Chapter 5

Diagonalization, Eigenvalues and Eigenvectors

5.1 Introduction

We now study the structure of a diagonalizable matrix $A \in \mathbb{F}^{n\times n}$. Recall that we define the diagonalizability of a matrix as follows:

Definition 5.1.1 A matrix $\boldsymbol{A} \in \mathbb{F}^{\boldsymbol{n\times n}}$ is said to be diagonalizable over $\mathbb{F}$ if there exists an ivertible matrix $\boldsymbol{P} \in \mathbb{F}^{\boldsymbol{n\times n}}$ such that $\boldsymbol{P^{-1}AP}$ is a diagonal matrix $\boldsymbol{D} \in \mathbb{F}^{\boldsymbol{n\times n}}$

Let us consider a diagonalizable matrix $A \in \mathbb{F}^{n\times n}$ and analyse what are the ingredients that make the matrix a diagonalizable matrix. Since A is diagonalizble we must have an invertible $P \in \mathbb{F}^{n\times n}$ such that

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{pmatrix} \in \mathbb{F}^{n\times n}$$

We can write this as

$$AP = PD$$

If we now denote the jth column of P as P_j then we get

$$A[P_1\ P_2\ \cdots\ \ P_j \cdots\ P_n] \;=\; [P_1\ P_2\ \cdots\ \ P_j \cdots\ P_n]D$$

From this we get

$$[AP_1\ AP_2\ \cdots\ \ AP_j \cdots\ AP_n] \;=\; [\lambda_1 P_1\ \ \lambda_2 P_2\ \cdots\ \cdots\ \lambda_j P_j \cdots\ \lambda_n P_n]$$

Comparing the jth columns on both sides we get,

$$AP_j \;=\; \lambda_j P_j \text{ for } 1 \leq j \leq n \tag{5.1.1}$$

We note that the column vectors $P_1, P_2, \cdots P_n$ are linearly independent vectors in $\mathbb{F}^n$, (since P is invertible). Thus we have,

Conclusion :

$\boldsymbol{A} \in \mathbb{F}^{\boldsymbol{n}\times\boldsymbol{n}}$ is diagonalizable over $\mathbb{F}$

$\Longrightarrow$

There exist $\boldsymbol{n}$ scalars $\boldsymbol{\lambda_1, \lambda_2, \cdots, \lambda_n}$ in $\mathbb{F}$ and $\boldsymbol{n}$ linearly independent vectors $\boldsymbol{P_1, P_2, \cdots, P_n}$ in $\mathbb{F}^{\boldsymbol{n}}$ such that

$$\boldsymbol{AP_j = \lambda_j P_j} \textbf{ for } \boldsymbol{1 \leq j \leq n}$$

Conversely, it is easy to see that if there exists n scalars $\lambda_1, \lambda_2, \cdots, \lambda_n$ in $\mathbb{F}$ and n linearly independent vectors $P_1, P_2, \cdots, P_n$ in $\mathbb{F}^n$ such that

$$AP_j = \lambda_j P_j \text{ for } 1 \leq j \leq n$$

then we can define $P \in \mathbb{F}^{n\times n}$ as the matrix whose jth column is P_j, and then we get $P^{-1}AP$ as the diagonal matrix whose n diagonal entries are respectively $\lambda_1, \lambda_2, \cdots \lambda_n$. Combining this with the above Conclusion we get

Theorem 5.1.1 **$\boldsymbol{A} \in \mathbb{F}^{\boldsymbol{n}\times\boldsymbol{n}}$ is diagonalizable over $\mathbb{F}$**

$\Longleftrightarrow$

There exist $\boldsymbol{n}$ scalars $\boldsymbol{\lambda_1, \lambda_2, \cdots, \lambda_n}$ in $\mathbb{F}$ and $\boldsymbol{n}$ linearly independent vectors $\boldsymbol{P_1, P_2, \cdots, P_n}$ in $\mathbb{F}^{\boldsymbol{n}}$ such that

$$\boldsymbol{AP_j = \lambda_j P_j} \textbf{ for } \boldsymbol{1 \leq j \leq n}$$

Thus if we have to diagonalize a matrix A we need n pairs (λ_j, P_j). where $\lambda_j \in \mathbb{F}$ and $P_j \in \mathbb{F}^n$ such that $AP_j = \lambda_j P_j$. This leads us to the notion of eigenvalues and eigenvectors.

