Chapter 3

Linear Independence and Basis

3.1 Finitely Generated Spaces

We shall now begin investigating the question of obtaining a spanning set of optimal size. We have introduced in the last chapter the notion of a finitely generated subspace. We had,

Definition 3.1.1 Let $\mathcal{V}$ be a vector space over a field $\mathcal{F}$. A subspace $\mathcal{W}$ of $\mathcal{V}$ is said to be finitely generated if there exists a finite spanning set for $\mathcal{W}$, that is, if there exists $S \subset \mathcal{W}$ such that S is finite and $\mathcal{L}[S] = \mathcal{W}$

We illustrate this by some examples.

Example 3.1.1 Consider the vector space $\mathbb{R}^3$. Let $\mathcal{W}$ be the subspace defined as

$$\mathcal{W} = \left\{ x = \begin{pmatrix} \alpha \\ \beta \\ \alpha + \beta \end{pmatrix} : \alpha,\ \beta \in \mathbb{R} \right\}$$

Clearly the set of vectors

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},\ u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

form a finite spanning set for $\mathcal{W}$. Hence $\mathcal{W}$ is a finitely generated subspace.

Clearly the set of vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

form a finite spanning set for $\mathbb{R}^3$ and hence the vector spsce $\mathbb{R}^3$ is itself finitely generated.

Example 3.1.2 Let $\mathcal{V}$ be the vector space, $\mathcal{F}_{\mathbb{R}}[\mathbb{R}]$ of all functions from $\mathbb{R}$ to $\mathbb{R}$. We have

$$\mathcal{F}_{\mathbb{R}}[\mathbb{R}] \;\; = \;\; \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$$

Consider the subspace $\mathcal{W} = \mathbb{R}[x]$ of all polynomials in x with real coefficients. Then $\mathcal{W}$ is not finitely generated. For, suppose it is finitely generated. This would then mean that there exists a finite spanning set

$$S = p_1, p_2, \cdots, p_k$$

for $\mathcal{W}$. Let

$$d = Max.\{degree\ p_j : 1 \leq j \leq k\}$$

Since S is a spanning set for $\mathcal{W}$ we have

$$\begin{aligned} p \in \mathcal{W} \;\; &\Longrightarrow \;\; p = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k, \; (\alpha_j \in \mathbb{R}, \; 1 \leq j \leq k) \\ &\Longrightarrow \;\; degree\ p \leq d \end{aligned}$$

This means that no polynomial in $\mathcal{W}$ can have degree greater than d. Thus is a contradiction, since for example, x^{d+1} is a polynomial of degree greater than d and is in $\mathcal{W}$. Thus $\mathcal{W}$ is not finitely generated. On the other hand, consider the subspace, $\mathcal{W} = \mathbb{R}_N[x]$, of all polynomials in $\mathcal{V}$ of degree less than or equal to N. Then clearly

$$S = \{p_n = x^n\}_{n=0}^{N}$$

is a finite spanning set for $\mathcal{W}$ and hence this subspace is finitely generated.

3.2 Linear Independence

We shall next introduce the notion of a linearly independent set. Consider a finite set of vectors

$$u_1, u_2, \cdots, u_r$$

in a vector space $\mathcal{V}$. Any linear combination of these vectors is of the form

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r$$

where $\alpha_j \in \mathcal{F}$, for $1 \leq j \leq r$. In particular,

$$0u_1 + 0u_2 + \cdots + 0u_r$$

is a linear combination of these vectors and is equal to $\theta_{\mathcal{V}}$. This linear combination is called the trivial linear combination of these vectors. Thus we find that given any finite set of vectors, we can obtain the zero vector $\theta_{\mathcal{V}}$, as a linear combination of these vectors.

Example 3.2.1 Consider the vector space $\mathcal{V} = \mathbb{R}^3$ and the set of vectors,

$$S = u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Then clearly we can write the zero vector $\theta_{\mathcal{V}}$ as the trivial linear combination of these vectors as

$$\theta_{\mathcal{V}} = 0u_1 + 0u_2$$

Further this is the only way we can express $\theta_{\mathcal{V}}$ as a linear combination of u_1, u_2. For, if a linear combination gives $\theta_{\mathcal{V}}$, then we must have,

$$\begin{aligned} \alpha_1 u_1 + \alpha_2 u_2 &= \theta_{\mathcal{V}} \\ &\Longrightarrow \\ \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \alpha_1 + \beta_1 \end{pmatrix} &= 0 \\ &\Longrightarrow \\ \alpha_1, \text{ and } \alpha_2 &= 0 \end{aligned}$$

On the other hand consider the set of vectors,

$$S = u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \; u_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Then we have the trivial linear combination

$$\theta_{\mathcal{V}} \;\; = \;\; 0u_1 + 0u_2 + 0u_3$$

We also have

$$1u_1 + 1u_2 + (-1)u_3 \;\; = \;\; \theta_{\mathcal{V}}$$

In fact, for any $\alpha \in \mathbb{R}$ we have

$$\alpha u_1 + \alpha u_2 + (-\alpha)u_3 \;\; = \;\; \theta_{\mathcal{V}}$$

Thus nontrivial linear combinations of u_1, u_2, u_3 also give rise to the zero vector.

From the above example it follows that given any finite subset S of a vector space $\mathcal{V}$, the following two possibilities arise:

1. EITHER $\theta_{\mathcal{V}}$ can be expressed ONLY as the trivial linear combination of the vectors in S,

2. OR $\theta_{\mathcal{V}}$ can also be expressed as a nontrivial linear combination of the vectors in S

We distinguish these two possibilities with the following definition:

Definition 3.2.1 Let $\mathcal{V}$ be a vector space over a field $\mathcal{F}$. A nonempty finite subset

$$S = u_1, u_2, \cdots, u_r$$

is said to be **linearly independent** if

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}} \implies \alpha_j = 0, \; 1 \leq \jmath \leq r$$

(that is, the only way to express the zero vector as a linear combination of the vectors in S is to express it as the trivial linear combination).
If S is not linearly independent it is said to be **linearly dependent**.

Remark 3.2.1 The set

$$S = u_1, u_2, \cdots, u_r$$

is linearly dependent means that there exist $\alpha_1, \alpha_2, \cdots, \alpha_r \in \mathcal{F}$, at least one of which is not zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}}$$

Example 3.2.2 In Example 3.2.1 above, the set

$$S = u_1, u_2$$

is linearly independent, whereas te set

$$S = u_1, u_2, u_3$$

is linearly dependent.

