Chapter 6

Jordan Canonical Form

6.1 Primary Decomposition Theorem

For a general matrix, we shall now try to obtain the analogues of the ideas we have in the case of diagonalizable matrices.

Diagonalizable Case	**General Case**
Minimal Polynomial: $$m_A(\lambda) = \prod_{j=1}^{k}(\lambda - \lambda_j)^{r_j}$$ where $r_j = 1$ for all j	Minimal Polynomial: $$m_A(\lambda) = \prod_{j=1}^{k}(\lambda - \lambda_j)^{r_j}$$ where $1 \leq r_j \leq a_j$ for all j
Eigenspaces: $$\mathcal{W}_j = Null\ Space\ of\ (A - \lambda_j I)$$ $dim(\mathcal{W}_j) = g_j$ and $g_j = a_j$ for all j	Eigenspaces: $$\mathcal{W}_j = Null\ Space\ of\ (A - \lambda_j I)$$ $dim(\mathcal{W}_j) = g_j$ and $g_j \leq a_j$ for all j Generalized Eigenspaces: $$\mathcal{V}_j = Null\ Space\ of\ (A - \lambda_j I)^{r_j}$$ $dim(\mathcal{V}_j) = a_j$
Decomposition: $$\mathbb{F}^n = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \mathcal{W}_j \oplus \cdots \oplus \mathcal{W}_k$$ $x \in \mathbb{F}^n \Longrightarrow \exists$ unique $x_j \in \mathcal{W}_j$, $1 \leq j \leq k$ such that $$x = x_1 + x_2 + \cdots + x_j + \cdots + x_k$$	Decomposition: $\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \mathcal{W}_j \oplus \cdots \oplus \mathcal{W}_k$ is a subspace of $\mathbb{F}^n$ **Primary Decomposition Theorem**: $$\mathbb{F}^n = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \mathcal{V}_j \oplus \mathcal{V}_k$$ $x \in \mathbb{F}^n \Longrightarrow \exists$ unique $v_j \in \mathcal{V}_j$, $1 \leq j \leq k$ such that $$x = v_1 + v_2 + \cdots + v_j + \cdots + v_k$$

6.2 Analogues of Lagrange Polynomial

Diagonalizable Case	**General Case**
Define $$f_j(\lambda) = \prod_{\substack{i=1 \\ i\neq j}}^{k} (\lambda - \lambda_i)$$ This is a set of coprime polynomials and hence there exist polynomials $g_j(\lambda)$, $1 \leq j \leq k$ such that $$\sum_{j=1}^{k} f_j(\lambda) g_j(\lambda) = 1$$ The polynomials $$\ell_j(\lambda) = f_j(\lambda) g_j(\lambda),\ 1 \leq j \leq k$$ are the Lagrange Interpolation polynomials	Define $$F_j(\lambda) = \prod_{\substack{i=1 \\ i\neq j}}^{k} (\lambda - \lambda_i)^{r_i}$$ This is a set of coprime polynomials and hence there exist polynomials $G_j(\lambda)$, $1 \leq j \leq k$ such that $$\sum_{j=1}^{k} F_j(\lambda) G_j(\lambda) = 1$$ The polynomials $$L_j(\lambda) = F_j(\lambda) G_j(\lambda),\ 1 \leq j \leq k$$ are the counterparts of the Lagrange Interpolation polynomials
$\ell_1(\lambda) + \ell_2(\lambda) + \cdots + \ell_k(\lambda) = 1$	$L_1(\lambda) + L_2(\lambda) + \cdots + L_k(\lambda) = 1$
$\lambda_1\ell_1(\lambda) + \lambda_2\ell_2(\lambda) + \cdots + \lambda_k\ell_k(\lambda) = \lambda$	$\lambda_1 L_1(\lambda) + \lambda_2 L_2(\lambda) + \cdots + \lambda_k L_k(\lambda) = d(\lambda)$ (a polynomial but it need not be $= \lambda$)
For $i \neq j$ the product $\ell_i(\lambda)\ell_j(\lambda)$ has a factor $m_{A}(\lambda)$	For $i \neq j$ the product $L_i(\lambda)L_j(\lambda)$ has a factor $m_{A}(\lambda)$

6.3 Matrix Decomposition

In the diagonalizable case, using the Lagrange polynomials we obtained a decomposition of the matrix A. We shall next look at the analogous situation for a general matrix.

