Solutions of Matrix Equations

Thir Raj Dangal

University of Southern Mississippi
The University of Southern Mississippi

SOLUTIONS OF MATRIX EQUATIONS

by

Thir Raj Dangal

A Thesis
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Science

May 2013
ABSTRACT

SOLUTIONS OF MATRIX EQUATIONS

by Thir Raj Dangal

May 2013

Two $n \times n$ matrices $A$ and $B$ can make many different matrix equations, e.g., $AB = BA$, $AB = AA$, $ABA = BAB$, and $AAB = BAA$. It is not always easy to describe solutions to these matrix equations. This thesis considers the problem of describing solutions to the matrix equations $A^2 = B^2$, $AB = A^2$ and $AB = B^2$. This problem is motivated by considering the properties of commutative matrices (i.e., $AB = BA$), as well as the matrix form of the Yang-Baxter equation, $ABA = BAB$.

For each of these equations, solutions are provided such that $A \neq B$. However, in the case of $A^2 = B^2$ (when $A$ has many distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that $\lambda_i \neq -\lambda_j$) and in the case of $AB = A^2$, the matrices $A$ and $B$ must have a common eigenvector. In addition, matrices arising from graphs are considered, and restrictions are determined which imply unique solutions to the matrix equation $A^2 = B^2$. 


# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td><strong>1 BACKGROUND</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Commuting Matrices</td>
<td>1</td>
</tr>
<tr>
<td>1.3 The Yang-Baxter Equation</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Other Matrix Equations</td>
<td>8</td>
</tr>
<tr>
<td><strong>2 THE MATRIX EQUATION $A^2 = B^2$</strong></td>
<td>9</td>
</tr>
<tr>
<td>2.1 An Example for $A^2 = B^2$, with $A \neq B$</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Common Eigenvectors</td>
<td>9</td>
</tr>
<tr>
<td>2.3 A Generalization of Theorem 2.2.2</td>
<td>13</td>
</tr>
<tr>
<td>2.4 Unique Square Roots</td>
<td>13</td>
</tr>
<tr>
<td><strong>3 THE MATRIX EQUATIONS $AB = B^2$ AND $AB = A^2$</strong></td>
<td>19</td>
</tr>
<tr>
<td>3.1 An Example for $AB = B^2$, with $A \neq B$</td>
<td>19</td>
</tr>
<tr>
<td>3.2 An Example for $AB = A^2$, with $A \neq B$</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Matrix Commutators</td>
<td>22</td>
</tr>
<tr>
<td>3.4 Common Eigenvectors</td>
<td>23</td>
</tr>
<tr>
<td><strong>4 CONCLUSION</strong></td>
<td>25</td>
</tr>
<tr>
<td>4.1 Results</td>
<td>25</td>
</tr>
<tr>
<td>4.2 Future Work</td>
<td>25</td>
</tr>
<tr>
<td><strong>BIBLIOGRAPHY</strong></td>
<td>27</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

Figure

2.1 The tree $T_1$, which has adjacency matrix $A(T_1)$ .................................................. 14
2.2 Illustration indicating the existence of 4-cycles and 3-cycles ............................................. 16
2.3 The tree $T$, illustrating Corollary 2.4.3 .............................................................................. 17
2.4 An example illustrating Lemma 2.4.4 ................................................................................... 18
LIST OF TABLES

Table

4.1  The unique matrix equations of two matrices taken three at a time  . . . . . . 26
Chapter 1

BACKGROUND

1.1 Introduction

As an introduction and motivation for this project, two important topics will be discussed. The first topic is that of commuting matrices. Square matrices \( A \) and \( B \) of the same dimensions commute with each other if and only if \( AB = BA \). Given a matrix \( A \), properties of a "solution" \( B \) are desired.

The second topic is the Yang-Baxter Equation, specifically the matrix form, which is \( ABA = BAB \). In this case, for a given matrix \( A \), the goal is to find a matrix \( B \) with the same size such that \( ABA = BAB \). Here the matrix \( B \) is the solution of the Yang-Baxter Equation corresponding to the matrix \( A \).

The goal of this introduction is to motivate consideration of other possible products of two matrices taken two at a time. In particular, the matrix equations to be considered are \( A^2 = B^2 \), \( AB = B^2 \), and \( AB = A^2 \).

1.2 Commuting Matrices

For two square matrices \( A \) and \( B \) of the same dimensions, the products \( AB \) and \( BA \) always exist. However, since matrix multiplication is equivalent to the composition of linear maps, it is unrealistic to expect that the products taken in any order will be equal. For example, if

\[
A = \begin{bmatrix} 9 & 3 \\ 7 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & 8 \\ 3 & 1 \end{bmatrix},
\]

then, the matrix products are

\[
AB = \begin{bmatrix} 54 & 75 \\ 41 & 58 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 101 & 31 \\ 34 & 11 \end{bmatrix},
\]

from which it is plain to see that \( AB \neq BA \). For any matrices \( A \) and \( B \) of the same size, \( Tr(AB) = Tr(BA) \) even if \( AB \neq BA \). Also, it is the case that \( Tr(A^T B) = Tr(AB^T) \).