Example 3.2.3 Consider the vector space $\mathcal{V} = \mathbb{R}[x]$ of all polynomials over $\mathbb{R}$.
a) Consider the set

$$S_1 = p_1, p_2, p_3$$

where

$$p_1 = 1,\ p_2 = x,\ p_3 = x^2$$

We have

$$\begin{aligned} \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 &= \theta_{\mathcal{V}} \\ \Longrightarrow & \\ \alpha_1 + \alpha_2 x + \alpha_3 x^2 &= \theta_{\mathcal{V}} \\ \Longrightarrow & \\ \alpha_1,\ \alpha_2 \text{ and } \alpha_3 &= 0 \end{aligned}$$

Hence the set S_1 is linearly independent.
b) Next we consider the set

$$S_2 = f_1, f_2, f_3$$

where

$$f_1 = 1 + x,\ f_2 = 1 + x^2,\ f_3 = 1 + x + x^2$$

We have

$$\begin{aligned}
\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
\alpha_1(1+x) + \alpha_2(1+x^2) + \alpha_3(1+x+x^2) &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
(\alpha_1+\alpha_2+\alpha_3) + (\alpha_1+\alpha_3)x + (\alpha_2+\alpha_3)x^2 &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
\left.\begin{array}{lll} \alpha_1+\alpha_2+\alpha_3 &=& 0 \\ \alpha_1+\alpha_3 &=& 0 \\ \alpha_2+\alpha_3 &=& 0 \end{array}\right\} \\
&\Longrightarrow \\
\alpha_1,\ \alpha_2 \text{ and } \alpha_3 &= 0
\end{aligned}$$

Hence the set S_2 is linearly independent.
c) Consider the set

$$S_3 = f_1, f_2, f_3$$

where

$$f_1 = 1+x,\ f_2 = x+x^2,\ f_3 = 1+x^2$$

We have

$$\begin{aligned}
\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
\alpha_1(1+x) + \alpha_2(x+x^2) + \alpha_3(1+x^2) &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
(\alpha_1+\alpha_3) + (\alpha_1+\alpha_2)x + (\alpha_2+\alpha_3)x^2 &= \theta_{\mathcal{V}} \\
&\Longrightarrow \\
\left.\begin{array}{lll} \alpha_1+\alpha_3 &=& 0 \\ \alpha_1+\alpha_2 &=& 0 \\ \alpha_2+\alpha_3 &=& 0 \end{array}\right\} \\
&\Longrightarrow \\
\alpha_1,\ \alpha_2 \text{ and } \alpha_3 &= 0
\end{aligned}$$

Hence the set S_3 is linearly independent
d) Consider the set

$$S_4 = f_1, f_2, f_3$$

where

$$f_1 = 1 - x,\ f_2 = 1 + x,\ f_3 = 1$$

This set is linearly dependent since we have

$$1f_1 + 1f_2 + (-2)f_3 = \theta_{\mathcal{V}}$$

a nontrivial linear combination giving rise to $\theta_{\mathcal{V}}$.

3.3 Properties of Linearly Dependent Sets

We shall now look at an useful property of a linearly dependent set. Consider a linearly dependent set

$$S = u_1, u_2, \cdots, u_r$$

(We arrange these vectors in S in some order as above).
1) Let $\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u_{(r+1)}, \cdots, u_k$ be any superset of $\mathcal{S}$. Since $\mathcal{S}$ is linearly dependent there exist $\alpha_1, \alpha_2, \cdots, \alpha_r$, at least one of which is not zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}}$$

$\Longrightarrow$

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r + 0u_{(r+1)} + \cdots + 0u_k = \theta_{\mathcal{V}}$$

Since at least one of the α_j is nonzero we have above a nontrivial linear combination of $\mathcal{S}_1$ vectors giving rise to the zero vector. Hence $\mathcal{S}_1$ is linearly dependent. Hence we can conclude,
Property 1:
Any superset of a linearly dependent set is linearly dependent

2) There exist $\alpha_1, \alpha_2, \cdots, \alpha_r \in \mathcal{F}$, at least one of which is not zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}}$$

Let k be the largest index such that $\alpha_k \neq 0$, that is, $\alpha_k \neq 0$ and $\alpha_j = 0$ if $j > k$. Then we have

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k = \theta_{\mathcal{V}}$$

Since $\alpha_k \neq 0$ we get

$$u_k = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_{(k-1)} u_{(k-1)}$$

where

$$\beta_j = \alpha^{-1}\alpha_j \text{ for } 1 \leq j \leq (k-1)$$

Thus we see that u_k is a linear combination of the preceeding vectors $u_1, u_2, \cdots, u_{(k-1)}$.
Thus we have the folowing property of a linearly dependent set:
Property 2:
If S is a finite linearly dependent set, (in a vector space $\mathcal{V}$), whose vectors are arranged in some order

$$\boldsymbol{S = u_1, u_2, \cdots, u_r}$$

then there exists a vector $\boldsymbol{u_k}$ such that it is a linear combination of the preceeding vectors $\boldsymbol{u_1, u_2, \cdots, u_{(k-1)}}$

3) We shall now use this property to remove the redundancies from a linearly dependent spanning set for a subspace.
Consider a finite set of vectors

$$S = u_1, u_2, \cdots, u_r$$

Without loss of generality let us assume that these vectors are all nonzero.
Case 1: S is linearly independent
In this case S is a linearly independent spanning set for $\mathcal{L}[S]$.
Case 2; S is linearly dependent
In this case, by the above property of linearly dependent sets, we must have a u_k such that it is a linear combination of the preceeding vectors $u_1, u_2, \cdots, u_{(k-1)}$. Let k_1 be the smallest index such that u_{k_1} is a linear combination of the preceeding vectors. (Since the vectors are all nonzero vectors we have $k_1 > 1$). Ths means that,
a) u_{k_1} is a linear combination of $u_1, u_2, \cdots, u_{(k_1-1)}$, and
b) u_j is NOT a linear combination of $u_1, u_2, \cdots, u_{(j-1)}$ for any $j < k_1 - 1$
Now any vector that can be written as a linear combination of $u_1, u_2, \cdots, u_r$ can also be written as a linear combinationof the set of vectors,

$$S_1 = u_1, u_2, \cdots, u_{(k_1-1)}, u_{(k_1+1)}, \cdots, u_r$$

obtained from S by removing the vector u_k. Thus we have

$$\mathcal{L}[S] = \mathcal{L}[\mathcal{S}_1]$$

If S_1 is linearly independent then it is a linearly independent spanning set for $\mathcal{L}[S]$ and $S_1 \subset S$.

If S_1 is linearly dependent, we repeat the above process with S_1 and remove one more vector to get a subset $S_2 \subset S_1 \subset S$ such that

$$\mathcal{L}[S_2] = \mathcal{L}[S_1] = \mathcal{L}[S]$$

If S_2 is linearly independent then it is a linearly independent spanning set for $\mathcal{L}[S]$. If not, we continue this process and in each step we remove one vector, and since S is a finite set, we get, after a finite number of steps, a subset $\tilde{S} \subset S$ such that $\tilde{S}$ is a linearly independent spanning set for $\mathcal{L}[S]$. Thus we have the folowing property of a linearly dependent set:

Property 3:

If S is a finite linearly dependent set in a vector space $\mathcal{V}$, there exists a subset $\tilde{S} \subset S$ such that, $\tilde{S}$ is a linearly independent spanning set for $\mathcal{L}[S]$.