Diagonalizable Case	**General Case**
Define the matrices $$A_j = \ell_j(A) \text{ for } 1 \le j \le k$$	Analogously define the matrices $$A_j = L_j(A) \text{ for } 1 \le j \le k$$
We have $$A_1 + A_2 + \cdots + A_k = I$$ since $$\ell_1(\lambda) + \ell_2(\lambda) + \cdots + \ell_k(\lambda) = 1$$	Analogously we have $$A_1 + A_2 + \cdots + A_k = I$$ since $$L_1(\lambda) + L_2(\lambda) + \cdots + L_k(\lambda) = 1$$
$$A_iA_j = 0_{(n\times n)} \text{ if } i \ne j$$ since, then $\ell_i(\lambda)\ell_j(\lambda)$ has the minimal polynomial as a factor	$$A_iA_j = 0_{(n\times n)}$$ since, then $L_i(\lambda)L_j(\lambda)$ has the minimal polynomial as a factor
$A_j^2 = A_j$ for $1 \le j \le k$	$A_j^2 = A_j$ for $1 \le j \le k$
Range of A_j is $\mathcal{W}_j$ and $A_jx = x$ for all $x \in \mathcal{W}_j$	Range of A_j is $\mathcal{V}_j$ and $A_jx = x$ for all $x \in \mathcal{V}_j$
We have $$\lambda_1A_1 + \lambda_2A_2 + \cdots + \lambda_kA_k = A$$ since $$\lambda_1\ell_1(\lambda) + \lambda_2\ell_2(\lambda) + \cdots + \lambda_k\ell_k(\lambda) = \lambda$$	We will not have $$\lambda_1A_1 + \lambda_2A_2 + \cdots + \lambda_kA_k = A$$ since $$\lambda_1L_1(\lambda) + \lambda_2L_2(\lambda) + \cdots + L_k(\lambda)\ ,$$ in general, may not be equal to λ. Let $$D = \lambda_1A_1 + \lambda_2A_2 + \cdots + \lambda_kA_k$$ Then D is a diagonalizable matrix (by Theorem 5.7.1).

We then look at by how much A differs form D. For this purpose we define $N = A - D$ so that $A = D + N$
(In the case where A is a diagonalizable matrix we have $N = 0_{(n\times n)}$ and $A = D$).
Thus, in the general case, in addition to a diagonalizable matrix, we also have to analyse the part $N = A - D$. We shall do this in the next section.

6.4 Nilpotent Matrices

We want to analyse the difference

$$N = A - D \tag{6.4.1}$$

where D is as defined in the previous section. Using the fact that

$$I = A_1 + A_2 + \cdots + A_k$$

we get

$$\begin{aligned} A &= AI_{n\times n} \\ &= AA_1 + AA_2 + \cdots + AA_k \end{aligned}$$

Combining this and the fact that

$$D = \lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k$$

we get

$$\begin{aligned} A - D &= (AA_1 + AA_2 + \cdots + AA_k) - (\lambda_1 A_1 + \lambda_2 A_2 + \cdots + \lambda_k A_k) \\ &= (A - \lambda_1 I_{n\times n})A_1 + (A - \lambda_2 I_{n\times n})A_2 + \cdots + (A - \lambda_k I_{n\times n})A_k \end{aligned}$$

All the A_j, being polynomials in A, commute with A, and furher, $A_i A_j = 0_{(n\times n)}$. Using these facts we get

$$(A - D)^r = (A - \lambda_1 I_{n\times n})^r A_1 + (A - \lambda_2 I_{n\times n})^r A_2 + \cdots + (A - \lambda_k I_{n\times n})^r A_k \tag{6.4.2}$$

If $r \geq \ Max.\{r_1, r_2, \cdots, r_j\}$ then $(A - \lambda_j I_{n\times n})^r A_j$ has a factor $m_{_A}(A)$, and hence

$$(A - \lambda_j I_{n\times n})^r A_j = 0_{n\times n} \text{ for all } j \text{ if } r \geq Max.\{r_1, r_2, \cdots, r_j\} \tag{6.4.3}$$

Hence we get from (6.4.2)

$$N^r = (A - D)^r \;\; = \;\; 0_{n\times n} \text{ for } r \geq Max.\{r_1, r_2, \cdots, r_j\} \qquad (6.4.4)$$

Thus the difference between A and the diagonalizable matrix D is such that a power of it vanishes. This leads us to the notion of a nilpotent matrix. We have

Definition 6.4.1 A matrix $N \in \mathbb{C}^{n\times n}$ is said to be a "NILPOTENT" matrix if there exists a positive integer r such that $N^r = 0_{n\times n}$. For a nilpotent matrix the smallest power r for which $N^r = 0_{n\times n}$ is called the "order of nilpotency" of N and is denoted by $\gamma_{_N}$.

Thus we have that the matrix $N = A - D$ is a nilpotent matrix the order of nilpotency being equal to $Max.\{r_1, r_2, \cdots, r_k\}$. We can summarise our analysis above as follows:

Theorem 6.4.1 Every matrix $A \in \mathbb{C}^{n\times n}$ can be expressed as the sum,

$$\boldsymbol{A} \;\; = \;\; \boldsymbol{D} + \boldsymbol{N}, \qquad (6.4.5)$$

of a diagonalizable matrix D and a nilpotent matrix N

We have already analysed diagonalizable matrices in the last chapter. We shall now analyse a nilpotent matrix.