In the most trivial case, if the matrices \( A \) and \( B \) are equal, then \( AB = BA \). The more general case is somewhat more complicated. There are some special cases when the products \( AB \) and \( BA \) are equal. Consider the \( n \times n \) matrices \( A = \{a_{ij}\} \) and \( B = \{b_{ij}\} \). Some examples of cases in which these matrices commute are discussed in the following list:
1. For any two diagonal matrices $A$ and $B$, $AB = BA$ [6].

2. Two diagonalisable matrices commute iff they are simultaneously diagonalisable, i.e., there exists a matrix $M$ such that $MAM^{-1} = L_1$ and $MBM^{-1} = L_2$ are diagonal matrices [6].

3. For any two symmetric matrices $A$ and $B$, $AB$ is symmetric iff $AB = BA$.

4. A matrix $B$ is called a polynomial of matrix $A$ if the matrix $B$ can be expressed as the linear combination of different powers of the matrix $A$ where the matrices $A$ and $B$ are of the same size. For example, if the square matrix $B$ is a polynomial of the square matrix $A$, it can be written as

$$B = a_0 I + a_1 A + a_2 A^2 + a_3 A^3 + \ldots + a_n A^n,$$

where $a_0, a_1, a_2, \ldots, a_n$ are real numbers. Then,

$$AB = A(a_0 I + a_1 A + a_2 A^2 + \ldots + a_n A^n),$$

$$= a_0 A + a_1 A^2 + a_2 A^3 + \ldots + a_n A^{n+1},$$

$$= (a_0 I + a_1 A + a_2 A^2 + \ldots + a_n A^n)A,$$

$$= BA.$$

Thus, every polynomial in the matrix $A$ commutes with $A$. But the converse is not always true. If a matrix $B$ commutes with matrix $A$, then the matrix $B$ is a polynomial in $A$ if and only if $A$ has equal characteristic and minimum polynomial. Furthermore, if the matrix $B$ commutes with all commuting matrices of $A$, then $B$ is a polynomial in $A$.

5. A scalar multiple of an identity matrix commutes with all matrices of the same size (however it is important to note that in general, an arbitrary diagonal matrix will not necessarily commute with another matrix of the same dimensions).

6. The multiplication of $n \times n$ rotation matrices is commutative in the plane.

7. A matrix $A$ is called a normal matrix if $AA^* = A^*A$, where $A^*$ is the Hermitian matrix of $A$. Therefore, by definition, a normal matrix $A$ commutes with $A^*$ [4].

8. For a matrix $A$, if each of the descending diagonal from left to right is constant, it is
called a Toeplitz matrix. For example,

\[
A = \begin{bmatrix}
  a & b & c & d & e \\
  f & a & b & c & d \\
  g & f & a & b & c \\
  h & g & f & a & b \\
  i & h & g & f & a \\
\end{bmatrix}
\]

is a Toeplitz matrix. All Toeplitz matrices commute asymptotically because they are simultaneously diagonalisable when the row and column dimensions tend to infinity [5].

9. A circulant matrix is a matrix in which the first column is specified by a vector and the remaining columns are the cyclic permutations of the first column. For example,

\[
C = \begin{bmatrix}
  a & b & c & d & e \\
  e & a & b & c & d \\
  d & e & a & b & c \\
  c & d & e & a & b \\
  b & c & d & e & a \\
\end{bmatrix}
\]

is a circulant matrix. A circulant matrix is a special kind of Toeplitz matrix. For any circulant matrices \(A\) and \(B\), \(AB=BA\) [5].

10. A doubly stochastic matrix is a square matrix of non-negative real numbers such that each of rows and columns sum to 1.

A doubly stochastic matrix of the form \(\gamma I_n + (1 - \gamma)J_n\), \(I_n\) is the \(n \times n\) identity matrix and \(J_n\) is \(n \times n\) matrix with all entries equal to \(n^{-1}\) where \(0 \leq \gamma \leq 1\). For \(n = 1\), \(\gamma\) is arbitrary, for \(n > 1\), \(-(n - 1)^{-1}\) \(\leq \gamma \leq 1\). Such doubly stochastic matrices commute with any doubly stochastic matrix [12].

11. A permutation matrix is a matrix with a single one in each row and each column and zero elsewhere. For example,

\[
A = \begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

is a permutation matrix. Since permutation matrices are orthogonal matrices, i.e., for any permutation matrix \(A\), \(A^T = A^{-1}\). Therefore, \(A^TA = AA^T\), so a permutation matrix commutes with its transpose [13].
In addition to specific cases of matrices or families of matrices satisfying the commutative property, the following theorems provide some general properties of commutative matrices. These theorems will be used as motivation for considering common eigenvectors in $A$ and $B$ when $A^2 = B^2$ and when $AB = A^2$.

**Theorem 1.2.1.** If $A$ and $B$ are defined on a finite dimensional vector space over an algebraically closed field and $AB = BA$, then the matrices $A$ and $B$ have at least one common eigenvector [6].

