ is equal to the geometric multiplicity g_i of

the eigenvalue λ_i .

Note

Hence, we observe that if $g_j = a_j$ then each subblock is just 1×1 and therefore $A^{(j)}$ will just be an $a_j \times a_j$ diagonal matrix all of whose diagonal entries are λ_j

Question 3:

How are these subblocks arranged?

These subblocks are arranged in such a way that as we go down the diagonal in $A^{(j)}$, the sizes of the subblocks are nonincreasing

It remains to answer the fourth question.

Question 4:

What are the sizes of these blocks?

The structure is analogous to what we did for nilpotent matrices. We define the following subspaces

$$V_r^{(j)} = Null \ Space \ of \ (A - \lambda_j I_{n \times n})^r \ \text{for} \ 1 \le r \le r_j$$
 (6.7.3)

We then define the dimensions of these subspaces,

$$d_r^{(j)} = dim.(V_r^{(j)}), \ 1 \le j \le r_j$$
 (6.7.4)

Then we have

$$g_j = d_1^{(j)} < d_2^{(j)} < \dots < d_{r_j-1}^{(j)} < d_{r_j}^{(j)} = a_j$$
 (6.7.5)

We then define

$$\alpha_r^{(j)} = d_r^{(j)} - d_{r-1}^{(j)} \text{ for } 2 \le r \le r_j$$

$$(6.7.6)$$

These $\alpha_r^{(j)}$ give us the extra dimension we get when we move up from $V_{(r-1)}^{(j)}$

to $V_r^{(j)}$. We shall define $\alpha_1^{(j)} = g_j$ and $\alpha_{r_j+1}^{(j)} = 0$. Then we can write

$$d_{1}^{(j)} = \alpha_{1}^{(j)} \text{ (the geometric multiplicity of the eigenvalue 0)} \\ d_{2}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} \\ d_{3}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} + \alpha_{3}^{(j)} \\ \cdots \cdots \cdots \\ d_{(r-1)}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} + \alpha_{3}^{(j)} + \cdots + \alpha_{(r-1)}^{(j)} \\ d_{r}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} + \alpha_{3}^{(j)} + \cdots + \alpha_{(r-1)}^{(j)} + \alpha_{r}^{(j)} \\ d_{(r+1)}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} + \alpha_{3}^{(j)} + \cdots + \alpha_{(r-1)}^{(j)} + \alpha_{r}^{(j)} + \alpha_{(r+1)}^{(j)} \\ \cdots \cdots \cdots \cdots \\ d_{(r_{j}-1)}^{(j)} = d_{1}^{(j)} + \alpha_{2}^{(j)} + \alpha_{3}^{(j)} + \cdots + \alpha_{(r-1)}^{(j)} + \alpha_{r}^{(j)} + \alpha_{(r+1)}^{(j)} + \cdots + \alpha_{(r_{j}-1)}^{(j)} + \alpha_{r_{j}}^{(j)} = a_{j} \\ \end{pmatrix}$$

$$(6.7.7)$$

These $\alpha_r^{(j)}$ have a nice structure as follows:

$$\alpha_1^{(j)} \ge \alpha_2^{(j)} \ge \dots \ge \alpha_r^{(j)} \ge \alpha_{(r+1)}^{(j)} \ge \dots \ge \alpha_{(r_i-1)}^{(j)} \ge \alpha_{r_i}^{(j)}$$
 (6.7.8)

This means that the extra dimension that we acquire from moving up from $V_{(r-1)}^{(j)}$ to $V_r^{(j)}$ is at least as much as the extra dimension that we acquire in the next stage of moving up from $V_r^{(j)}$ to $V_{(r+1)}^{(j)}$.

We now look at the differences of the dimensions gained at two successive stages. More precisely we define,

(We observe that $\alpha_{r_j}^{(j)}$ is at least 1 and hence $n_{r_j}^{(j)}$ is at least one). We now state the following theorem (without proof):

Theorem 6.7.1 With the above notations, there exists an invertible matrix, $P \in \mathbb{F}^{n \times n}$ such that $P^{-1}AP = A_{can}$, where A_{can} is an $n \times n$ matrix which is built as follows: It is made up of diagonal blocks of matrices $A^{(1)}, A^{(2)}, \dots, A^{(j)}, \dots, A^{(k)}$. For each j, $(1 \leq j \leq k)$, the $A^{(j)}$ is divided into a number of subblocks of different sizes such that,

- 1. The subblocks are arranged in non increasing order as we go down the diagonal,
- 2. The leading subblock is $J_{r_j}(\lambda_j)$, the canonical $r_j \times r_j$ Jordan matrix with diagonal λ_j , (and all other blocks can be at most of this size),
- 3. For $1 \le r \le r_i$, there are $n_r^{(j)}$ blocks of size $r \times r$
- 4. The total number of subblocks in $A^{(j)}$ is equal to g_j , the geometric multiplicity of λ_j

 A_{can} is called the "Jordan Canonical Form" of the matrix A.