6.5 Nilpotent Matrices

Let $N \in \mathbb{F}^{n\times n}$ be a nilpotent matrix with order of nilpotency $\gamma_{_N}$. This means that

$$N^{\gamma_{_N}} \;\; = \;\; 0_{n\times n} \text{ and} \qquad (6.5.1)$$

$$N^r \;\; \neq \;\; 0_{n\times n} \text{ for any } r < \gamma_{_N} \qquad (6.5.2)$$

From (6.5.1) we also get,

$$N^r \;\; = \;\; 0_{n\times n} \text{ for all } r \geq \gamma_{_N} \qquad (6.5.3)$$

From this it follows that the polynomial $p(\lambda) = \lambda^{\gamma_{_N}}$ is an annihilating polynomial for N. Hence the minimal polynomial must divide $\lambda^{\gamma_{_N}}$. Thus the minimal polynomial must be of the form $m_{_N}(\lambda) = \lambda^r$ where $r \leq \gamma_{_N}$. However, from (6.5.2) we have that $N^r \neq 0_{n\times n}$ for any $r < \gamma_{_N}$. Hence we get that the minimal polynomial of N is given by

$$m_{_N}(\lambda) \;\;=\;\; (\lambda)^{\gamma_N}$$

Thus we see that **$\boldsymbol{\lambda_1 = 0}$ is the only eigenvalue of any nilpotent matrix $\boldsymbol{N}$**. Since the characteristic polynomial of N is a monic polynomial of degree n and its roots are the eigenvalues we get that the characteristic polynomial of N is given by

$$c_{_N}(\lambda) \;\;=\;\; \lambda^n$$

Hence we have,

Theorem 6.5.1 If $\boldsymbol{N \in \mathbb{F}^{n\times n}}$ is a nilpotent matrix with order of nilpotency as $\boldsymbol{\gamma_{_N}}$ then its characteristic and minimal polynomials are given by

$$\boldsymbol{c_{_N}(\lambda)} \;\;=\;\; \boldsymbol{\lambda^n}, \textbf{ and} \tag{6.5.4}$$

$$\boldsymbol{m_{_N}(\lambda)} \;\;=\;\; \boldsymbol{\lambda^{\gamma_N}} \tag{6.5.5}$$

From the above simple properties we now obtain a useful result. Suppose $A \in \mathbb{F}^{n\times n}$ is both nilpotent and diagonalizable. Since it is nilpotent matrix its minimal polynomial is of the form $m_{_A} = \lambda^{\gamma_A}$, where $\gamma_{_A}$ is its order of nilpotency. Combining this with the fact that A is diagonalizable we get that its minimal polynomial must be $m_{_A}(\lambda) = \lambda$. Hence $m_{_A}(A) = 0_{n\times n}$ gives us that $A = 0_{n\times n}$. Thus it follows that

Theorem 6.5.2 The only matrix in $\boldsymbol{\mathbb{F}^{n\times n}}$ which is both diagonalizable and nilpotent is the zero matrix

We shall now see an example to show that the sum and product of two nilpotent matrices need not be nilpotent.

Example 6.5.1 . Consider the matrices $N_1, N_2 \in \mathbb{C}^{n\times n}$ given below:

$$N_1 \;\;=\;\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{6.5.6}$$

$$N_2 \;\;=\;\; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{6.5.7}$$

It is easy to see that

$$\begin{aligned} N_1^2 &= 0_{2\times 2} \\ N_2^2 &= 0_{2\times 2} \end{aligned}$$

Thus both are nilpotent matrices both with order of nilpotency as 2. However, we have,

$$N_1 + N_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

This matrix has characteristic polynomial given by

$$\lambda^2 - 1 = 0$$

Since there are two eigenvalues $\lambda_1 = 1,\ \lambda_2 = -1$, both eigenvalues having algebraic and geometric multiplicity both as 1, the matrix is diagonalizable. Since this is not the zero matrix, by the Theorem 6.5.2, this matrix cannot be nilpotent. Thus we see that the sum $N_1 + N_2$ of the two nilpotent matrices N_1 and N_2 is not nilpotent. Similarly we have,

$$N_1 N_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and this is a diagonal matrix. Again by Theorem 6.5.2 it follows that N_1N_2 is not nilpotent. Thus the product N_1N_2 of the two nilpotent matrices N_1 and N_2 is not nilpotent.