**Proof.** Let $A$ and $B$ be two matrices of same size such that $AB = BA$. Let $u$ be the eigenvector of the matrix $B$ corresponding to the eigenvalue $\lambda$. Then, $Bu = \lambda u$. Now,

$$A(Bu) = A(\lambda u) = \lambda (Au).$$

But $A(Bu) = B(Au)$. Therefore $B(Au) = \lambda (Au)$ which means that $Au$ is also the eigenvector of $B$ corresponding to the same eigenvalue $\lambda$. Thus, for any eigenvector $u$, another eigenvector $Au$ can be found with the same eigenvalue. Let

$$X = \{x : Bx = \lambda x\}$$

be the eigenspace to the matrix $B$ corresponding to the eigenvalue $\lambda$. Then for any $x$, $Ax \in X$. Thus, the vector $Ax$ must be expressed as a linear combination of the basis vectors of $X$. Hence, there must be at least one $x \in X$ such that $Ax = \alpha x$ for some scalar $\alpha$ because every operator over an algebraically closed field has an eigenvector. Thus, $x$ is an eigenvector of both the matrices $A$ and $B$. Hence, the matrices $A$ and $B$ must have at least one common eigenvector if they are commuting. \qed

Theorem 1.2.1 can be used to prove the following well-known result.

**Theorem 1.2.2.** If $A$ and $B$ be two commuting operators on a finite dimensional vector space $V$, then a suitable basis can be chosen for which the matrices corresponding to both the operators $A$ and $B$ are upper triangular.

**Proof.** Let $V$ be a vector space with dimension $n$ and $A$ and $B$ be two commuting operators on $V$. Proceed by induction on $n$. By Theorem 1.2.1, there is at least one eigenvector $v_1 \in V$ such that $Av_1 = \lambda v_1$ and $Bv_1 = \alpha v_1$. Let $W$ be a subspace spanned
by \( v_1 \). Then dimension of the quotient set \( V/W \) is \( n - 1 \). Now form a basis for \( V \) from a given basis for the quotient set \( V/W \). Choose a basis for the quotient set \( V/W \) as

\[ u_1 + W, u_2 + W, u_3 + W, \ldots, u_k + W. \]

The next step is to show that \( u_1, u_2, u_3, \ldots, u_k, v_1 \) forms a basis for \( V \). One component of this is to show that the vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \) are linearly independent in \( V \). Let \( a_1, a_2, a_3, \ldots, a_k \) be scalars such that

\[ a_1 u_1 + a_2 u_2 + a_3 u_3 + \ldots + a_k u_k + c v_1 = 0. \quad (1.1) \]

Then

\[ a_1 u_1 + a_2 u_2 + \ldots + a_k u_k = -c v_1. \]

So,

\[ a_1 (u_1 + W) + a_2 (u_2 + W) + a_3 (u_3 + W) + \ldots + a_k (u_k + W) = 0 + W \]

which implies that \( a_1 = a_2 = a_3 = \ldots = a_k = 0 \). Also, using the Equation 1.1, it must be the case that

\[ 0 = -c v_1. \]

So, \( c = 0 \) as well. Thus, the vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \) are linearly independent in \( V \). Now, it remains to show that the set of vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \) spans \( V \).

Choose a vector \( v_1 \in V \) and add vectors to \( v_1 \) to form a basis \( v_1, x_1, x_2, x_3, \ldots, x_{n-1} \) in \( V \). Then, for any vector \( v \in V \), it is the case that \( v = x + c v_1 \) where \( x \) and \( v_1 \) are linearly independent. Consider the coset \( x + W \). Then, \( a_1 u_1 + a_2 u_2 + a_3 u_3 + \ldots + a_k u_k + W = x + W \). Thus, for some scalar \( c_1 \),

\[ a_1 u_1 + a_2 u_2 + a_3 u_3 + \ldots + a_k u_k + c_1 v_1 = x. \]

Adding \( c v_1 \) on both sides of this relation yields

\[ a_1 u_1 + a_2 u_2 + a_3 u_3 + \ldots + a_k u_k + c_1 v_1 + c v_1 = x + c v_1 = v. \]

Since the vector \( v \in V \) was arbitrary, every vector of the vector space \( V \) can be written as a linear combination of the vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \). Hence the set of the vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \) spans \( V \). Therefore, the set of vectors \( u_1, u_2, u_3, \ldots, u_k, v_1 \) forms a
basis for \( V \). The matrix corresponding to the operator \( A \) on the quotient set \( V/W \) is upper triangular, and as a consequence, for any \( u_i \in V \),

\[
A(u_i + W) = (a_1u_1 + a_2u_2 + a_3u_3 + \ldots + a_iu_i) + W,
\]

\[
A(u_i + cv_1) = (a_1u_1 + a_2u_2 + a_3u_3 + \ldots + a_iu_i) + c_1v_1,
\]

\[
A(u_i) + cA(v_1) = (a_1u_1 + a_2u_2 + a_3u_3 + \ldots + a_iu_i) + c_1v_1,
\]

\[
A(u_i) = (a_1u_1 + a_2u_2 + a_3u_3 + \ldots + a_iu_i) + c_1v_1 - cA(v_1),
\]

\[
A(u_i) = (a_1u_1 + a_2u_2 + a_3u_3 + \ldots + a_iu_i) + (c_1 - c\lambda)v_1.
\]

Hence, the matrix corresponding to the operator \( A \) on \( V \) is upper triangular. Similarly, the matrix corresponding to the operator \( B \) on \( V \) is upper triangular.