Analogous to the nilpotent matrices we can easily see that

$$n_r^{(j)} = number\ of\ r \times r\ subblocks\ in\ A^{(j)}$$

= $2dim.(V_r^{(j)}) - dim.(V_{(r-1)}^{(j)}) - dim.(V_{(r+1)}^{(j)})$

Reason:

We have

$$\begin{array}{lll} n_r^{(j)} & = & \alpha_r^{(j)} - \alpha_{(r+1)}^{(j)} \\ & = & (d_r^{(j)} - d_{(r-1)}^{(j)}) - (d_{(r+1)}^{(j)} - d_r^{(j)}) \\ & = & 2d_r^{(j)} - d_{(r-1)}^{(j)} - d_{(r+1)}^{(j)} \\ & = & 2dim.(V_r^{(j)}) - dim.(V_{(r-1)}^{(j)}) - dim.(V_{(r+1)}^{(j)}) \end{array}$$

Example 6.7.1 Suppose $A \in \mathbb{F}^{14 \times 14}$ has characteristic polynomial and minimal polynomial given by,

$$c_A(\lambda) = (\lambda - 2)^5 (\lambda + 4)^3 (\lambda - 6)^6$$

 $m_A(\lambda) = (\lambda - 2)^3 (\lambda + 4)^2 (\lambda - 6)^3$

Then the distinct eigenvalues are $\lambda_1 = 2$, $\lambda_2 = -4$ and $\lambda_3 = 6$, with algebraic multiplicities given by

$$a_1 = 5, \ a_2 = 3, a_3 = 6$$

Since there are three distinct eigenvalues there will be three major blocks $A^{(1)}$, $A^{(2)}$ and $A^{(3)}$ corresponding to the three distinct eigenvalues $\lambda_1 = 2, \lambda_2 = -4$ and $\lambda_3 = 6$, We have

$$A_{can} = \left(\begin{array}{c|c} A^{(1)} & & \\ \hline & A^{(2)} & \\ \hline & & A^{(3)} \end{array}\right)$$

The algebraic multiplicities of the eigenvalues give us the fact that $A^{(1)}$ is 5×5 , $A^{(2)}$ is 3×3 and $A^{(3)}$ is 6×6 . So we have

$$A_{can} = \left(\begin{array}{c|c} (A^{(1)})_{5\times5} & & \\ \hline & (A^{(2)})_{3\times3} & \\ \hline & & (A^{(3)})_{6\times6} \end{array}\right)$$
(6.7.10)

The minimal polynomial gives us

$$r_1 = 3$$
, $r_2 = 2$ and $r_3 = 3$

This means that,

1. The leading subblock in $A^{(1)}$ is

$$J_3(\lambda_1) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

2. The leading subblock in $A^{(2)}$ is

$$J_2(\lambda_2) = \begin{pmatrix} -4 & 1 \\ 0 & -4 \end{pmatrix}$$

3. The leading subblock in $A^{(3)}$ is

$$J_3(\lambda_3) = \begin{pmatrix} 6 & 1 & 0 \\ 0 & 6 & 1 \\ 0 & 0 & 6 \end{pmatrix}$$

To determine the remaining subblocks we need to look at the relevant subspaces $V_r^{(j)} = Null\ Space\ of\ (A - \lambda_j I_{14\times14})^r$. Suppose we are given the following data:

$$d_1^{(1)} = dim.(V_1^{(1)}) = Null Space of (A - 2I_{14\times14}) = 3$$

 $d_2^{(1)} = dim.(V_2^{(1)}) = Null Space of (A - 2I_{14\times14})^2 = 4$
 $d_3^{(1)} = dim.(V_3^{(1)}) = Null Space of (A - 2I_{14\times14})^3 = 5$

$$d_1^{(2)} = dim.(V_1^{(2)}) = Null Space of (A + 4I_{14\times14}) = 2$$

 $d_2^{(2)} = dim.(V_2^{(2)}) = Null Space of (A + 4I_{14\times14})^2 = 3$

$$d_1^{(3)} = dim.(V_1^{(3)}) = Null Space of (A - 6I_{14\times14}) = 3$$

 $d_2^{(3)} = dim.(V_2^{(3)}) = Null Space of (A - 6I_{14\times14})^2 = 5$
 $d_3^{(3)} = dim.(V_3^{(3)}) = Null Space of (A - 6I_{14\times14})^3 = 6$