However, we can show that the following is true:
when two nilpotent matrices commute with each other, then their sum and product will also be nilpotent
We shall now use this result to investigate the uniqueness of the decomposition of a matrix as the sum $D + N$ of a diagonalizable matrix D and a nilpotent matrix N obtained in Section 6.3. The diagonalizable matrix D we got is a polynomial in A. Further the nilpotent matrix N we obtained was defined as $N = A - D$, and hence N is also a polynomial in A. Thus D, N and A all commute with each other. Thus we have the decomposition

$$A = D + N \text{ where} \tag{6.5.8}$$

$$DN = ND \tag{6.5.9}$$

Suppose now there exists another decomposition of A as

$$A = D_1 + N_1 \tag{6.5.10}$$

where D_1 is diagonalizable and N_1 is nilpotent and D_1 and N_1 commute with each other, that is,

$$D_1 N_1 = N_1 D_1 \tag{6.5.11}$$

Then we get from (6.5.13),

$$\begin{aligned} D_1 A &= D_1(D_1 + N_1) \\ &= D_1^2 + D_1 N_1 \\ &= D_1^2 + N_1 D_1 \text{ by (6.5.14)} \\ &= (D_1 + N_1) D_1 \\ &= A D_1 \end{aligned}$$

Hence D_1 commutes with A and hence with any polynomial in A. Thus D_1 commutes with D and N. Similarly it can be shown that N_1 commutes with D and N. Thus we have $N - N_1$ is nilpotent. We can also show that the sum of two commuting diagonalizable matrices is diagonalizable and hence $D - D_1$ is diagonalizable. Now we have from (6,2,11) and (6,2,13),

$$\begin{aligned} A - A &= (D - N) - (D_1 - N_1) \\ \Longrightarrow & \\ 0_{n\times n} &= (D - D_1) - (N - N_1) \\ \Longrightarrow & \\ D - D_1 &= N - N_1 \end{aligned}$$

The lhs above is a diagonalizable whereas the rhs is nilpotent, and since they are equal we must have both equal to $0_{n\times n}$, since we have seen that the only matrix which is both diagonalizable and nilpotent is the zero matrix. Thus we have,

$$\begin{aligned} D &= D_1 &(6.5.12) \\ N &= N_1 &(6.5.13) \end{aligned}$$

Thus the decomposition of a matrix $A \in \mathbb{F}^{n\times n}$ as the sum of a diagonalizable and nilpotent matrix is unique if these two matrices have to commute. Thus we have

Theorem 6.5.3 Every matrix $\boldsymbol{A} \in \mathbb{F}^{n \times n}$, can be decomposed as the sum,

$$\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{N} \tag{6.5.14}$$

of a diagonalizable matrix $\boldsymbol{D}$ and a nilpotent matrix $\boldsymbol{N}$. Moreover such a decomposition is unique if

$$\boldsymbol{DN} = \boldsymbol{ND} \tag{6.5.15}$$

We the call $\boldsymbol{D}$ as the "Diagonalizable Part" of $\boldsymbol{A}$ and $\boldsymbol{N}$ as the "Nilpotent Part" of $\boldsymbol{A}$

Remark 6.5.1 The decomposition of A as the sum of a diagonalizable matrix and a nilpotent matrix is NOT unique if we do not insist that the diagonalizable and Nilpotent matrices in the decomposition commute. For example for the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

we have the following decompositions of the matrix:

$$A = D + N$$

where

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{and} N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and also the decomposition

$$A = D_1 + N_1$$

where

$$D_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } N = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$$

We have $DN = ND$ but we do not have $D_1N_1 = N_1D_1$. Thus the second decomposition is a decomposition where the diagonalizable matrix D_1 and the nilpotent matrix N_1 in the decomposition do not commute.

6.6 Canonical Form of a Nilpotent Matrcix

We shall first look at simple examples of nilpotent matrices

Example 6.6.1 The matrix

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

is a nilpotent matrix with order of nilpotency 2.
The matrix

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

is a nilpotent matrix with order of nilpotency 3.
The matrix

$$N = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

is a nilpotent matrix with order of nilpotency 2.
The matrix

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

is a nilpotent matrix with order of nilpotency 2.
In general, the matrix

$$N_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \in \mathbb{C}^{n\times n}$$

is a nilpotent matrix with order of nilpotency n. This is called the canonical $(n \times n)$ (nilpotent) matrix. (N is the $n \times n$ matrix whose $(j, j+1)$th entry is 1 for $j = 1, 2, \cdots, (n-1)$ and all other entries are zero). Thus we have

$$N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Suppose $N \in \mathbb{F}^{n\times n}$ is nilpotent. Then we know that its characteristic polynomial must be $c_{_A}(\lambda) = \lambda^n$. The order of nilpotency $\gamma_{_N}$ satisfies,

$$1 \leq \gamma_{_N} \leq n \tag{6.6.1}$$

When the order of nilpotency is 1 we have $N^1 = 0_{n\times n}$ and hence N is the zero matrix. Thus **the only nilpotent matrix with order of nilpotency 1 is the zero matrix**.