Remark 5.1.1 While seeking these n pairs (λ_j, P_j), it is not necessary that the scalars λ_j be distinct. Some of them may even be repeated. However, the vectors P_j that we are seeking must be linearly independent and hence they must all be nonzero vectors.

5.2 Eigenvalues and Eigenvectors

We begin with the definition of eigenvalues and eigenvectors.

Definition 5.2.1 A scalar $\boldsymbol{\lambda} \in \mathbb{F}$ is said to be an eigenvalue of a matrix $\boldsymbol{A} \in \mathbb{F}^{n\times n}$ if there exists a nonzero vector $\boldsymbol{\varphi} \in \mathbb{F}^n$ such that

$$\boldsymbol{A\varphi} \;=\; \boldsymbol{\lambda\varphi} \tag{5.2.1}$$

If $\boldsymbol{\lambda} \in \mathbb{F}$ is an eigenvalue of $\boldsymbol{A}$ then any nonzero vector $\boldsymbol{\varphi} \in \mathbb{F}^n$ such that $\boldsymbol{A\varphi} = \boldsymbol{\lambda\varphi}$ is called an eigenvector corresponding to the eigenvalue $\boldsymbol{\lambda}$. We shall call an eigenvalue-eigenvector pair $(\boldsymbol{\lambda}, \boldsymbol{\varphi})$, as an eigenpair

Given any $A \in \mathbb{F}^{n\times n}$ can we find n such eigenpairs in which the vectors in the n pairs are all linearly independent? We shall first look at some examples

Example 5.2.1 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined as follows:

$$A \;=\; \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \tag{5.2.2}$$

Consider $\lambda_1 = 1 \in \mathbb{R}$ and the nonzero vector $\varphi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^2$. Then we have

$$\begin{aligned} A\varphi_1 &= \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= 1\varphi_1 \\ &= \lambda_1\varphi_1 \end{aligned}$$

Hence $(1, \varphi_1)$ is one eigenpair for this matrix. Next, consider $\lambda_2 = 2$ and $\varphi_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then we have

$$\begin{aligned} A\varphi_1 &= \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= 2\varphi_2 \\ &= \lambda_2\varphi_2 \end{aligned}$$

Thus we see that $(2, \varphi_2)$ is another eigenpair. Moreover φ_1, φ_2 are linearly independent in $\mathbb{R}^2$. Hence we have two eigenpairs as required. Hence the matrix A is diagonalizable over $\mathbb{R}$. If we define

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

then it is easy to check that $P \in \mathbb{R}^{2\times 2}$ is invertible and

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \text{ a diagonal matrix in } \mathbb{R}^{2\times 2} \end{aligned}$$

Example 5.2.2 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined as follows:

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

We shall see that this matrix is not diagonalizable over $\mathbb{R}$. Suppose A is diagonalizable over $\mathbb{R}$. Then there must exist an invertible $P \in \mathbb{R}^{2\times 2}$ such that $P^{-1}AP$ is a diagonal matrix $D \in \mathbb{R}^{2\times 2}$. Let

$$\begin{aligned} P &= \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathbb{R}^{2\times 2} \\ D &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathbb{R}^{2\times 2} \end{aligned}$$

Then we have

$$\begin{aligned} P^{-1}AP &= D \\ &\Longrightarrow \\ AP &= PA \\ &\Longrightarrow \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} &= \begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \\ &\Longrightarrow \\ \begin{pmatrix} -r & -s \\ p & q \end{pmatrix} &= \begin{pmatrix} \lambda_1 p & \lambda_2 q \\ \lambda_1 r & \lambda_2 s \end{pmatrix} \end{aligned}$$

Comparing the first column on both sides we get

$$\begin{aligned}
-r &= \lambda_1 p \\
p &= \lambda_1 r \\
\Longrightarrow & \\
-r &= \lambda_1^2 r \\
\Longrightarrow & \\
(1+\lambda_1^2) r &= 0 \\
\Longrightarrow & \\
r &= 0 \text{ (since } \lambda_1 \text{ is real)}
\end{aligned}$$