Remark 3.3.1 We call the above process of getting a linearly independent spanning set out of a linearly dependent spanning set, in short, as the "scanning" (from the left) process

3.4 Basis

Consider a finitely generated subspace $\mathcal{W}$ of a vector space $\mathcal{V}$. Since $\mathcal{W}$ is finitely generated there must be a finite spanning set, say

$$S = u_1, u_2, \cdots, u_r$$

Since S is a spanning set for $\mathcal{W}$, we have $\mathcal{L}[S] = \mathcal{W}$. If S is linearly independent then we have a linearly independent spanning set for $\mathcal{W}$. If S is linearly dependent, then by Property 3 of the previous section we can get a linearly independent subset $\tilde{S} \subset S$ such that $\mathcal{L}[\tilde{S}] = \mathcal{L}[S] = \mathcal{W}$. Hence $\tilde{S}$ is a linearly independent spanning set. Thus, in any case, we see that a finitely generated subspace must possess a linearly independent, finite, spanning set. This leads us to the following definition:

Definition 3.4.1 A finite linearly independent spanning set for a finitely generated subspace is called a **BASIS** for the subspace.

Remark 3.4.1 If the vector space $\mathcal{V}$ is itself finitely generated then it will a have finite, linearly independent, spanning set and such a spanning set is called a basis for $\mathcal{V}$

We shall now look at some properties of linearly independent sets and Basis.

1) It is easy to see that,
if $\mathcal{S} = u_1, u_2, \cdots, u_r$ is a linearly independent set then any nonempty subset of $\mathcal{S}$ is also linearly independent

2) Suppose $\mathcal{W}$ is a finitely generated subspace of $\mathcal{V}$ and has basis having k vectors,

$$\mathcal{B} = u_1, u_2, \cdots, u_k$$

Consider any linearly independent set in $\mathcal{W}$ having r vectors, say

$$\mathcal{S} = v_1, v_2, \cdots, v_r$$

We then consider

$$S_1 = v_1, u_1, u_2, \cdots, u_k$$

Since $v_1 \in \mathcal{W}$ and $\mathcal{B}$ is a basis we must have v_1 as a linear combination of the vectors in $\mathcal{B}$. Hence S_1 must be linearly dependent. Hence by Property 3 of linearly dependent spanning sets obtained in the previous section, we must have a subset $\tilde{S}_1 \subset S_1$ such that $\tilde{S}_1$ is a linearly independent spanning set for $\mathcal{W}$, that is, $\tilde{S}_1$ is a basis for $\mathcal{W}$. This is got by the process of removing the redundancy in the linearly dependent spanning set, S_1, using the scanning process described in the previous section. Clearly the process does not remove v_1 from the set S_1. Hence there must be a proper subset $\mathcal{B}'$ of $\mathcal{B}$, (obtained by removing at least one vector from $\mathcal{B}$), such that

$$\mathcal{B}_1 = v_1, \mathcal{B}'$$

is a basis for $\mathcal{W}$. Now we let

$$S_2 = v_2, v_1, \mathcal{B}'$$

Since this is a linearly dependent set we can repeat the above argument to S_2 to obtain a proper subset $\mathcal{B}'_1$ of $\mathcal{B}_1$, (and hence a proper subset of $\mathcal{B}$), such that

$$\mathcal{B}_2 = v_2, v_1, \mathcal{B}'_1$$

is a basis for $\mathcal{W}$. We continue this process. There arise two possibilities.
Possibility 1: All the vectors from $\mathcal{S}$ have been exhaustsed.
In this case we get a basis

$$\mathcal{B}_r = v_r, v_{(r-1)}, \cdots, v_1, \mathcal{B}'_{(r-1)}$$

where $\mathcal{B}'_{(r-1)}$ is a proper subset of $\mathcal{B}$. Since in each step, we remove at least one of the vectors in $\mathcal{B}$ and append one vector from $\mathcal{S}$, we must have at least r vectors in $\mathcal{B}$, that is,

$$r \leq k \tag{3.4.1}$$

Possibility 2: $r > k$. All the $\mathcal{B}$ vectors are removed first
In this case suppose at the jth stage (where, clearly $j \leq r$), we have a basis

$$\mathcal{B}_j = v_j, v_{(j-1)}, \cdots, v_1$$

for $\mathcal{W}$, and $j \leq r$. Hence we have

$$\mathcal{W} = \mathcal{L}[\mathcal{B}_j] \text{ and } v_{(j+1)} \in \mathcal{W}$$

Hence $v_{(j+1)}$ must be a linear combination of $v_1, v_2, \cdots, v_j$, which is a contradiction, since S is linearly independent. Thus this possibility cannot arise. Hence we have (3.4.1). Thus we have
Property 2:
If a subspace $\mathcal{W}$ has a basis consisting of k vectors then any linearly independent set in $\mathcal{W}$ can have at most k vectors
or equivalently we can say that,
If a finitely generated subspace $\mathcal{W}$ has a basis consisting of k vectors, then any subset of $\mathcal{W}$ having more than k vectors must be linearly dependent

3) Suppose now $\mathcal{W}$ is a finitely generated subspace and

$$\mathcal{B} = u_1, u_2, \cdots, u_d$$

$$\mathcal{B}_1 = u_1, u_2, \cdots, u_k$$

are any two bases for $\mathcal{W}$. Then we have since $\mathcal{B}$ is a basis and $\mathcal{B}_1$ being a basis is linearly independent, we must have by the Property 2 above that $\mathcal{B}'$ must have at most d vectors, that is

$$k \leq d \tag{3.4.2}$$

Similarly, since $\mathcal{B}'$ is a basis and $\mathcal{B}$ is a linearly independent set, we must have by above property that $\mathcal{B}$ has at most k vectors, that is,

$$d \leq k \tag{3.4.3}$$

Combining (3.4.3) and (3.4.4) we get

$$k = d \tag{3.4.4}$$

Thus we have,
Property 3:
If $\mathcal{W}$ is a finitely generated subspace then all bases for $\mathcal{W}$ must have the same number of vectors

4) Thus we see that with every finitely generated subspace $\mathcal{W}$ there is a finite number associated, namely, the number of vectors in a basis for $\mathcal{W}$. This leads us to the following definition:

Definition 3.4.2 If $\mathcal{W}$ is a finitely generated subspace then the number of vectors in a basis is called the "**dimension**" of $\mathcal{W}$.

From now on, we shall therefore refer to a finitely generated subspace as a finite dimensional subspace. If the vector space is itself finite dimensional we refer to it as a finite dimensional vector space.

5) Consider a finite set $\mathcal{S} = u_1, u_2, \cdots, u_r$ of linearly independent vectors in $\mathcal{V}$. Let $u \in \mathcal{V}$ be such that $u \notin \mathcal{L}[\mathcal{S}]$. We shall now show that $\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u$, the set obtained by appending u to $\mathcal{S}$ is linearly independent. We have
$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r + \alpha u = \theta_{\mathcal{V}} \Longrightarrow$
$\alpha = 0$ for otherwise u will be a linear combination of $\mathcal{S}$ vectors and hence will be in $\mathcal{L}[\mathcal{S}]$ which is not so
$\Longrightarrow \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}} \Longrightarrow$
$\Longrightarrow \alpha_j = 0,\ 1 \leq j \leq r$ since $\mathcal{S}$ is linearly independent
$\Longrightarrow \mathcal{S}_1$ is linearly independent. Thus we have,
Property 4:
$\mathcal{S} = u_1, u_2, \cdots, u_r$ is linearly independent in $\mathcal{V}$ and $u \notin \mathcal{L}[\mathcal{S}] \Longrightarrow$ $\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u$ is also linearly independent.