These properties associated with commuting matrices are used as motivation and direction for the following chapters. In addition, the following problem is also related. For matrices \( A \) and \( B \) of sizes \( m \times m \) and \( n \times n \), solve the matrix equation \( AX = XB \) for some matrix \( X \) of size \( m \times n \).

**Theorem 1.2.3.** [3] Let \( \mathbb{F} \) be an algebraically closed field and let \( A \in \mathbb{F}_{m \times m} \) and \( B \in \mathbb{F}_{n \times n} \). Then the matrix equation \( AX = XB \) has a nonzero solution if and only if \( A \) and \( B \) have a common eigenvalue.

Replacing the matrix \( B \) by the matrix \( A \) in the above matrix equation yields \( AX =XA \). Then the set \( C(A) = \{ X \in F_{n \times n} : AX =XA \} \) is a subspace of \( F_{n \times n} \) and is called centralizer of \( A \).

Consider the \( 2 \times 2 \) matrix \( A \). The set of all matrices \( X \) such that \( AX =XA \) can be examined directly. For example, let

\[
A = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}
\]

and let

\[
X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.
\]

Then, \( AX =XA \) gives

\[
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.
\]
This matrix equation results in the following list of equations:

\[ a + 3c = a + 5b, \]
\[ b + 3d = 3a + 7b, \]
\[ 5a + 7c = c + 5d, \]
\[ 5b + 7d = 3c + 7d. \]

Solving these equations yields the following:

\[ a = s - 6t/5, \]
\[ b = 3t/5, \]
\[ c = t, \]
\[ d = s. \]

Hence the matrix \( X \) commutes with the matrix \( A \), where the entries of \( A \) are chosen as above, and \( s \) and \( t \) are free variables.

Let \( A \) and \( B \) be two square matrices of the same dimensions. These matrices can be multiplied, taken two at a time, in four different ways, namely \( AB, BA, AA, \) and \( BB \). When \( AB = BA \), the matrices \( A \) and \( B \) are said to commute. However, there are other possible relations between these product matrices (the topic of this thesis), namely

\[ AB = AA, AB = BB, \text{ and } BB = AA. \]

If \( A \) and \( B \) are identity matrices or zero matrices or equal matrices, all of the above relations are satisfied. If \( A \neq B \), this may not be true. If the matrix \( A \) is a nilpotent matrix of index 2 and the matrix \( B \) is a zero matrix, all of the above relations are true. Furthermore, if the matrix \( A \) is an idempotent matrix, i.e. \( AA = A \) and \( B \) is an identity matrix, then we can see that

\[ AB = BA = AA. \]

Similarly, one can consider different arrangements of products of two square matrices taken three at a time. Let \( A \) and \( B \) be two square matrices of same size. There are twenty eight different relationships that can be constructed between the eight different products of the matrices \( A \) and \( B \) taken three at a time. For example, the following relationships can be considered:

\[ ABA = BAB, AAB = ABB, \text{ and } AAA = BBB. \]

One such equation is \( ABA = BAB \), which is referred to as the matrix form of the Yang-Baxter Equation.
1.3 The Yang-Baxter Equation

The Yang-Baxter Equation was first introduced by C.N. Yang in 1967 when he published two papers on the study of a simple one-dimensional quantum many-body problem, and considered the following:

$$\sum_i p_i^2 + 2c \sum_{i>j} \delta(x_i - x_j).$$

The equation was $A(u)B(u+v)A(v) = B(u)A(u+v)B(v)$, where $A(u)$ and $B(v)$ are rational functions of $u$ and $v$.

The importance of the same equation was noticed by R.J. Baxter in 1972 when he was solving some classical statistical mechanical problems in two dimensions. The equation was established as the Yang-Baxter Equation by Faddeev in 1981 to denote a principle of integrability, i.e., exact solvability in a wide variety of problems in physics and mathematics [8]. It has been widely used in statistical mechanics, quantum field theory, knot theory, braid theory, quantum group theory, and other disciplines.

1.4 Other Matrix Equations

In the next sections, we consider problems which are a compromise between the matrix version of the Yang-Baxter Equation and the problem of determining properties of commuting matrices. As discussed previously, we look at the following matrix equations:

$$A^2 = B^2,$$
$$AB = B^2,$$
$$AB = A^2.$$

In some cases, there are simple solutions. As an example, $B = A$ is a solution to all equations, $B = 0$ is a solution to (1) and (2) and $B = -A$ is a solution for (3). Furthermore, if $A$ is invertible, solutions to (2) are restricted to $B = A$. Also, if solutions must be invertible matrices, (1) implies $B = A$.

However, in general, it may not be easy to determine all solutions, or properties of all solutions that satisfy these three matrix equations. Therefore, the goal in the following sections is three-fold. First, we examine solutions to the equations above such that $A \neq B$. Second, we determine general properties that must hold for a solution $B$, in terms of the matrix $A$. Third, to make further progress, we will place restrictions on the domain of both $A$ and $B$. 
Chapter 2

THE MATRIX EQUATION $A^2 = B^2$

2.1 An Example for $A^2 = B^2$, with $A \neq B$

In this chapter, we discuss matrices $A$ and $B$ that satisfy the property that their squares are equal. To begin with, we highlight an important example of matrices $A$ and $B$ (with $A \neq B$) which satisfy this equation.