We shall next look at nilpotent matrices whose order of nilpotency is $\boldsymbol{n}$, that is $\gamma_{_N} = n$. For such matrices we have

$$\begin{aligned} N^n &= 0_{n\times n} \text{ and} && (6.6.2)\\ N^r &\neq 0_{n\times n} \text{ for } 1 \leq r \leq n-1 && (6.6.3)\end{aligned}$$

In particular we have $N^{(n-1)} \neq 0_{n\times n}$. Hence there exists a nonzero vector $x \in \mathbb{F}^n$ such that $N^{(n-1)}x \neq \theta_n$. Now we define n nonzero vectors $\{v_j\}_{j=1}^{n}$ as follows:

$$\left.\begin{array}{lcl} v_1 & = & N^{(n-1)}x \\ v_2 & = & N^{(n-2)}x \\ \cdots & \cdots & \cdots \\ v_j & = & N^{((n-j))}x \\ \cdots & \cdots & \cdots \\ v_{(n-1)} & = & Nx \\ v_n & = & x \end{array}\right\} \tag{6.6.4}$$

We now show the following:

<u>**CLAIM**</u>:

$\boldsymbol{v_1, v_2, \cdots, v_n}$ **are linearly independent**

We shall prove this Claim later. But suppose this claim is true then we get that $v_1, v_2, \cdots, v_n$ form a basis for $\mathbb{F}^n$. We then have, from (6.3.4),

$$\left.\begin{array}{lcl} Nv_1 & = & \theta_n \\ Nv_2 & = & v_1 \\ \cdots & \cdots & \cdots \\ Nv_j & = & v_{(j-1)} \\ \cdots & \cdots & \cdots \\ Nv_n & = & v_{(n-1)} \end{array}\right\} \tag{6.6.5}$$

If we now define the matrix P as the matrix whose jth column is v_j, for $j = 1, 2, \cdots, n$, that is,

$$P = [v_1\ v_2\ \cdots\ v_j\ \cdots\ v_n] \tag{6.6.6}$$

then P is invertible, (since the columns are linearly independent), and,

$$\begin{aligned}
NP &= [Nv_1\ Nv_2\ \cdots\ Nv_j\ \cdots\ Nv_n] \\
&= [\theta_n\ v_1\ v_2\ \cdots\ v_{(j-1)}\ \cdots\ v_{(n-1)}] \\
&= P\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \\
\Longrightarrow & \\
P^{-1}NP &= \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \\
&= N_n
\end{aligned}$$

Thus we have,

Theorem 6.6.1 If $N \in \mathcal{F}^{n\times n}$ is nilpotent with order of nilpotency as n, then it is similar to the canonical $n \times n$ nilpotent matrix N_n, that is, there exists an invertible matrix $P \in \mathcal{F}^{n\times n}$ such that

$$P^{-1}NP = N_n \tag{6.6.7}$$

We say that "canonical form of N is N_n and write $N_{can} = N_n$".

We shall complete the proof of the above argument by proving the Claim.
<u>**Proof of Claim**</u>:
We have

$$\begin{aligned}
\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \theta_n \Longrightarrow \\
\alpha_1 N^{(n-1)}x + N^{(n-2)}x + \cdots + \alpha_{(n-1)}Nx + \alpha_n x = \theta_n \Longrightarrow \\
N^{(n-1)}(\alpha_1 N^{(n-1)}x + N^{(n-2)}x + \cdots + \alpha_{(n-1)}Nx + \alpha_n x) = \theta_n \\
\text{(since all terms except the last have factor } N^n \text{ which is } 0_{n\times n}) \Longrightarrow \\
\alpha_n N^{(n-1)}x = \theta_n \Longrightarrow \\
(\text{since } N^{(n-1)}x \neq \theta_n)\ \alpha_n = 0 \Longrightarrow
\end{aligned}$$

$$\alpha_1 N^{(n-1)}x + N^{(n-2)}x + \cdots + \alpha_{(n-1)}Nx = \theta_n \Longrightarrow$$
$$N^{(n-2)}(\alpha_1 N^{(n-1)}x + N^{(n-2)}x + \cdots + \alpha_{(n-1)}Nx) = \theta_n \Longrightarrow$$
$$\text{(as above) } \alpha_{(n-1)} = 0$$

Continuing this process step by step we get all α_j as 0 and hence $v_1, v_2, \cdots, v_n$ are linearly independent, thus proving the Claim.

We next look at nilpotent matrices for which the order of nilpotency is between 1 and n, that is,

$$1 < \gamma_{_N} < n \tag{6.6.8}$$

We then have,

$$N^{\gamma_N} = 0_{n\times n} \text{ and} \tag{6.6.9}$$
$$N^r \neq 0_{n\times n} \text{ for } 1 \leq r \leq \gamma_{_N} - 1 \tag{6.6.10}$$

From this we get

$$\textit{Null Space of } N^{\gamma_N} = \mathcal{F}^n \tag{6.6.11}$$
$$\textit{Null Space of } N^r \neq \mathcal{F}^n \text{ for } 1 \leq r \leq \gamma_{_N} - 1 \tag{6.6.12}$$