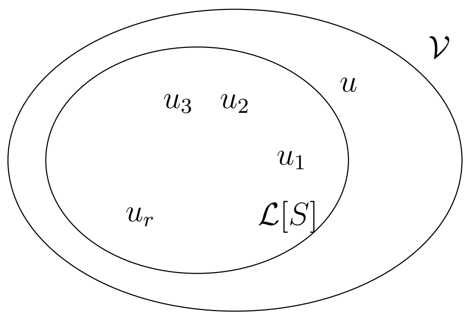
6) Suppose now $\mathcal{W}$ is a finite dimensional subspace and dimension of $\mathcal{W}$ is d. Then any basis for $\mathcal{W}$ has exactly d vectors. Suppose $\mathcal{S}$ is any linearly independent set having d vectors. Then $\mathcal{S}$ must be a basis for $\mathcal{W}$. For,
$\mathcal{S}$ NOT a basis for $\mathcal{W} \Longrightarrow \mathcal{S}$ is not a spanning set for $\mathcal{W}$
$\Longrightarrow \mathcal{S}$ is properly contained in $\mathcal{W}$
$\Longrightarrow$ There exists a $w \in \mathcal{W}$ such that $w \notin \mathcal{L}[\mathcal{S}]$
$\Longrightarrow \mathcal{S} \cup \{w\}$ is linearly independent and has $d+1$ vectors

a contradiction because any $d+1$ vectors in $\mathcal{W}$ must be linearly dependent, (by Property 4 above). Thus we have

Property 5:

$\mathcal{W}$ has dimension $d \Longrightarrow$ Every linearly independent set having d vectors must be a basis for $\mathcal{W}$

7) Let $\mathcal{V}$ be a vector space and let $\mathcal{S} = u_1, u_2, \cdots, u_r$ be linearly independent vectors in $\mathcal{V}$. Let $u \in \mathcal{V}$ be such that $u \notin \mathcal{S}$



Now consider the set $\mathcal{S}_1$ obtained by appending u to $\mathcal{S}$, that is,

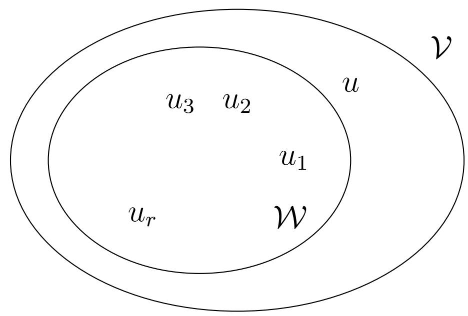
$$\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u$$

We shall show that $\mathcal{S}_1$ is linearly independent. We have
$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r + \alpha u = \theta_{\mathcal{V}} \Longrightarrow$
$\alpha = 0$ (for otherwise u will become a linear combination of $u_1, u_2, \cdots, u_r$ and hence will be in $\mathcal{L}[\mathcal{S}]$ which will be a contradiction since we have chosen $u \notin \mathcal{L}[\mathcal{S}]$) $\Longrightarrow$
$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_r u_r = \theta_{\mathcal{V}} \Longrightarrow$
$\alpha_j = 0$ for $1 \leq j \leq r$, (since $u_1, u_2, \cdots, u_r$ are linearly independent) $\Longrightarrow$
$\mathcal{S}_1$ is linearly independent
Thus we have,

Property 6:

If $\mathcal{S} = u_1, u_2, \cdots, u_r$ is linearly independent in $\mathcal{V}$ and $u \in \mathcal{V}$ is such that $u \notin \mathcal{L}[\mathcal{S}]$ then $\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u$ is also linearly independent

8) Let $\mathcal{W}$ be a subspace and let $\mathcal{S} = u_1, u_2, \cdots, u_r$ be linearly independent vectors in $\mathcal{W}$. Let $u \in \mathcal{V}$ be such that $u \notin \mathcal{W}$



Then we have, $\mathcal{L}[\mathcal{S}] \subseteq \mathcal{W}$ and hence $u \notin \mathcal{L}[\mathcal{S}]$ (since we are given that $u \notin \mathcal{W}$) $\Longrightarrow$ by Property 6 above, the set $\mathcal{S}_1 = u_1, u_2, \cdots, u_r, u$ is linearly independent. Thus we have

Property 7

$u_1, u_2, \cdots, u_r \in \mathcal{W}$ linearly independent and $u \notin \mathcal{W} \Longrightarrow u_1, u_2, \cdots, u_r, u$ is linearly independent in $\mathcal{V}$

This leads us to the following important fact:

9) Let $\mathcal{W}$ be a finite dimensional subspace. Let dimension of $\mathcal{W}$ be d. Let $\mathcal{S} = u_1, u_2, \cdots, u_r$ be any linearly independent set in $\mathcal{W}$. Since any $d+1$ vectors in $\mathcal{W}$ must be linearly dependent we must have $r \leq d$.

Case 1: $r = d$

Then $\mathcal{S}$ is a basis for $\mathcal{W}$ since any d vectors in a d dimensional subspace must be a basis for that subspace.

Case 2: $r < d$

Let $d - r = k$.

Since $\mathcal{S}$ is a subspace of $\mathcal{W}$ and is properly contained in $\mathcal{W}$, there exists a vector $w_1 \in \mathcal{W}$ such that $w_1 \notin \mathcal{L}[\mathcal{S}]$. By the above property we have

$$\mathcal{S}_1 = u_1, u_2, \cdots, u_r, w_1$$

is a linearly independent set and is in $\mathcal{W}$. Then look at $\mathcal{L}[\mathcal{S}_1]$ and get a $w_2 \in \mathcal{W}$ such that $w_2 \notin \mathcal{L}[\mathcal{S}_1]$. Then we have

$$\mathcal{S}_2 = u_1, u_2, \cdots, u_r, w_1, w_2$$

linearly independent in $\mathcal{W}$. Continuing this process, at the kth step we get

$$\mathcal{S}_k = u_1, u_2, \cdots, u_r, w_1, w_2 \cdots, w_k$$

linearly independent in $\mathcal{W}$. since $k = d - r$, this gives us a linearly independent set having d vectors in the d dimensional subspace. Hence $\mathcal{S}_k$ is a basis for $\mathcal{W}$ by Property 5 above, and this basis is obtained by "extending" the given linearly independent set by appending $d-r$ more vectors suitably from $\mathcal{W}$. Thus we have

Property 8:

A linearly independent in a finite dimensional subspace $\mathcal{W}$ is either a Basis for $\mathcal{W}$ or can be extended to be a basis for $\mathcal{W}$

In particular,

if $\mathcal{V}$ is a finite dimensional vector space having dimension n, then any linearly independent set $u_1, u_2, \cdots, u_r$ $(r < n)$ in $\mathcal{V}$ can be extended to a basis by appending suitable $n-r$ vectors