Let $A$ and $B$ be two matrices of same size. If $A = B$, it is obvious that $A^2 = B^2$. However, $A^2 = B^2$ is certainly not true in general if $A \neq B$.

For example, let $A$ be an identity matrix and $B$ be a zero matrix. Then, $A^2$ is again an identity matrix and $B^2$ is a zero matrix. In this case, it is easy to see that $A^2 \neq B^2$. Alternatively, the example below gives two matrices $A$ and $B$ such that $A^2 = B^2$, and $A \neq B$.

Example 1. Let $A$ and $B$ be defined as follows:

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

and

$$B = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}.$$  

Then,

$$A^2 = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$B^2 = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$  

2.2 Common Eigenvectors

Following in the direction of the results provided for commuting matrices, the next thing to consider is the eigenvectors of $A$ and $B$. As indicated in Theorem 1.2.1, if $AB = BA$, then the matrices $A$ and $B$ have a common eigenvector. However, this may not be the case if $A^2 = B^2$.

To demonstrate this, consider the matrices given in Example 1. The eigenvectors of $A$ are

$$\left\{ a \begin{bmatrix} 1 \\ -1/2 \end{bmatrix} : a \in \mathbb{R} \right\} \cup \left\{ a \begin{bmatrix} 0 \\ 1 \end{bmatrix} : a \in \mathbb{R} \right\}.$$
The eigenvectors of $B$ are

$$\left\{ b \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} : b \in \mathbb{R} \right\} \cup \left\{ b \begin{bmatrix} 1 \\ 0 \end{bmatrix} : b \in \mathbb{R} \right\}.$$ 

These two sets are disjoint, indicating $A$ and $B$ have no common eigenvectors.

However, if $A$ (or $B$) have many "distinct" eigenvalues (to be more precise, the squares of the eigenvalues should be distinct, that is, $\lambda_i \neq -\lambda_j$ for $\lambda_i$ and $\lambda_j$ to be "distinct"), then it can be shown that $A$ and $B$ have a common eigenvector. The first step is to show that for any eigenvector $u$ of $A$, both $u$ and $B(u)$ are eigenvectors of $B^2$.

**Lemma 2.2.1.** Suppose $A^2 = B^2$. If $\lambda$ and $u$ are an eigenvalue and eigenvector of $A$, then $u$ and $B(u)$ are eigenvectors of $B^2$ corresponding to the eigenvalue $\lambda^2$.

**Proof.** Suppose that $u$ is an eigenvector of $A$ with eigenvalue $\lambda$. By definition, $Au = \lambda u$. Therefore,

$$A^2(u) = A(Au) = A(\lambda u) = \lambda (Au) = \lambda (\lambda u) = \lambda^2(u)$$

Thus, $B^2u = \lambda^2 u$, since $A^2 = B^2$, implying $u$ is an eigenvector of $B^2$ corresponding to the eigenvalue $\lambda^2$. Furthermore,

$$B^2(B(u)) = B(B^2(u)) = B(\lambda^2 u) = \lambda^2(B(u)).$$

Therefore, $Bu$ is an eigenvector of $B^2$ corresponding to the eigenvalue $\lambda^2$. Hence, $u$ and $B(u)$ are eigenvectors of $B^2$ corresponding to the eigenvalue $\lambda^2$. 

This lemma can now be used to show that when there are many eigenvalues of $A$ such that $\lambda_i \neq -\lambda_j$, then there must be a common eigenvector.

**Theorem 2.2.2.** Suppose a matrix $A$ of dimension $n$ has more than $n/2$ distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$ such that $\lambda_i \neq -\lambda_j$ for any $i, j$. If $A^2 = B^2$, then $A$ and $B$ have a common eigenvector.

**Proof.** Let $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_{n/2}, \lambda_{n/2+1}$ be $n/2 + 1$ distinct eigenvalues of $A$. Lemma 2.2.1, implies that there are two eigenvectors of $B^2$ corresponding to each eigenvalue. Since $A$ has more than $n/2$ distinct eigenvalues, applying Lemma 2.2.1 to each eigenvalue yields a list of more than $n$ eigenvectors of $B^2$. Therefore, this list of eigenvectors cannot be linearly independent, and eigenvectors corresponding to distinct eigenvalues must be linearly independent. Thus, for some $i$, the eigenvectors $u_i$ and $B(u_i)$ are linearly dependent, that is, $B(u_i) = ku_i$ for some scalar $k$. Thus, $u_i$ is also an eigenvector of $B$ which proves that $A$ and $B$ have a common eigenvector. 

$\square$
2.2.1 Sharpness

Furthermore, we can show that the condition required on the number of distinct eigenvalues cannot be improved. Recall again the matrices $A$ and $B$ from Example 1. The eigenvalues for $A$ are 1 and $-1$ (as are the eigenvalues for $B$) corresponding to $n/2$ distinct eigenvalues (with $n = 2$) but no common eigenvector. The goal of the construction below is to extend this to larger $n$.