We further have,

$$\left.\begin{array}{l} \textit{Null Space of } N^r \text{ is a proper subspace of} \\ \textit{Null Space of } N^{(r+1)} \text{ for } 1 \leq r \leq \gamma_{_N} - 1 \end{array}\right\} \tag{6.6.13}$$

Thus if we define,

$$V_j = \textit{Null Space of } N^j \text{ for } 1 \leq j \leq \gamma_{_N} \tag{6.6.14}$$

then the subspaces are strictly increasing, that is,

$$V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_j \subsetneq V_{(j+1)} \subsetneq \cdots \subsetneq V_{\gamma_{_N}-1} \subsetneq V_{\gamma_{_N}} \tag{6.6.15}$$

Let

$$d_j = \textit{dimension of } V_j \tag{6.6.16}$$

Then by (6.6.15) we get

$$d_1 < d_2 < \cdots < d_{(j-1)} < d_j < d_{(j+1)} < \cdots < d_{(\gamma_N - 1)} < d_{\gamma_N} = n \quad (6.6.17)$$

(Clearly d_1 is the Nullity of the matrix N). We now look at the increase in dimension at each stage. We define

$$\alpha_1 = d_1 \quad (6.6.18)$$
$$\alpha_j = d_j - d_{(j-1)} \text{ for } j = 2, 3, \cdots, \gamma_N \quad (6.6.19)$$
$$(6.6.20)$$

Hence we can write

$$d_1 = \alpha_1 \quad (6.6.21)$$
$$d_2 = \alpha_1 + \alpha_2 \quad (6.6.22)$$
$$d_j = \alpha_1 + \alpha_2 + \cdots + \alpha_{(j-1)} + \alpha_j \text{ for } 2 \leq j \leq \gamma_N \quad (6.6.23)$$

We now state the following Lemma without proof:

Lemma 6.6.1 The $\boldsymbol{\alpha_j}$ are non increasing, that is,

$$\boldsymbol{\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{(j-1)} \geq \alpha_j \geq \alpha_{(j+1)} \geq \cdots \geq \alpha_{\gamma_N - 1} \geq \alpha_{\gamma_N}} \quad (6.6.24)$$

We now define

$$n_j = \alpha_j - \alpha_{(j+1)} \text{ for } 1 \leq j \leq \gamma_N - 1 \quad (6.6.25)$$
$$n_{\gamma_N} = \alpha_{\gamma_N} \quad (6.6.26)$$

By (6.6.24) we get $n_j \geq 0$ for each j. We now define a matrix N_{can} with the following properties:

1. N_{can} is an $n \times n$ matrix with diagonal blocks arranged such that the block sizes are non increasing as we go down the diagonal
2. The number of blocks is equal to α_1 - the nullity of N
3. Each block is a canonical nilpotent matrix
4. The first block is of size $\gamma_N \times \gamma_N$ and hence all other blocks are of size less than or equal to $\gamma_N \times \gamma_N$

5. The number of blocks of size $j \times j$ is equal to n_j, for $j = 1, 2, \cdots, \gamma_{_N}$

The matrix N_{can} is called the "canonical form" of the matrix N. We shall now state the main theorem (without proof) for nilpotent matrices:

Theorem 6.6.2 Every nilpotent matrix $\boldsymbol{N} \in \mathbb{C}^{n \times n}$ is similar to its canonical form, that is,

$$\boldsymbol{N} \in \mathbb{C}^{n\times n} \Longrightarrow \exists \text{ invertible } \boldsymbol{P} \in \mathbb{C}^{n \times n} \text{ such that } \boldsymbol{P}^{-1}\boldsymbol{N}\boldsymbol{P} = \boldsymbol{N_{can}}$$

Example 6.6.2 Consider a nilpotent matrix $N \in \mathbb{C}^{2\times 2}$. We have
<u>**CASE 1**</u>: $\gamma_{_N} = 1$
In this case N has to be $0_{2\times 2}$
<u>**CASE 2**</u>: $\gamma_{_N} = 2$
Since the leading block has to be of size 2×2 we get

$$N_{can} = N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{the canoncal } 2 \times 2 \text{ nilpotent matrix}$$

Example 6.6.3 Consider a nilpotent matrix $N \in \mathbb{C}^{3\times 3}$. We have
<u>**CASE 1**</u>: $\gamma_{_N} = 1$
In this case N has to be $0_{3\times 3}$
<u>**CASE 2**</u>: $\gamma_{_N} = 2$
Since the leading block has to be of size 2×2 we get

$$N_{can} = \left(\begin{array}{c|c} N_2 & 0_{2\times 1} \\ \hline 0_{1\times 2} & 0_{1\times 1} \end{array}\right) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