3.5 Rank Nullity Theorem

Let $\mathbb{F}$ be any field and consider the vector space $\mathbb{F}^k$. Let

$$e_j = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(k\times 1)}$$

be the vector in $\mathbb{F}^k$ with 1 in jth position and 0 elsewhere. The set of vectors,

$$\mathcal{B} = e_1, e_2, \cdots, e_k$$

form a basis for $\mathbb{F}^k$ as it is easy to see that they form a linearly independent set in $\mathbb{F}^k$ and form a spanning set for $\mathbb{F}^k$. Hence we have

$$dim(\mathbb{F}^k) \;=\; k \tag{3.5.1}$$

Now consider a matrix $A \in \mathbb{F}^{m\times n}$. For such $m \times n$ matrix over $\mathbb{F}$ we have introduced the following four fundamental subspaces:

1. Null Space of A denoted by $\mathcal{N}_{_A}$
2. Range Space of A denoted by $\mathcal{R}_{_A}$

3. Null Space of A^T denoted by $\mathcal{N}_{A^T}$

4. Range Space of A denoted by $\mathcal{R}_{A^T}$

We have seen that,

1. $\mathcal{R}_A = Col(A) = Row(A^T)$

2. $\mathcal{R}_{A^T} = Col(A^T) = Row(A)$

The subspaces $\mathcal{R}_{A^T}$ and $\mathcal{N}_A$ are subspaces of $\mathbb{F}^n$ and hence are finite dimensional. We define

Definition 3.5.1

$$Nullity\ of\ A \stackrel{def}{=} dimension\ of\ \mathcal{N}_A \text{ and is denoted by } \nu_A \quad (3.5.2)$$

$$Rank\ of\ A \stackrel{def}{=} dimension\ of\ \mathcal{R}_A \text{ and is denoted by } \rho_A \quad (3.5.3)$$

Similarly we have

$$Nullity\ of\ A^T \stackrel{def}{=} dimension\ of\ \mathcal{N}_{A^T} \text{ and is denoted by } \nu_{A^T} \quad (3.5.4)$$

$$Rank\ of\ A^T \stackrel{def}{=} dimension\ of\ \mathcal{R}_{A^T} \text{ and is denoted by } \rho_{A^T} \quad (3.5.5)$$

We shall now see an important relation between these numbers:
Let us consider the zero matrix $0 \in \mathbb{F}^{m\times n}$. Then clearly we have

$$\begin{aligned} \mathcal{N}_0 &= \mathbb{F}^n \text{ and} \\ \mathcal{R}_0 &= \{\theta_n\} \end{aligned}$$

Hence we have $\nu_0 = n$ and $ranko = 0$. Thus we get

$$\rho_0 + \nu_0 = n, \text{ the number of columns in } 0$$

Next let $A \in \mathbb{F}^{m\times n}$ be a nonzero matrix. Any basis for $\mathcal{N}_A$ will have ν_A vectors. Let

$$\mathcal{B}_{\mathcal{N}_A} = \varphi_1, \varphi_2, \cdots, \varphi_{\nu_A}$$

be a basis for $\mathcal{N}_A$, (where $\nu_A < n$). By the Property 8 obtained in Section 3.4, we can extend this to a basis

$$\mathcal{B} = \varphi_1, \varphi_2, \cdots, \varphi_{\nu_A}, v_1, v_2, \cdots, v_{(n-\nu_A)}$$

for $\mathbb{F}^n$, by appending suitable vectors $v_1, v_2, \cdots, v_{(n-\nu_A)}$. Now any vector $b \in \mathcal{R}_A$ is of the form Ax for some $x \in \mathbb{F}^n$, and any $x \in \mathbb{F}^n$ is a linear combination of the vectors in the basis $\mathcal{B}$. Therefore we have,

$$\begin{aligned}
b \in \mathcal{R}_A &\implies \exists x \in \mathbb{F}^n \ni b = Ax \\
&\implies b = A\left(\sum_{j=1}^{\nu} {}_A\alpha_j\varphi_j + \sum_{k=1}^{(n-\nu_A)} \beta_k v_k\right) \\
&\quad\;\; (\text{where } \alpha_j, \beta_k \in \mathbb{F},\ 1 \leq j \leq \nu_A,\ 1 \leq k \leq n - \nu_A) \\
&\implies b = \sum_{j=1}^{\nu_A} \alpha_j(A\varphi_j) + \sum_{k=1}^{(n-\nu_A)} \beta_k(Av_k) \ (\text{since } A\varphi_j = \theta_n) \\
&\implies b = \sum_{k=1}^{(n-\nu_A)} \beta_k u_k \text{ where } u_k = Av_k \in \mathcal{R}_A
\end{aligned}$$

Thus we see that the set of vectors,

$$S = u_1, u_2, \cdots, u_k$$

is in $\mathcal{R}_A$ and every vector in $\mathcal{R}_A$ is a linear combination of these vectors. Hence S is a spanning set for $\mathcal{R}_A$. If we show that S is also linearly independent then it will become a linearly independent spanning set and hence a basis for $\mathcal{R}_A$. We now proceed to prove that S is linearly independent. We have,

$$\begin{aligned}
\sum_{k=1}^{(n-\nu_A)} \beta_k u_k = \theta_m &\implies \sum_{k=1}^{(n-\nu_A)} \beta_k(Av_k) = \theta_m \ (\text{since } u_k = Av_k) \\
&\implies A\left(\sum_{k=1}^{(n-\nu_A)} \beta_k v_k\right) = \theta_m \\
&\implies \sum_{k=1}^{(n-\nu_A)} \beta_k v_k \in \mathcal{N}_A \\
&\implies \sum_{k=1}^{(n-\nu_A)} \beta_k v_k = \sum_{j=1}^{\nu_A} \gamma_j\varphi_j, \text{ since } \mathcal{B}_{\mathcal{N}A} \text{ is a basis for } \mathcal{N}_A \\
&\implies \sum_{j=1}^{\nu_A} \gamma_j\varphi_j + \sum_{k=1}^{(n-\nu_A)} (-\beta_k)v_k = \theta_n
\end{aligned}$$

$$\Longrightarrow \quad \gamma_j = 0,\ \beta_k = 0,\ 1 \leq j \leq \nu_{_A},\ 1 \leq k \leq n - \nu_{_A}$$