Similarly, comparison of the second columns gives

$$s = 0$$

Hence the second row of P is zero row and hence P is not invertible - a contradiction. Thus this matrix is not diagonalizable over $\mathbb{R}$

Example 5.2.3 Consider the same matrix A of the above example but now we treat A as a matrix in $\mathbb{C}^{2\times 2}$. Consider $\lambda_1 = i$ and $\varphi_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. Then we have

$$\begin{aligned}
A\varphi_1 &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} \\
&= \begin{pmatrix} i \\ 1 \end{pmatrix} \\
&= i \begin{pmatrix} 1 \\ -i \end{pmatrix} \\
&= i\varphi_1
\end{aligned}$$

Thus we see that (i, φ_1) is an eigenpair. Similarly, we can verify that $(-i, \varphi_2)$ is an eigenpair, where, $\varphi_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$. We further note that φ_1, φ_2 are linearly independent in $\mathbb{C}^2$. Thus if we define

$$P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

it is easy to verify that

$$P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \text{ a diagonal matrix over } \mathbb{C}^{2\times 2}$$

Thus A is diagonalizable over $\mathbb{C}$

From the above examples, it follows that we have the following possibilities:

1. There are some matrices for which we can get these requisite number of eigenpairs and hence these matrices are diagonalizable

2. There are some matrices for which we cannot get these requisite number of eigenpairs and hence we cannot diagonalize these

3. There are some real matrices which are diagonalizable over $\mathbb{C}$ but not over $\mathbb{R}$

The main question that remains still is that of finding these eigenpairs. If we know an eigenvalue λ then we can find the possible eigenvectors corresponding to this eigenvalue by solving the homogeneous system

$$(A - \lambda I)x = \theta_n$$

Hence we search for the eigenvalues. How do we locate these eigenvalues? We shall discuss this next.

5.3 Characteristic Polynomial and Algebraic Multiplicity

Where should we look for the eigenvalues of a matrix? We have the following:

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A \in \mathbb{F}^{n\times n} &\iff \exists \varphi \in \mathbb{F}^n \ni A\varphi = \lambda\varphi \\ &\iff A_\lambda \varphi = \theta_n \\ &\iff \text{The homogeneous system } A_\lambda x = \theta_n \text{ has} \\ & \text{ nontrivial solution } \varphi \\ &\iff |A_\lambda| = 0 \text{ where } A_\lambda = \lambda I - A \end{aligned}$$

Thus the eiegnvalues are the roots of the function $c_A(\lambda)$ where

$$c_A(\lambda) = |\lambda I - A| \tag{5.3.1}$$

We have

$$c_A(\lambda) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & \cdots & -a_{1j} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & \cdots & -a_{2j} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_{j1} & -a_{j2} & \cdots & \cdots & \lambda - a_{jj} & \cdots & -a_{jn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \cdots & -a_{nj} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

When we expand this determinant we get a polynomial,

$$c_A(\lambda) = \lambda^n - (Trace(A))\lambda^{(n-1)} + \cdots + (-1)^n|A| \tag{5.3.2}$$

where

$$Trace(A) = \sum_{i=1}^{n} a_{ii} \text{ (sum of the diagonal entries of } A) \tag{5.3.3}$$

$c_A(\lambda)$ is a MONIC polynomial of degree n over $\mathbb{F}$, that is, $c_A(\lambda) \in \mathbb{F}[\lambda]$. This polynomial is called the **Characteristic Polynomial** of the matrix A. The eigenvalues that we are looking for are precisely the roots of this polynomial in $\mathbb{F}$. We shall now look at some examples.

Example 5.3.1 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined below:

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

We have

$$\begin{aligned} c_A(\lambda) &= |\lambda I - A| \\ &= \begin{vmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 2 \end{vmatrix} \\ &= (\lambda - 1)(\lambda - 2) \end{aligned}$$

Hence we have

$$c_A(\lambda) \;=\; (\lambda-1)(\lambda-2)$$

The roots are $\lambda_1 = 1$ and $\lambda_2 = 2$ and both are real. Thus we are able to get two eigenvalues in $\mathbb{R}$ for this $A \in \mathbb{R}^{2\times 2}$.