Let us take the matrices $A$ and $B$ from the Example 1. Then,

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

and

$$B = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Now, given the matrix $A$, define $A_k$ as follows:

$$A_k = \begin{bmatrix} A & 0 & 0 & \ldots & 0 \\ 0 & 1/2A & 0 & \ldots & 0 \\ 0 & 0 & 1/3A & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/(k-1)A & 0 \\ 0 & 0 & \ldots & \ldots & 1/kA \end{bmatrix}.$$

Given $B$, define $B_k$ as follows:

$$B_k = \begin{bmatrix} B & 0 & 0 & \ldots & 0 \\ 0 & 1/2B & 0 & \ldots & 0 \\ 0 & 0 & 1/3B & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/(k-1)B & 0 \\ 0 & 0 & \ldots & \ldots & 1/kB \end{bmatrix}.$$

Alternatively, the Kronecker product of two matrices $A$ and $B$ is denoted by $A \otimes B$ and is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \ldots & a_{mn}B \end{bmatrix},$$

where the matrices $A$ and $B$ have the sizes $m \times n$ and $p \times q$ respectively and the size of the product matrix $A \otimes B$ is $mp \times nq$ [7]. More succinctly, $A_k$ can be written as $A_k = C_k \otimes A$ and
$B_k = C_k \otimes B$ where

$$C_k = \begin{bmatrix}
1 & 0 & 0 & \ldots & 0 \\
0 & 1/2 & 0 & \ldots & 0 \\
0 & 0 & 1/3 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1/(k-1) & 0 \\
0 & 0 & \ldots & 1/k & 0
\end{bmatrix}$$

As $A^2 = B^2$, then $A_k^2 = B_k^2$. The eigenvalues of the matrix $C_k$ are just the diagonal entries, namely $1, 1/2, 1/3, \ldots, 1/k$. The eigenvalues for both $A$ and $B$ were determined earlier, and the eigenvalues of both $A$ and $B$ were found to be $1$ and $-1$. Hence, by a property of the Kronecker product of the matrices, the eigenvalues of both the matrices $A_k = C_k \otimes A$ and $B_k = C_k \otimes B$ are $1, -1, 1/2, -1/2, 1/3, -1/3, \ldots, 1/k, -1/k$. Alternatively, this can be shown directly without much difficulty. Consider the eigenvectors of a matrix $X_k$ given by

$$X_k = \begin{bmatrix}
X & 0 & 0 & \ldots & 0 \\
0 & 1/2X & 0 & \ldots & 0 \\
0 & 0 & 1/3X & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1/(k-1)X & 0 \\
0 & 0 & \ldots & 1/kX & 0
\end{bmatrix}$$

where $X$ is a square matrix with eigenvalues of $1$ and $-1$. Then considering the eigenvectors of $X_k$, we get,

$$\begin{bmatrix}
X & 0 & 0 & \ldots & 0 \\
0 & (1/2)X & 0 & \ldots & 0 \\
0 & 0 & (1/3)X & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & (1/k-1)X & 0 \\
0 & 0 & \ldots & 0 & (1/k)X
\end{bmatrix} \begin{bmatrix}
u_1 \\
u_2 \\
u_3 \\
\vdots \\
u_k
\end{bmatrix} = \lambda \begin{bmatrix}
u_1 \\
u_2 \\
u_3 \\
\vdots \\
u_k
\end{bmatrix}$$

This yields the following equations:

$$X u_1 = \lambda u_1 \Rightarrow X u_1 = \lambda u_1$$

$$(1/2)X u_2 = \lambda u_2 \Rightarrow X u_2 = 2 \lambda u_2$$

\vdots

$$(1/k)X u_k = \lambda u_k \Rightarrow X u_k = k \lambda u_k.$$ 

Thus, each $u_i$ is an eigenvector of $X$ or zero, and the eigenvalues of $X$ are $1, -1, 1/2, -1/2, \ldots, 1/k, -1/k$. As there are no common eigenvectors between $A$ and $B$, then $A_k$ and $B_k$ have no common eigenvectors.
Furthermore, both $A_k$ and $B_k$ have $k$ "distinct" eigenvalues. Since $A$ and $B$ are $2 \times 2$ matrices, then the $n \times n$ matrices $A_k$ and $B_k$ have $n = 2k$. So $k = n/2$. This indicates the sharpness of the result.

### 2.3 A Generalization of Theorem 2.2.2

The conclusion of Theorem 2.2.2 holds for a broader class of matrix equations. Consider the matrix equation $A^m = B^m$ for any integer $m \geq 2$. Suppose $A^m = B^m$ and $u$ is an eigenvector of $A$. Then,

\[
A^m(u) = A^{m-1}A(u),
\]

\[
= A^{m-1}(\lambda u),
\]

\[
= \lambda A^{m-1}(u),
\]

\[
= \lambda \lambda A^{m-2}(u),
\]

\[\vdots\]

\[
= \lambda^{m-1}A(u),
\]

\[
= \lambda^m u.
\]

However, since $A^m = B^m$, then $A^m u = \lambda^m u$ implies $B^m u = \lambda^m u$. Hence, $u$ is also an eigenvector of $B^m$ corresponding to the eigenvalue $\lambda^m$. Also,

\[
B^m(B(u)) = B(B^m u)
\]

\[
= B(A^m u)
\]

\[
= B(\lambda^m u)
\]

\[
= \lambda^m (B(u)).
\]

Therefore, if $A^m = B^m$ and $u$ is an eigenvalue of $A$, then the vectors $u$ and $B(u)$ are the eigenvectors of the matrix $B^m$. Thus, as in Theorem 2.2.2, if the matrix $A$ has more than $n/2$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that $\lambda_i \neq -\lambda_j$ for any $i, j$ and $A^m = B^m$, the matrices $A$ and $B$ have common eigenvector.