(since $\mathcal{B}$ is a basis and hence linearly independent)

$$\Longrightarrow \quad S \text{ is linearly independent}$$

Thus S is a linearly independent spanning set for $\mathcal{R}_{_A}$ and hence basis for $\mathcal{R}_{_A}$. Since there are $n - \nu_{_A}$ vectos in S we get

$$Dimension\ \mathcal{R}_{_A} = n - \nu_{_A}$$

But the dimension of $\mathcal{R}_{_A}$ is $\rho_{_A}$, the rank of A. Thus we get

$$\rho_{_A} + \nu_{_A} = \text{number of columns of } A \tag{3.5.6}$$

Similarly we get

$$\rho_{_{A^T}} + \nu_{_{A^T}} = \text{number of columns of } A^T \tag{3.5.7}$$

Thus we have,

Theorem 3.5.1 **<u>Rank Nullity Theorem</u>:**
For any matrix $\boldsymbol{A} \in \mathbb{F}^{\boldsymbol{m \times n}}$, we have

$$\boldsymbol{Rank\ of\ A + Nullity\ of\ A = Number\ of\ Columns\ in\ A}$$

Thus we have for any $A \in \mathbb{F}^{m \times n}$,

$$\nu_{_A} + \rho_{_A} = n \tag{3.5.8}$$

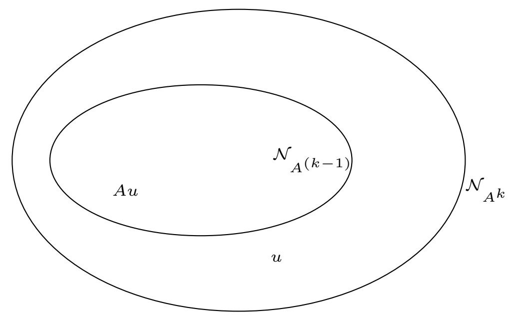
$$\rho_{_{A^T}} + \nu_{_{A^T}} = m \tag{3.5.9}$$

3.6 Some Properties of the Null Space of a Square Matrix and Its Powers

Let $\mathbb{F}$ be any field and $A \in \mathbb{F}^{n \times n}$. Clearly, for any positive integers k, ℓ we have that every vector in $\mathcal{N}_{_{A^k}}$ is also in $\mathcal{N}_{_{A^\ell}}$ whenever $k < \ell$. On the other hand all vectors in $\mathcal{N}_{_{A^\ell}}$ may not be in $\mathcal{N}_{_{A^k}}$. We now observe some consequences of such situations:

1) Let $A \in \mathbb{F}^{n \times n}$. Let k be any positive integer ≥ 2. Suppose $u \in \mathcal{N}_{_{A^k}}$ and $u \notin \mathcal{N}_{_{A^{(k-1)}}}$. Then the vector $Au \in \mathcal{N}_{_{A^{(k-1)}}}$ since

$$A^{(k-1)}(Au) = A^k u = \theta_n \text{ since } u \in \mathcal{N}_{_{A^k}}$$



Since $u \notin \mathcal{N}_{A^{(k-1)}}$ it follows that $u \notin \mathcal{N}_A$ and hence $Au \neq \theta_n$. Thus Au, being a nonzero vector in $\mathcal{N}_{A^{(k-1)}}$, forms a linearly independent set in the subspace $\mathcal{N}_{A^{(k-1)}}$. Since u does not belong to this subspace, it follows that Au, u form a linearly independent set in $\mathcal{N}_{A^k}$, (by Property 7 in Section 3.4). Thus we have

Property 1

$\boldsymbol{k \geq 2,\ u \in \mathcal{N}_{A^k}}$ **and** $\boldsymbol{u \notin \mathcal{N}_{A^{(k-1)}}}$

$\Longrightarrow$

a) $\boldsymbol{Au}$ **is a linearly independent set in** $\boldsymbol{\mathcal{N}_{A^{(k-1)}}}$ **and**

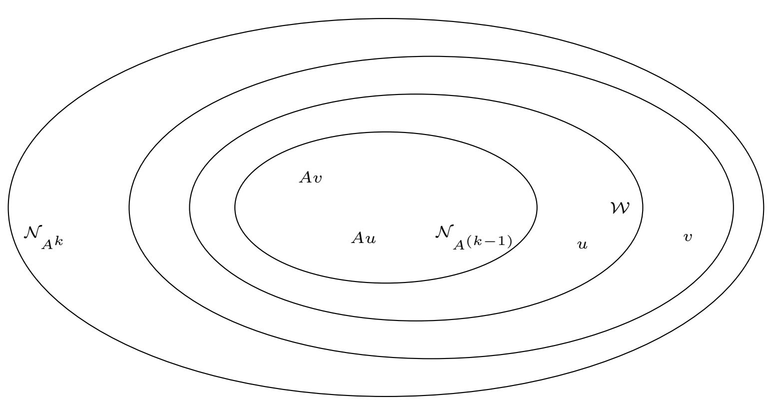
b) $\boldsymbol{Au, u}$ **is a linearly independent set in** $\boldsymbol{\mathcal{N}_{A^k}}$

We shall now look at a simple generalization of this.

2) Let k be any positive integer $k \geq 2$. Let u, v be any two vectors in $\mathcal{N}_{A^k}$ which are not in $\mathcal{N}_{A^{(k-1)}}$. Let $\mathcal{W} = \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u]$. Suppose

$$v \ \notin \ \mathcal{W} \tag{3.6.1}$$

Clearly the vectors Au and Av are in $\mathcal{N}_{A^{(k-1)}}$ since $A^{(k-1)}(Au) = A^k u = \theta_n$ and $A^{(k-1)}(Av) = A^k v = \theta_n$ as $u, v \in \mathcal{N}_{A^k}$ and as above they are nonzero vectors.



Further u, v are nonzero vectors since they are outside $\mathcal{N}_{_A}$ and θ_n is in $\mathcal{N}_{_A}$. We shall first see that the vectors Au, Av are linearly independent vectors in $\mathcal{N}_{_{A^{(k-1)}}}$. We have

$$\begin{array}{lcl}
\alpha Au + \beta Av & = & \theta_n \\
\Longrightarrow & & \\
A(\alpha u + \beta v) & = & \theta_n \\
\Longrightarrow & & \\
\alpha u + \beta v & \in & \mathcal{N}_{_A} \\
\Longrightarrow & & \\
\alpha u + \beta v & \in & \mathcal{N}_{_{A^{(k-1)}}} \text{ (since } \mathcal{N}_{_A} \subseteq \mathcal{N}_{_{A^{(k-1)}}}) \\
\Longrightarrow & & \\
\alpha u + \beta v & = & x \text{ where } x \in \mathcal{N}_{_{A^{(k-1)}}} \\
\Longrightarrow & & \\
\beta v & = & x + (-\alpha)u \\
\Longrightarrow & & \\
\beta v & \in & \mathcal{N}_{_{A^{(k-1)}}} + \mathcal{L}[u] = \mathcal{W} \\
\Longrightarrow & & \\
\beta & = & 0 \text{ (since otherwise } v \text{ will be in } \mathcal{W} \text{ violating (.1))} \\
\Longrightarrow & & \\
\alpha u & = & \theta_n \\
\Longrightarrow & & \\
\alpha & = & 0 \text{ since } u \text{ is a nonzero vector}
\end{array}$$

Thus we have

$$\alpha Au + \beta Av = \theta_n \implies \alpha = \beta = 0$$

Hence Au, Av are linearly independent vectors.
Now consider the set of vectors,

$$\mathcal{S} = Au, Av, u$$

This is a linearly independent set of vectors since Au, Av are linearly independent vectors in the subspace $\mathcal{N}_{_{A^{(k-1)}}}$ and u is outside this subspace. Consequently,

$$\mathcal{S} = Au, Av, u, v$$

is a linearly independent set of vectors since Au, Av, u are linearly independent vectors in the subspace $\mathcal{W}$ and v is outside this subspace. Thus we have