Example 5.3.2 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined below:

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

We have

$$\begin{aligned} c_A(\lambda) &= |\lambda I - A| \\ &= \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} \\ &= \lambda^2 + 1 \end{aligned}$$

Thus the characteristic polynomial is given by

$$c_A(\lambda) \;=\; \lambda^2 + 1$$

The polynomial has no real roots and hence there are no eigenvalues in $\mathbb{R}$ for this matrix.
However, if we consider A as matrix in $\mathbb{C}^{2\times 2}$ then we have the two roots for the characteristic polynomial as $\lambda_1 = i$ and $\lambda_2 = -i$. Thus this matrix has eigenvalues over the field $\mathbb{C}$ but not over the field $\mathbb{R}$.

Example 5.3.3 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined below:

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

We have

$$\begin{aligned} c_A(\lambda) &= |\lambda I - A| \\ &= \begin{vmatrix} \lambda & -1 \\ 0 & \lambda \end{vmatrix} \\ &= \lambda^2 \end{aligned}$$

Thus the characteristic polynomial is given by

$$c_{_A}(\lambda) \;=\; \lambda^2$$

Hence the characteristic equation is

$$\lambda^2 \;=\; 0$$

It has repeated roots $\lambda_1 = \lambda_2 = 0$

Example 5.3.4 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ defined below:

$$A \;=\; \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

We have

$$\begin{aligned} c_{_A}(\lambda) &= |\lambda I - A| \\ &= \begin{vmatrix} \lambda - 1 & -1 & 0 \\ 0 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} \\ &= (\lambda - 1)(\lambda - 2)^2 \end{aligned}$$

Hence we have

$$c_{_A}(\lambda) \;=\; (\lambda - 1)(\lambda - 2)^2$$

The roots are $\lambda_1 = 1$ and $\lambda_2 = 2$ and both are real. The eigenvalue $\lambda_2 = 2$ is repeated twice. Thus we are able to get three eigenvalues in $\mathbb{R}$ for this $A \in \mathbb{R}^{2\times 2}$ when we count the multiplicities also.

From the above examples, it follows that, in general , the characteristic polynomial of $A \in \mathbb{F}^{n\times n}$ need not have n roots in $\mathbb{F}$, and that when there are roots these may be repeated. However, consider a diagonalisable matrix $A \in \mathbb{F}^{n\times n}$. Then we have an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $P^{-1}AP = D$ where $D \in \mathbb{F}^{n\times n}$ is a diagonal matrix, whose diagonal entries, $d_1, d_2, \cdots, d_n$ are in $\mathbb{F}$ and may not be distinct. So we have,

$$A \;=\; PDP^{-1}$$

$$
\begin{aligned}
&= P \begin{pmatrix} d_1 & & & & \\ & d_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & d_n \end{pmatrix} P^{-1} \\
\Longrightarrow & \\
c_{_A}(\lambda) &= \left|\lambda I - PDP^{-1}\right| \\
&= \left|\lambda PIP^{-1} - PDP^{-1}\right| \\
&= \left|P(\lambda I - D)P^{-1}\right| \\
&= |P|\,|\lambda I - D|\,\left|P^{-1}\right| \\
&= c_{_A}(D) \text{ (since } |P|\,|P^{-1}| = 1) \\
&= (\lambda - d_1)(\lambda - d_2)\cdots(\lambda - d_n)
\end{aligned}
$$

Hence we have that for a diagonalizable matrix $A \in \mathbb{F}^{n\times n}$, the characteristic polynomial, $c_{_A}(\lambda)$, can be completely factorized into linear factors, (some of which may be repeated). Hence for a diagonalisable matrix the characteristic polynomial can be written as

$$c_{_A}(\lambda) \;=\; (\lambda - \lambda_1)^{a_1}(\lambda - \lambda_2)^{a_2}\cdots(\lambda - \lambda_k)^{a_k}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct roots in $\mathbb{F}$ and $a_1, a_2, \cdots, a_k$ are their multiplicities. We shall, therefore, first look at matrices for which the characteristic polynomial can be factorized as above.