### 2.4 Unique Square Roots

As noted in Example 1, if $A^2 = B^2$, it may not be the case that $A = B$. In fact, matrices $A$ and $-A$, when squared, give the same matrix $A^2$. However, if solutions are restricted to a particular domain, it may be possible to find unique square roots. For example, a positive definite matrix has a unique square root which is also positive definite [6]. Let $C_n$ denote a
n-cycle in a graph. Then, in the sense of graph theory, the square of any adjacency matrix in
the set below has a unique square root.

\[ D = \{ A : A \text{ is the adjacency matrix of a graph with no } C_3 \text{ or } C_4 \} \]

**Definition 2.4.1.** An **adjacency matrix** is a matrix which represents the relation between
the adjacent vertices of a given graph. Let \( a_{ij} = 1 \) if \( i^{th} \) and \( j^{th} \) vertices are adjacent to each
other and 0 otherwise. For our purpose, we will also set \( a_{ii} = 1 \) for each \( i \) [2].

For example,

\[
A(T_1) = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1
\end{bmatrix}
\]

is an adjacency matrix for the graph in Figure 2.1.

![Figure 2.1: The tree \( T_1 \), which has adjacency matrix \( A(T_1) \)](image)

**Definition 2.4.2.** Let \( G \) be a graph. Then, the **square** of a graph \( G \) is the graph \( G^2 \) such
that \( V(G^2) = V(G) \) and \( u, v \in E(G^2) \) if and only if \( 1 \leq d_G(u, v) \leq 2 \) where \( d_G(u, v) \) is the
distance between the vertices \( u \) and \( v \) [2].

Then the adjacency matrix corresponding to the square of the graph \( T_1 \) from Figure 2.1
is given by

\[
A(T_1^2) = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1
\end{bmatrix}
\]

Multiplying the matrix \( T_1 \) with itself in the usual sense of matrix multiplication, yields
Here, both the matrices $A(T_i^2)$ and $A(T_i)^2$ have zeroes in the same position, and the non-zero elements of both of the matrices represent paths between two vertices on the graph, which are of length 1 or 2.

To start with, some useful results about the squares of the adjacency matrices of graphs are derived.

**Lemma 2.4.1.** If $A(G) = \{a_{ij}\}$ represents the adjacency matrix of a graph $G$ (with ones on the diagonal), then in $A(G)^2 = \{b_{ij}\}$, the diagonal entry $b_{ii}$ will be equal to the degree of vertex plus one.

**Proof.** Let $G$ be a graph and $A(G)$ be its adjacency matrix. For undirected graphs with vertices $v_1, v_2, \ldots, v_n$, the adjacency matrix is always symmetric and the diagonal element of the square of the graph has the form $b_{ii} = a_{i1}a_{1i} + a_{i2}a_{2i} + \ldots + a_{ii}a_{ii} + \ldots + a_{in}a_{ni}$ where $a_{ii}a_{ii} = 1$ because there is 1 on diagonal and $a_{ij}a_{ji} = 1$ if $u_i$ and $v_j$ are adjacent and zero otherwise. In other words,

$$b_{ii} = \sum_{a_{ji} \text{ and } a_{ij} \text{ are adjacent or } v_i = v_j} 1$$

If the entry $a_{ii}$ corresponds to the vertex $v_i$, it is clear that the diagonal element $b_{ii}$ is equal to the degree of vertex $v_i + 1$. Hence, the lemma is proved.  

In fact, restricting the adjacency matrices considered to those corresponding to graphs containing no 4-cycle or triangle has the following important consequence.

**Lemma 2.4.2.** If $G$ contains no 4-cycle or triangle, then for $i \neq j$, $[M(G)]^2_{ij} = 2$ if and only if $v_i$ and $v_j$ are adjacent in $G$.

**Proof.** Suppose $[M(G)]^2_{ij} = 2$ but suppose $v_i$ and $v_j$ are not adjacent. Thus, there must be two disjoint paths of length two from $v_i$ to $v_j$, say $v_i, v_k, v_j$ and $v_i, v'_k, v_j$ which do not use a loop at $v_i$ and $v_j$. However, this would result in a situation shown in Figure 2.2, implying a 4-cycle, so $v_i$ and $v_j$ must be adjacent.