Thus we have,

$$\alpha_1^{(1)} = d_1^{(1)} = 3$$

$$\alpha_2^{(1)} = d_2^{(1)} - d_1^{(1)} = 1$$

$$\alpha_3^{(1)} = d_3^{(1)} - d_2^{(1)} = 1$$

Hence we get

$$n_3^{(1)} = \alpha_3^{(1)} = 1$$

$$n_2^{(1)} = \alpha_2^{(1)} - \alpha_3^{(1)} = 0$$

$$n_1^{(1)} = \alpha_1^{(1)} - \alpha_2^{(1)} = 2$$

Thus in $A^{(1)}$, there is one subblock of 3×3 , (because $n_3^{(1)} = 1$), zero sublocks of size 2×2 , (because $n_2^{(1)} = 0$) and two subblocks of size 1×1 , (because

 $n_1^{(1)} = 2$). Thus we get,

Similarly we get

$$\alpha_1^{(2)} = d_1^{(2)} = 2$$

$$\alpha_2^{(2)} = d_2^{(2)} - d_1^{(2)} = 1$$

Hence we get

$$n_2^{(2)} = \alpha_2^{(2)} = 1$$

 $n_1^{(1)} = \alpha_1^{(2)} - \alpha_2^{(2)} = 1$

Thus in $A^{(2)}$, there is one subblock of 2×2 , (because $n_2^{(2)} = 1$), and one subblock of size 1×1 , (because $n_1^{(2)} = 1$). Thus we get,

$$A^{(2)} = \begin{pmatrix} -4 & 1 \\ 0 & -4 \\ \hline & | -4 \end{pmatrix}$$

$$= \begin{pmatrix} J_2(-4) \\ \hline & | J_1(-4) \end{pmatrix}$$
(6.7.13)

For the third eigenvalue we get,

$$\begin{array}{rcl} \alpha_1^{(3)} & = & d_1^{(3)} = 3 \\ \alpha_2^{(2)} & = & d_2^{(3)} - d_1^{(3)} = 2 \\ \alpha_3^{(3)} & = & d_3^{(3)} - d_2^{(3)} = 1 \end{array}$$

Hence we get

$$n_3^{(3)} = \alpha_3^{(3)} = 1$$

$$n_2^{(3)} = \alpha_2^{(3)} - \alpha_3^{(3)} = 1$$

$$n_1^{(3)} = \alpha_1^{(3)} - \alpha_2^{(3)} = 1$$

Thus in $A^{(3)}$, there is one subblock of 3×3 , (because $n_3^{(3)}=1$), one sublock of size 2×2 , (because $n_2^{(3)}=1$) and one subblock of size 1×1 , (because $n_1^{(3)}=1$). Thus we get,

Thus the canonical form of A is given by (6.7.10), where $A^{(1)}, A^{(2)}$ and $A^{(3)}$ are given by (6.4.12), (6.4.14) and (6.4.16). Thus we have

$$A_{can} = \begin{pmatrix} J_3(2) & & & & & & & & & \\ & J_1(2) & & & & & & & \\ & & & J_1(2) & & & & & & \\ & & & & & J_2(-4) & & & & \\ & & & & & & J_1(-4) & & & \\ & & & & & & & J_2(6) & \\ & & & & & & & & & J_1(6) \end{pmatrix}$$

6.8 Functions of a General Matrix

Just as we used the diagonal form to define f(A) for a diagonalizable matrix $A \in \mathbb{C}^{n \times n}$, for any function $f(\lambda)$ analytic in a disc containing the eigenvalues

of A, we can use the Jordan Canonical Form to define f(A) for a general matrix $A \in \mathbb{C}^{n \times n}$. We first observe that, if $J_m(a)$ is a canonical Jordan matrix, then

$$(J_m(a))^2 = \begin{pmatrix} a^2 & 2a & 1 & & & & \\ & a^2 & 2a & 1 & & & & \\ & & a^2 & 2a & 1 & & & \\ & & & \ddots & \ddots & & & \\ & & & & a^2 & 2a & 1 & \\ & & & & & a^2 & 2a \\ & & & & & & a^2 \end{pmatrix}_{m \times m}$$

By induction we can then show that, $(J_m(a))^r$ is the $m \times m$ matrix whose first row is such that the jth entry in the first row is given by

$$[(J_m(a))^r]_{1j} = {r \choose (j-1)} a^{r-(j-1)}$$
(6.8.1)