Example 6.6.4 Consider a nilpotent matrix $N \in \mathbb{C}^{4\times 4}$. We have
<u>**CASE 1**</u>: $\gamma_{_N} = 1$
In this case N has to be $0_{4\times 4}$
<u>**CASE 2**</u>: $\gamma_{_N} = 2$
The leading block is 2×2. The remaining can be either one 2×2 block, (which will hold if the nullity is 2 because then there will be only two blocks), or two 1×1 blocks (which will hold when the nullity is 3 because then there will be three blocks). Hence we have
<u>**CASE 2A**</u>: Nullity of N is 2

$$N_{can} = \left(\begin{array}{c|c} N_2 & 0_{2\times 2} \\ \hline 0_{2\times 2} & N_2 \end{array}\right) = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

CASE 2B: Nullity of N is 3

$$N_{can} = \left(\begin{array}{c|c|c} N_2 & 0_{2\times1} & 0_{2\times1} \\ \hline 0_{1\times2} & 0_{1\times1} & 0_{1\times1} \\ \hline 0_{1\times2} & 0_{1\times1} & 0_{1\times1} \end{array}\right) = \left(\begin{array}{cc|c|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

CASE 3: $\gamma_{_N} = 3$
Since now the leading block has to be 3×3 the remaining has to be a 1×1 block. Hence we get

$$N_{can} = \left(\begin{array}{c|c} N_3 & 0_{3\times1} \\ \hline 0_{1\times3} & 0_{1\times1} \end{array}\right) = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

In this case the nullity has to be 2
CASE 4: $\gamma_{_N} = 4$
Since the leading block itself is 4×4 there will be only one block. We then have

$$N_{can} = N_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Example 6.6.5 Consider a nilpotent matrix $N \in \mathbb{C}^{6\times6}$
CASE 1: $\gamma_{_N} = 1$
N has to be $0_{6\times6}$
CASE 2: $\gamma_{_N} = 2$
All blocks have to be of size smaller than or equal to 2×2 and the leading block must be 2×2. Hence there must be at least 3 boocks. This means that nullity of N is at least 3
CASE 2A: Nullity is 3:

$$N_{can} = \left(\begin{array}{c|c|c} N_2 & & \\ \hline & N_2 & \\ \hline & & N_2 \end{array}\right)$$

CASE 2B: Nullity is 4:

$$N_{can} = \left(\begin{array}{c|c|c|c} N_2 & & & \\ \hline & N_2 & & \\ \hline & & N_1 & \\ \hline & & & N_1 \end{array}\right)$$

CASE 2C: Nullity is 5

$$N_{can} = \left(\begin{array}{c|c|c|c|c} N_2 & & & & \\ \hline & N_1 & & & \\ \hline & & N_1 & & \\ \hline & & & N_1 & \\ \hline & & & & N_1 \end{array}\right)$$

Nullity cannot be 6 when $\gamma_{_N} = 2$(Why?)
CASE 3: $\gamma_{_N} = 3$
The leading block is 3×3. The remaining can be
(i) one 3×3 block (when nullity is 2), or
(ii) one 2×2 block and one 1×1 block (when nullity is 3), or
iii) three 1×1 blocks (when nullity is 4)
Nullity cannot be $1, 5$ or 6 when $\gamma_{_N} = 3$ (Why?)
CASE 3A: Nullity is 2

$$N_{can} = \left(\begin{array}{c|c} N_3 & \\ \hline & N_3 \end{array}\right)$$

CASE 3B: Nullity is 3

$$N_{can} = \left(\begin{array}{c|c|c} N_3 & & \\ \hline & N_2 & \\ \hline & & N_1 \end{array}\right)$$

CASE 3C: Nullity is 4

$$N_{can} = \left(\begin{array}{c|c|c|c} N_3 & & & \\ \hline & N_1 & & \\ \hline & & N_1 & \\ \hline & & & N_1 \end{array}\right)$$

CASE 4: $\gamma_{_N} = 4$
The leading block is 4×4 and the remaining can be,
i) one 2 block (when nullity is 2), or
ii) two 1×1 blocks (when nullity is 3
CASE 4A: Nullity is 2

$$N_{can} = \left(\begin{array}{c|c} N_4 & \\ \hline & N_2 \end{array}\right)$$

CASE 4B: Nullity is 3

$$N_{can} = \left(\begin{array}{c|c|c} N_4 & & \\ \hline & N_1 & \\ \hline & & N_1 \end{array}\right)$$

CASE 5: $\gamma_{_N} = 5$
The leading block is 5×5 and hence the remaining is a 1×1 block.

$$N_{can} = \left(\begin{array}{c|c} N_5 & \\ \hline & N_1 \end{array}\right)$$

In this case the nullity has to be 2.
CASE 6: $\gamma_{_N} = 6$
Clearly there is only one 6×6 block and we have

$$N_{can} = N_6$$

Example 6.6.6 Consider a nilpotent matrix $N \in \mathbb{C}^{12\times 12}$ whoae canonical form is given by

$$N_{can} = \left(\begin{array}{c|c|c|c|c} N_5 & & & & \\ \hline & N_3 & & & \\ \hline & & N_2 & & \\ \hline & & & N_1 & \\ \hline & & & & N_1 \end{array}\right)$$