Thus we consider the following type of matrices for the rest of this chapter:

$A \in \mathbb{F}^{n\times n}$ such that the characteristic polynomial of A is of the form,

$$c_{_A}(\lambda) \;=\; (\lambda - \lambda_1)^{a_1}(\lambda - \lambda_2)^{a_2}\cdots(\lambda - \lambda_k)^{a_k} \tag{5.3.4}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct elements of $\mathbb{F}$ and $a_1, a_2, \cdots, a_k$ are positive integers such that

$$a_1 + a_2 + \cdots + a_k \;=\; n \tag{5.3.5}$$

(Clearly, this is possible for all matrices in $\mathbb{C}^{n\times n}$).

For such matrices $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the distinct eigenvalues and $a_1, a_2, \cdots, a_k$ are their multiplicities as roots of the characteristic polynomial. The multiplicity a_j is called the **Algebraic Multiplicity** of the eigenvalue λ_j.

5.4 Eigenspaces and Geometric Multiplicity

Let $A \in \mathbb{F}^{n\times n}$ and let its characteristic polynomial be as in (5.3.4). Hence the distinct eigenvalues of A are $\lambda_1, \lambda_2, \cdots, \lambda_k$ with algebraic multiplicities, $a_1, a_2, \cdots, a_k$ respectively. Having found the eigenvalues we next look at the eigenvectors. We define

$$\begin{aligned} \mathcal{W}_j &= Null\ Space\ of\ A - \lambda_j I &\quad (5.4.1)\\ &= \{x \in \mathbb{F}^n : Ax = \lambda_j x\} &\quad (5.4.2) \end{aligned}$$

Any nonzero vector in $\mathcal{W}_j$ is an eigenvector corresponding to the eigenvalue λ_j. Since $\mathcal{W}_j$ is the Null Space of the matrix $A - \lambda_j I$, it is a subspace of $\mathbb{F}^n$. This subspace is called the **Eigenspace** corresponding to the eigenvalue λ_j. The dimension of this subspace is called the **Geometric Multiplicity** of the eigenvalue λ_j and we denote this by g_j. Thus we have

$$\text{geometric multiplicity, } g_j = dimension\ of\ \mathcal{W}_j \qquad (5.4.3)$$

We now look at some simple examples.

Example 5.4.1 For the matrix A in the Example 5.3.1 we have

$$\lambda_1 = 1,\ \lambda_2 = 2,\ a_1 = 1,\ a_2 = 1$$

For the eigenspaces we have,

$$\begin{aligned} \mathcal{W}_1 &= Null\ Space\ of\ A_I \\ &= Null\ Space\ of \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \left\{ x \in \mathbb{R}^2 : x = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \alpha \in \mathbb{R} \right\} \\ \mathcal{W}_2 &= Null\ Space\ of\ A - 2I \\ &= Null\ Space\ of \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \\ &= \left\{ x \in \mathbb{R}^2 : x = \begin{pmatrix} \beta \\ \beta \end{pmatrix}, \beta \in \mathbb{R} \right\} \end{aligned}$$

We see that

$$\begin{aligned} g_1 &= dimension\ of\ \mathcal{W}_1 = 1 = a_1 \\ g_2 &= dimension\ of\ \mathcal{W}_2 = 1 = a_2 \end{aligned}$$

Thus, in this case, for each eigenvalue, the algebraic multiplicity is the same as the geometric multiplicity.

Example 5.4.2 For the matrix A of Example 5.3.3 we have $\lambda_1 = 0$ is the only eigenvalue and its algebraic multiplicity is $a_1 = 2$. For the eigenspace we have

$$\begin{aligned} \mathcal{W}_1 &= Null\ Space\ of\ A - 0I \\ &= Null\ Space\ of\ A \\ &= Null\ Space\ of\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ &= \left\{ x \in \mathbb{R}^2 : x = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \alpha \in \mathbb{R} \right\} \end{aligned}$$

Thus the dimension of $\mathcal{W}_j$ is 1 and hence we heve $g_1 = 1$. Thus, in this case we have an eigenvalue whose geometric multiplicity is smaller than its algebraic multiplicity.