Property 2:
Let k be any integer ≥ 2. Then
$\boldsymbol{u, v \in \mathcal{N}_{A^k}}$ **and** $\boldsymbol{\notin \mathcal{N}_{A^{(k-1)}}}$ **and** $\boldsymbol{v \notin \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u]}$
$\Longrightarrow$
a) The set of vectors $\boldsymbol{\mathcal{S}_1 = Au, Av}$ is a linearly independent set in $\boldsymbol{\mathcal{N}_{A^{(k-1)}}}$
b) The set of vectors $\boldsymbol{\mathcal{S} = Au, Av, u, v}$ is a linearly independent set in $\boldsymbol{\mathcal{N}_{A^k}}$ We can easily follow the same arguments to prove the following property:

Property 3
Let k be any integer ≥ 2. Then
$\boldsymbol{u_1, u_2, \cdots, u_r \in \mathcal{N}_{A^k}}$ **and** $\boldsymbol{\notin \mathcal{N}_{A^{(k-1)}}}$ **are such that**
$\boldsymbol{u_1 \notin \mathcal{N}_{A^{(k-1)}}}$ **and**
$\boldsymbol{u_j \notin \mathcal{W}_j = \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(u-1)}]}$ **for $\boldsymbol{2 \leq j \leq r}$**
$\Longrightarrow$
a) The set of vectors $\boldsymbol{\mathcal{S}_1 = Au_1, Au_2, \cdots, Au_r}$ is a linearly independent set in $\boldsymbol{\mathcal{N}_{A^{(k-1)}}}$
b) The set of vectors $\boldsymbol{\mathcal{S} = Au_1, Au_2, \cdots, Au_r, u_1, u_2, \cdots, u_r}$ is a linearly independent set in $\boldsymbol{\mathcal{N}_{A^k}}$
We shall now see further generalisations of this property.

3) Let k be any positive integer ≥ 3. For any $A \in \mathbb{F}^{n \times n}$ we have

$$\mathcal{N}_{A^{(k-2)}} \subseteq \mathcal{N}_{A^{(k-1)}} \subseteq \mathcal{N}_{A^k} \tag{3.6.2}$$

Suppose the matrix $A \in \mathbb{F}^{n \times n}$ is such that

$$\mathcal{N}_{A^{(k-2)}} \neq \mathcal{N}_{A^{(k-1)}} \neq \mathcal{N}_{A^k} \tag{3.6.3}$$

Let $u_1, u_2, \cdots, u_r$ be vectors in $\mathcal{N}_{A^k}$ such that

$$u_1 \notin \mathcal{N}_{A^{(k-1)}} \text{ and} \tag{3.6.4}$$
$$u_j \notin \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(j-1)}] \tag{3.6.5}$$

We define

$$\mathcal{W}_j \stackrel{def}{=} \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(j-1)}]$$

Thus we have

$$u_j \notin \mathcal{W}_j \text{ for } 2 \leq j \leq r$$

Clearly $u_1, u_2, \cdots, u_r$ are all nonzero vectors.
First we consider the vectors

$$\mathcal{S}_1 = A^2u_1, A^2u_2, \cdots, A^2u_r$$

These vectors are all in $\mathcal{N}_{A^{(k-2)}}$ since $A^{(k-2)}(A^2u_j) = A^ku_j = \theta_n$ for $1 \leq j \leq r$.
We shall now show that this is a linearly independent set of vectors in $\mathcal{N}_{A^{(k-2)}}$.
We have,

$$\begin{aligned}
\alpha_1A^2u_1 + \alpha_2A^2u_2 + \cdots + \alpha_rA^2u_r &= \theta_n \\
&\Longrightarrow \\
A^2\left[\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_ru_r\right] &= \theta_n \\
&\Longrightarrow \\
\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_ru_r &\in \mathcal{N}_{A^2} \\
&\Longrightarrow \\
\alpha_1u_1 + \alpha_2u_2 + \cdots + \alpha_ru_r &\in \mathcal{N}_{A^{(k-1)}} \\
& \left(\text{since } k \text{ being} \geq 3 \text{ we have } \mathcal{N}_{A^2} \subseteq \mathcal{N}_{A^{(k-1)}}\right) \\
&\Longrightarrow \\
\alpha_ru_r &= \text{an } \mathcal{N}_{A^{(k-1)}} \text{ vector} + \text{ a vector in } \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}] \\
&\Longrightarrow \\
\alpha_ru_r &\in \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}] \\
&\Longrightarrow \\
\alpha_r &= 0 \text{ (since otherwise } u_r \in \mathcal{W}_r \text{ - a contradiction)}
\end{aligned}$$

Hence we get

$$\alpha_1A^2u_1 + \alpha_2A^2u_2 + \cdots + \alpha_{(r-1)}A^2u_{(r-1)} = \theta_n$$

Repeating the above argument to this sum now we get $\alpha_{(r-1)} = 0$. Continuing this process we get

$$\alpha_j = 0 \text{ for all } j = 2, 3, \cdots, r$$

Hence we get $\alpha_1u_1 \in \mathcal{N}_{A^{(k-1)}}$ which gives $\alpha_1 = 0$ (since otherwise $u_1 \in \mathcal{N}_{A^{(k-1)}}$ - a contradiction). Thus we see that

$$\alpha_1A^2u_1 + \alpha_2A^2u_2 + \cdots + \alpha_rA^2u_r = \theta_n \implies \alpha_j = 0 \text{ for } 1 \leq j \leq r$$

Thus we get

$$\left.\begin{array}{l}\text{The set } \mathcal{S}_1 = A^2u_1, A^2u_2, \cdots, A^2u_r \\ \text{is a linearly independent set in } \mathcal{N}_{A^{(k-2)}}\end{array}\right\} \qquad (3.6.6)$$

Next we consider the set of vectors

$$\mathcal{S}_2 = Au_1, Au_2, \cdots, Au_r$$

These vectors are clearly in $\mathcal{N}_{A^{(k-1)}}$ since $A^{(k-1)}(Au_j) = A^k u_j = \theta_n$ for $1 \leq j \leq r$. We shall now show that these vectors are linearly independent. We have

$$\begin{array}{rcl}\displaystyle\sum_{j=1}^{r} \alpha_j A u_j &=& \theta_n \\ \Longrightarrow && \\ A\left(\displaystyle\sum_{j=1}^{r} \alpha_j u_j\right) &=& \theta_n \\ \Longrightarrow && \\ \displaystyle\sum_{j=1}^{r} \alpha_j u_j &\in& \mathcal{N}_A \\ \Longrightarrow && \\ \displaystyle\sum_{j=1}^{r} \alpha_j u_j &\in& \mathcal{N}_{A^{(k-1)}} \text{ (since } \mathcal{N}_A \subset \mathcal{N}_{A^{(k-1)}}) \\ \Longrightarrow && \\ \alpha_r u_r &=& (\text{a } \mathcal{N}_{A^{(k-1)}} \text{ vector}) + (\text{a vector in } \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}]) \\ \Longrightarrow && \\ \alpha_r &=& 0 \text{ (since otherwise } u_r \in \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}] \text{)}\end{array}$$