From the canonical form we can make the following conclusions:

1. Since there are 5 blocks we have Nullity of N must be 5. Thus we have

$$\alpha_1 = 5 \qquad (6.6.27)$$

2. Since the leading block is N_5 it follows that the order of nilpotency of N is 5. Thus we have

$$\gamma_{_N} = 5 \tag{6.6.28}$$

3. We have

$$\begin{aligned}
n_1 = \alpha_1 - \alpha_2 &= 2 \text{ (since there are two } 1\times 1 \text{ blocks)} && (6.6.29)\\
n_2 = \alpha_2 - \alpha_3 &= 1 \text{ (since there is one } 2\times 2 \text{ block)} && (6.6.30)\\
n_3 = \alpha_3 - \alpha_4 &= 1 \text{ (since there is one } 3\times 3 \text{ block)} && (6.6.31)\\
n_4 = \alpha_4 - \alpha_5 &= 0 \text{ (since there are no } 4\times 4 \text{ blocks)} && (6.6.32)\\
n_5 = \alpha_5 &= 1 \text{ (since there is one } 5\times 5 \text{ blocks)} && (6.6.33)
\end{aligned}$$

4. We have (6.6.27) and (6.6.29) $\Longrightarrow$

$$\alpha_2 = 3 \tag{6.6.34}$$

Now (6.6.30) and (6.6.34) *Longrightarrow*

$$\alpha_3 = 2 \tag{6.6.35}$$

Now (6.6.31) and (6.6.35) *Longrightarrow*

$$\alpha_4 = 1 \tag{6.6.36}$$

Now (6.6.32) and (6.6.36) *Longrightarrow*

$$\alpha_5 = 1 \tag{6.6.37}$$

5. We have $V_j = Null\ Space\ of\ N^j$ for $1 \le j \le 5$. From the above values for α_j we get

$$\begin{aligned}
dimension\ of\ V_1 &= \alpha_1 &&= 5\\
dimension\ of\ V_2 &= \alpha_1 + \alpha_2 &&= 5+3=8\\
dimension\ of\ V_3 &= \alpha_1 + \alpha_2 + \alpha_3 &&= 5+3+2=10\\
dimension\ of\ V_4 &= \alpha_1 + \alpha_2 + \alpha_3 + a_4 &&= 5+3+2+1=11\\
dimension\ of\ V_5 &= \alpha_1 + \alpha_2 + \alpha_3 + a_4 + \alpha_5 &&= 5+3+2+1+1=12
\end{aligned}$$

Example 6.6.7 Consider a nilpotent matrix $N \in \mathbb{C}^{17\times 17}$ of order of nilpotency 5 and $V_j = Null\ Space\ of\ N^j$ for $1 \le j \le 5$. Suppose

$$\begin{aligned}
dimension\ of\ V_1 &= 8\\
dimension\ of\ V_2 &= 12\\
dimension\ of\ V_3 &= 14\\
dimension\ of\ V_4 &= 16\\
dimension\ of\ V_5 &= 17
\end{aligned}$$

Then, the characteristic and minimal polynomials must be

$$\begin{aligned}
c_{_N}(\lambda) &= \lambda^{17}\\
m_{_N}(\lambda) &= \lambda^5
\end{aligned}$$

Further,

$$\begin{aligned}
\alpha_1 &= d_1 = 8\\
\alpha_2 &= d_2 - d_1 = 12 - 8 = 4\\
\alpha_3 &= d_3 - d_2 = 14 - 12 = 2\\
\alpha_4 &= d_4 - d_3 = 16 - 14 = 2\\
\alpha_5 &= d_5 - d_4 = 17 - 16 = 1
\end{aligned}$$

Hence we get

$$\begin{aligned}
n_1 &= \alpha_1 - \alpha_2 = 8 - 4 = 4\\
n_2 &= \alpha_2 - \alpha_3 = 4 - 2 = 2\\
n_3 &= \alpha_3 - \alpha_4 = 2 - 2 = 0\\
n_4 &= \alpha_4 - \alpha_5 = 2 - 1 = 1\\
n_5 &= \alpha_5 = 1
\end{aligned}$$

Hence there is

$$\begin{array}{lll}
one & 5\times 5 & block,\\
one & 4\times 4 & block,\\
no & 3\times 3 & blocks,\\
two & 2\times 2 & blocks,\ and\\
four & 1\times 1 & blocks.
\end{array}$$

Hence the canonical form of N is given by

$$N_{can} = \left(\begin{array}{c|c|c|c|c|c|c|c} N_5 & & & & & & & \\ \hline & N_4 & & & & & & \\ \hline & & N_2 & & & & & \\ \hline & & & N_2 & & & & \\ \hline & & & & N_1 & & & \\ \hline & & & & & N_1 & & \\ \hline & & & & & & N_1 & \\ \hline & & & & & & & N_1 \end{array}\right)$$