Example 5.4.3 For the matrix of Example 5.3.4 we have

$$c_A(\lambda) = (\lambda - 1)(\lambda - 2)^2$$

and hence

$$\lambda_1 = 1,\ \lambda_2 = 2 \text{ and } a_1 = 1,\ a_2 = 2$$

We now find the eigenspaces:

$$\begin{aligned} \mathcal{W}_1 &= Null\ Space\ of\ (A - \lambda_1 I) \\ &= Null\ Space\ of\ (A - I) \\ &= Null\ Space\ of\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \left\{ \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\} \end{aligned}$$

Hence we have

$$g_1 = dimension\ \mathcal{W}_1 = 1$$

For $\lambda_2 = 2$ we have

$$\begin{aligned} \mathcal{W}_2 &= Null\ Space\ of\ (A - 2I) \\ &= Null\ Space\ of \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \left\{ \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\} \end{aligned}$$

Hence we have

$$g_2 = dimension\ \mathcal{W}_2 = 2$$

Thus in this example we have

$$a_1 = 1 = g_1 \text{ and } a_2 = 2 = g_2$$

In the above examples we found that in Examples 5.4.1 and 5.4.3 we have the geometric multiplicity of every eigenvalue same as the algebraic multiplicity, and in Example 5.4.2 there is an eigenvalue with its geometric multiplicity smaller than its algebraic multiplicity. We shall now see that these are the only two possibilities and that the geometric multiplicity of no eigenvalue can exceed its algebraic multiplicity.
Since the dimension of $\mathcal{W}_j$ is g_j, any basis for $\mathcal{W}_j$ consists of g_j vectors. Let

$$\mathcal{B}_j \ : \ \varphi_1, \varphi_2, \cdots, \varphi_{g_j}$$

be a basis for $\mathcal{W}_j$. We can extend this to a basis for $\mathbb{F}^n$ by appending $n - g_j$ suitable linearly independent vectors. Let

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

be a basis for $\mathbb{F}^n$ which is an extension of $\mathcal{B}_j$. Let P be the matrix whose columns are these n basis vectors. We have

$$P = \left[\varphi_1\ \varphi_2 \cdots\ \varphi_{g_j}\ v_1\ v_2 \cdots\ v_{(n-g_j)}\right]$$

Clearly P is invertible since the columns are all linearly independent. Further, we have,

$$
\begin{aligned}
AP &= \left[A\varphi_1\ A\varphi_2 \cdots\ A\varphi_{g_j}\ Av_1\ Av_2 \cdots\ Av_{(n-g_j)}\right] \\
&= \left[\lambda_j\varphi_1\ \lambda_j\varphi_2 \cdots\ \lambda_j\varphi_{g_j}\ Av_1\ Av_2 \cdots\ Av_{(n-g_j)}\right] \\
&= \left[\varphi_1\ \varphi_2 \cdots\ \varphi_{g_j}\ v_1\ v_2 \cdots\ v_{(n-g_j)}\right] \left(\begin{array}{c|c} \lambda_j I_{(g_j\times g_j)} & \mathrm{K}_{g_j\times(n-g_j)} \\ \hline 0_{(n-g_j)\times g_j} & \mathrm{L}_{(n-g_j)\times(n-g_j)} \end{array}\right) \\
\Longrightarrow & \\
A &= P \left(\begin{array}{c|c} \lambda_j I_{(g_j\times g_j)} & \mathrm{K}_{g_j\times(n-g_j)} \\ \hline 0_{(n-g_j)\times g_j} & \mathrm{L}_{(n-g_j)\times(n-g_j)} \end{array}\right) P^{-1}
\end{aligned}
$$

Hence we get

$$
c_{_A}(\lambda) \;=\; (\lambda-\lambda_j)^{g_j} c_{_L}(\lambda)
$$

Thus we see that λ_j must be a root of $c_{_A}(\lambda)$ of multiplicity at least g_j. Thus we have $g_j \leq a_j$. Further, since the subspace $\mathcal{W}_j$ has at least one nonzero vector, (because λ_j is an eigenvalue), it follows that $1 \leq g_j$. Thus we have

$$
1 \leq g_j \leq a_j \text{ for } 1 \leq j \leq k \tag{5.4.4}
$$

Thus ,
The geometric multiplicity of any eigenvalue is less than or equal to its algebraic multiplicity, and the geometric multiplicity is at least one.
We shall now look at some properties of the eigenvectors.
1) We have

$$
\begin{aligned}
x \in \mathcal{W}_j \;\Longrightarrow\;\; & Ax = \lambda_j x \\
& A^2x = A(Ax) = A(\lambda_j x) = \lambda_j^2 x
\end{aligned}
$$