Hence we get

$$\sum_{j=1}^{(r-1)} \alpha_j A u_j \in \mathcal{N}_{A^{(k-1)}}$$

Applying the above argument repeatedly we get

$$\alpha_j = 0 \text{ for } 2 \leq j \leq r$$

Hence we get

$$\begin{aligned} \alpha_1 u_1 &\in \mathcal{N}_{A^{(k-1)}} \\ \text{This} &\implies \\ \alpha_1 &= 0 \text{ (since by our choice } u_1 \notin \mathcal{N}_{A^{(k-1)}} \text{)} \end{aligned}$$

Thus we have

$$\sum_{j=1}^{r} \alpha_j A u_j = \theta_n \implies \alpha_j = 0 \text{ for } 1 \le j \le r$$

Hence

$$\left.\begin{array}{c} \text{The set } \mathcal{S}_2 = Au_1., Au_2, \cdots, Au_r \\ \text{is a linearly independent set in } \mathcal{N}_{A^{(k-1)}} \end{array}\right\} \tag{3.6.7}$$

We next consider the set obtained by taking all the vectors in $\mathcal{S}_1$ and in $\mathcal{S}_2$ to get

$$\mathcal{S}_3 = A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r. \tag{3.6.8}$$

Those are all vectors in $\mathcal{N}_{A^{(k-1)}}$. We shall this set is also a linearly independent set. We have

$$\begin{aligned} \sum_{j=1}^{r} \alpha_j A^2 u_j + \sum_{j=1}^{r} \beta_j A u_j &= \theta_n \\ &\implies \\ A\left[\sum_{j=1}^{r} \alpha_j A u_j + \sum_{j=1}^{r} \beta_j u_j\right] &= \theta_n \\ &\implies \\ \sum_{j=1}^{r} \alpha_j A u_j + \sum_{j=1}^{r} \beta_j u_j &\in \mathcal{N}_A \end{aligned}$$

Let

$$x = \sum_{j=1}^{r} \alpha_j A u_j + \sum_{j=1}^{r} \beta_j u_j$$

We then have from above that $x \in \mathcal{N}_A$ and hence $x \in \mathcal{N}_{A^{(k-1)}}$. This gives

$$\beta_r u_r = x + \sum_{j=1}^{r} (-\alpha_j) A u_j + \sum_{j=1}^{r-1} \beta_j u_j$$

$$
\begin{aligned}
&= y + \sum_{j=1}^{r-1} \beta_j u_j \text{ where} \\
y &= x + \sum_{j=1}^{r} (-\alpha_j) A u_j \in \mathcal{N}_{A^{(k-1)}}
\end{aligned}
$$

This $\Longrightarrow$

$$
\beta_r u_r \in \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}]
$$

$\Longrightarrow$

$$
\beta_r u_r \in \mathcal{W}_r
$$

$\Longrightarrow$

$$
\beta_r = 0 \text{ (since otherwise} u_r \in \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(r-1)}] \text{ - a contradiction)}
$$

Thus we get $\beta_r = 0$ and hence we get

$$
\sum_{j=1}^{r} \alpha_j A^2 u_j + \sum_{j=1}^{(r-1)} \beta_j A u_j = \theta_n
$$

Continuing this process we get all the β_j as zero. Hence we get

$$
\sum_{j=1}^{r} \alpha_j A^2 u_j = \theta_n
$$

But this gives us all $\alpha_j = 0$ since we have already shown that the set $\mathcal{S}_1$ is linearly independent. Thus we have

$$
\left.\begin{array}{l}\text{The set } \mathcal{S}_3 = A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r \\ \text{is linearly independent}\end{array}\right\} \quad (3.6.9)
$$

Now the set

$$
A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r, u_1
$$

is linearly independent since all vectors except u_r are linearly independent vectors in the subspace $\mathcal{N}_{A^{(k-1)}}$ and u_1 is outside this subspace. Next the set

$$
A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r, u_1, u_2
$$

is linearly independent since all vectors except u_2 are in the subspace $\mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1]$ and u_2 is outside this subspace. Continuing this process we get,

$$
\left.\begin{array}{l}\text{The set} \\ A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r, u-1, u_2, \cdots, u_r \\ \text{is linearly independent}\end{array}\right\} \quad (3.6.10)
$$

Thus we have

Property 4:

(Let k be an integer ≥ 3). Then

$u_1, u_2, \cdots, u_r$ are vectors in $\mathcal{N}_{A^k}$ such that $u_j \notin \mathcal{N}_{A^2} + \mathcal{L}[u_1, u_2, \cdots, u_{(j-1)}]$

$\Longrightarrow$

a) The vectors $A^2u_1, A^2u_2, \cdots, A^2u_r$ are linearly independent in $\mathcal{N}_{A^{(k-2)}}$

b) The vectors $Au_1, Au_2, \cdots, Au_r$ are linearly independent in $\mathcal{N}_{A^{(k-1)}}$

c) The vectors $A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r$ are linearly independent in $\mathcal{N}_{A^{(k-1)}}$

d) The vectors

$$A^2u_1, A^2u_2, \cdots, A^2u_r, Au_1, Au_2, \cdots, Au_r, u_1, u_2, \cdots, u_r$$

are linearly independent in $\mathcal{N}_{A^k}$ Analogously, we can prove the following generalisation:

Property 5:

Let k be any positive integer.

$u_1, u_2, \cdots, u_r$ are vectors in $\mathcal{N}_{A^k}$ such that

i) $u_1 \notin \mathcal{N}_{A^{(k-1)}}$ and

ii) $u_j \notin \mathcal{N}_{A^{(k-1)}} + \mathcal{L}[u_1, u_2, \cdots, u_{(j-1)}]$

$\Longrightarrow$

a) The vectors $A^{(k-1)}u_1, A^{(k-1)}u_2, \cdots, A^{(k-1)}u_r$ form a linearly independent set in $\mathcal{N}_A$

b) The vectors $A^{(k-2)}u_1, A^{(k-2)}u_2, \cdots, A^{(k-2)}u_r$ form a linearly independent set in $\mathcal{N}_{A^2}$, and in general,

c) In general the vectors $A^{(k-j)}u_1, A^{(k-j)}u_2, \cdots, A^{(k-j)}u_r$ form a linearly independent set in $\mathcal{N}_{A^{(k-j)}}$ for $j = 1, 2, \cdots, (k-1)$

d) The vectors $\left\{\left\{A^{(k-j)}u_\ell\right\}_{\ell=1}^{r}\right\}_{j=1}^{k}$ form a linearly independent set in $\mathcal{N}_{A^k}$

e) For any j, $(1 \leq j \leq k)$, the vectors $\left\{A^{(k-j)}u_\ell\right\}_{\ell=1}^{r}$ form a linearly independent set in $\mathcal{N}_{A^{(k-j)}}$