Thus the first row of $(J_m(a))^r$ is of the form,

$$\begin{pmatrix} a^r & ra^{(r-1)} & \begin{pmatrix} r \\ 2 \end{pmatrix} a^2 \cdot \dots \cdot \begin{pmatrix} r \\ (j-1) \end{pmatrix} a^{r-(j-1)} \cdot \dots \begin{pmatrix} r \\ (n-1) \end{pmatrix} a^{r-(n-1)} \end{pmatrix}$$

The next row is obtained by shifting the above to the right and so on. Thus the *i*th row of $(J_m(a))^r$ will have all entries upto the diagonal entry as zero and then start off with a^r and proceed as above. We have the *i*th row of $(J_m(a))^r$ is give by

$$\left(0\ 0\ \cdots\ \underbrace{a^r}_{i-th\ entry} \quad ra^{(r-1)} \quad \left(\begin{array}{c} r\\ 2 \end{array}\right)a^{r-2} \cdots \cdot \left(\begin{array}{c} r\\ (n-i) \end{array}\right)a^{r-(n-i)}\right)$$

Using this we can show that if $p(\lambda)$ is any poynomial in $\mathbb{F}[\lambda]$ then $p(J_m(a))$ is the matrix whose *i*th row is of the form

$$0 \ 0 \ \cdots \underbrace{p(a)}_{i-th \ entry} \left(\begin{array}{c} r \\ 1 \end{array}\right) p'(a) \quad \left(\begin{array}{c} r \\ 1 \end{array}\right) p''(a) \ \cdots \left(\begin{array}{c} r \\ (n-i) \end{array}\right) p^{(n-i)}(a)$$

where $p^{(m)}$ denotes the *m*th derivative of *p*.

Finally, using the fact that any function $f(\lambda)$ analytic in a disc containing a

can be approximated by a sequence of polynomials, we see that we can define $f(J_m(a))$ as the $m \times m$ matrix whose *i*th row is of the form

$$0 \ 0 \ \cdots \underbrace{f(a)}_{i-th \ entry} \left(\begin{array}{c} r \\ 1 \end{array}\right) f'(a) \quad \left(\begin{array}{c} r \\ 2 \end{array}\right) f''(a) \ \cdots \cdots \left(\begin{array}{c} r \\ (n-i) \end{array}\right) f^{(n-i)}(a)$$

Since the Jordan canonical form of any matrix A consists of diagonal subblocks of canonical Jordan matrices we define the Jordan canonical form of f(A) by taking f(subblock) for each of the subblocks, provided f is analytic in a disc containing all the eigenvalues of A. We note that if $\mu_1, \mu_2, \dots, \mu_n$ are the eigenvalues of A, written including their multiplicities, the eigenvalues of f(A) are $f(\mu_1), f(\mu_2), \dots, f(\mu_n)$.

Example 6.8.1 For the matrix of Example 6.7.1 we had the eigenvalues 2, -4 and 6. Hence if f is analytic in a disc D_R of radius R centred at the origin, where R > 6, then we can define f(A) as follows: We have

$$f(A^{(2)}) = \begin{pmatrix} f(-4) & f'(-4) & \\ 0 & f(-4) & \\ \hline & & f(-4) \end{pmatrix}$$
 (6.8.3)

Then the canonical form of f(A) is given by

$$f(A)_{can} = \left(\begin{array}{c|c} f(A^{(1)}) & & \\ \hline & f(A^{(2)}) & \\ \hline & & f(A^{(3)}) \end{array}\right)$$
(6.8.5)

where $f(A^{(1)}), f(A^{(2)})$ and $f(A^{(3)})$ are as in (6.8.1),(6.8.2) and (6.8.3). Thus, for example, if $f(\lambda) = \sin(\pi \lambda)$ we get from above,

$$sin(\pi A^{(1)}) = \begin{pmatrix} 0 & \pi & 0 & | \\ 0 & 0 & \pi & | \\ 0 & 0 & 0 & | \\ \hline & & & 0 \end{pmatrix}$$
 (6.8.6)

$$sin(\pi A^{(2)}) = \begin{pmatrix} 0 & \pi & \\ 0 & 0 & \\ \hline & & 0 \end{pmatrix}$$
 (6.8.7)

$$sin(\pi A^{(3)}) = \begin{pmatrix} 0 & \pi & 0 & & & \\ 0 & 0 & \pi & & & & \\ 0 & 0 & 0 & & & & \\ \hline & & & 0 & 0 & & \\ \hline & & & & 0 & 0 \end{pmatrix}$$
(6.8.8)

Finally we have

$$sin(\pi A) = \begin{pmatrix} sin(\pi A^{(1)} | & & \\ & sin(\pi A^{(2)}) | & \\ & & sin(\pi A^{(3)}) \end{pmatrix}$$
 (6.8.9)