Analogously we get

$$
x \in \mathcal{W}_j \;\Longrightarrow\; A^\ell x = \lambda_j^\ell x \tag{5.4.5}
$$

2) Let

$$p(\lambda) = a_r\lambda^r + a_{(r-1)}\lambda^{(r-1)} + \cdots + a_1\lambda + a_0$$

be any polynomial over $\mathbb{F}$. We then define

$$p(A) = a_rA^r + a_{(r-1)}A^{(r-1)} + \cdots + a_1A + a_0I$$

We have

$$\begin{aligned} x \in \mathcal{W}_j \quad \Longrightarrow \quad p(A)x &= \sum_{\ell=0}^{r} a_\ell A^\ell x \\ &= \sum_{\ell=0}^{r} a_\ell \lambda_j^\ell x \\ &= p(\lambda_j)x \end{aligned}$$

Thus we have

$$\left.\begin{array}{c} x \in \mathcal{W}_j \Longrightarrow p(A)x = p(\lambda_j)x \\ \text{for every polynomial } p(\lambda) \text{ over } \mathbb{F} \end{array}\right\} \tag{5.4.6}$$

3)Consider the polynomial

$$P_j(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_{j-1)})(\lambda - \lambda_{(j+1)})\cdots(\lambda - \lambda_k)$$

We have

$$P_j(\lambda_\ell) \quad = \quad \begin{cases} 0 & \text{if} \quad \ell \neq j \\ P_j(\lambda_j) \neq 0 & \text{if} \quad \ell = j \end{cases} \tag{5.4.7}$$

Hence we have

$$\begin{aligned} x \in \mathcal{W}_\ell \quad \Longrightarrow \quad P_j(x) &= P_j(\lambda_\ell)x \\ &= \begin{cases} \theta_n & \text{if} \quad \ell \neq j \\ P_j(\lambda_j)x & \text{if} \quad \ell = j \end{cases} \end{aligned}$$

Thus we have

$$x \in \mathcal{W}_\ell \quad \Longrightarrow \quad \begin{cases} \theta_n & \text{if} \quad \ell \neq j \\ P_j(\lambda_j)x & \text{if} \quad \ell = j \end{cases} \tag{5.4.8}$$

We shall now see an important consequence of this fact.
4) We have

$$\begin{aligned}
x_\ell \in \mathcal{W}_\ell \text{ for } 1 \le \ell \le k \text{ and } x_1 + x_2 + \cdots + x_k = \theta_n \\
\Longrightarrow \\
P_j(A)(x_1 + x_2 + \cdots + x_k) = P_j(A)(\theta_n) = \theta_n \\
\Longrightarrow \\
P_j(A)(x_1 + x_2 + \cdots + x_k) = \sum_{\ell=1}^{k} P_j(A)x_\ell = \theta_n \\
\Longrightarrow \\
\sum_{\ell=1}^{k} P_j(\lambda_\ell)x_\ell = \theta_n \\
\Longrightarrow \\
P_j(\lambda_j)x_j = \theta_n \text{ since } P_j(\lambda_\ell) = 0 \text{ if } \ell \ne j \\
\Longrightarrow \\
x_j = \theta_n \text{ since } P_j(\lambda_j) \ne 0
\end{aligned}$$

Thus we have

$$\left.\begin{array}{l}
x_j \in \mathcal{W}_j, ; 1 \le \mathrm{J} \le k \text{ and } x_1 + x_2 + \cdots + x_k = \theta_n \Longrightarrow \\
x_j = \theta_n \text{ for every } j = 1, 2, \cdots, k
\end{array}\right\}$$

Thus the sum $\mathcal{W}_1 + \mathcal{W}_2 + \cdots + \mathcal{W}_k$ must be a Direct Sum. Hence we have a subspace $\mathcal{W}$ which is a Direct Sum of the eigenspaces:

$$\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \cdots \oplus \mathcal{W}_k$$

Will this subspace $\mathcal{W}$ obtained as the Direct Sum of the eigenspaces be all of $\mathbb{F}^n$ or will this be a proper subspace of $\mathbb{F}^n$? We shall see that the answer to this question is directly related to the fact whether A is diagonalizable or not. We shall see that $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if

$$\mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \cdots \oplus \mathcal{W}_k = \mathbb{F}^n$$