Alternatively, suppose $v_i$ and $v_j$ are adjacent. Then there are two paths of length two by using the loop at $v_i$ (or $v_j$) and the edge between them. If there is any additional path
between $v_i$ and $v_j$, the graph would involve a triangle or a 4-cycle. This can also be seen in Figure 2.2. Hence, $[M(G)]_{ij}^2 = 2$.

\[ \text{Corollary 2.4.3. If } A(G_1)^2 = A(G_2)^2 \text{ where } G_1 \text{ and } G_2 \text{ do not contain any 4-cycle or triangle, then } G_1 = G_2. \]

\[ \text{Proof. Let } G \text{ be a } C_3 \text{-free, } C_4 \text{-free graph and let } M(G) \text{ be the adjacent matrix of } G. \text{ Square the matrix } M(G) \text{ and consider the 2's in the non-diagonal entries of the matrix. By Lemma 2.4.2, there is a 2 on the } i^{th} \text{ row and } j^{th} \text{ column of the square of the matrix if and only if there is 1 in the same position in the adjacency matrix. This means that the vertices } i \text{ and } j \text{ are adjacent to each other. So the square of the adjacency matrix represents a unique } C_4 \text{-free, } C_3 \text{-free graph } G. \text{ Thus, if there are two adjacency matrices arising from graphs containing no 4-cycle or triangle such that the square of their adjacency matrices are equal, then the two adjacency matrices must be equal.} \]

As a special case, all trees satisfy the condition that they do not contain $C_4$ or $C_3$, and the following example illustrates Corollary 2.4.3 in the case of a tree. Let $T$ be the tree shown in Figure 2.3. Then the adjacency matrix corresponding to the tree is the following:

\[ M(T) = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1
\end{bmatrix}. \]

\[ \text{Figure 2.2: Illustration indicating the existence of 4-cycles and 3-cycles} \]
When we square this matrix, we get,

\[
\begin{bmatrix}
2 & 2 & 1 & 1 & 1 & 0 & 0 \\
2 & 5 & 2 & 2 & 2 & 1 & 1 \\
1 & 2 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 2 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 4 & 2 & 2 \\
0 & 1 & 0 & 0 & 2 & 2 & 1 \\
0 & 1 & 0 & 0 & 2 & 1 & 2 \\
\end{bmatrix}
\]

We see that there are five 2's on the diagonal of the matrix \(M(T)^2\). But from the Lemma 2.4.1, we know that the diagonal element of the square matrix is equal to the degree of the vertex plus one. Thus if we have 2 on the \(i^{th}\) row and \(i^{th}\) column in this square matrix, we understand that the \(i^{th}\) vertex of this tree has only one neighbor. Thus it is clear that the given tree has five leaves. Also, the leaves are uniquely associated with their neighborhoods. Therefore, the matrix uniquely represents a tree. Hence, if the square of adjacency matrices of two trees are equal, the tree is unique.

An alternative approach, relying only on the nature of entries as zero or nonzero, can be built on the following lemma.

**Lemma 2.4.4.** For a rooted tree, a leaf at the lowest level has exactly two non-leaf neighborhoods in the square (unless \(T\) is a star).

**Proof.** Let \(T\) be a rooted tree and \(v\) is the vertex at the lowest level of the tree and let \(u\) be its parent. Then the neighbors of \(u\) in \(T\) are the neighbors of \(v\) in the square of \(T\). Let \(w\) be the parent of the vertex \(u\). Then, \(v\) is adjacent to \(u\) and \(w\) in the square; both of which are not leaves. However, all other neighbors of \(u\) are leaves, because a non-leaf child would imply another level. \(\square\)
A graph which demonstrates the property of the lowest level vertex is given in Figure 2.2.

Figure 2.4: An example illustrating Lemma 2.4.4

When the square of the given tree is considered, it can clearly be seen that a leaf at the lowest level has exactly two non-leaf neighborhoods in the square.

Furthermore, if $G$ is not a star or double star, one can determine an induced path $uvw$, which allows use of the following lemma.

**Lemma 2.4.5.** Let $H_1$ and $H_2$ be two graphs on the vertex set $V$. Suppose that $G = H_1^2 = H_2^2$ and that $u, v, w \in V$ are three vertices such that $uvw$ is a path in both $H_1$ and $H_2$. Then, $H_1 = H_2$ [1].

Let $T$ be a tree with the vertex $u$ at its lowest level. Also, let $v$ be the parent of $u$ and child of the vertex $w$. Then $N(v) \subset N(w)$ in the square of the tree $T$. This can be used to determine the order of vertices in the $uvw$ path, unless $N(v) = N(w)$. In this case, the graph must be a double star. If the tree is not a double star, then Lemma 2.4.5 can be used to determine the rest of the adjacency matrix.
Chapter 3

THE MATRIX EQUATIONS $AB = B^2$ AND $AB = A^2$

3.1 An Example for $AB = B^2$, with $A \neq B$

In this chapter, we consider matrices which satisfy the property $AB = B^2$ and $AB = A^2$, or in some cases, matrices corresponding to the operators which satisfy $AB = B^2$ and $AB = A^2$.

Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be two linear mappings. If $AB = B^2$, then the mappings $A$ and $B$ are equal when restricted to the range of $B$. This can be seen clearly by noting that for any $u \in \mathbb{R}^n$,

$$A(B(u)) = B(B(u)).$$

If $B$ is invertible, then $AB = B^2$ implies $A = B$, as $A = ABB^{-1} = B^2B^{-1} = B$. However, if $B$ is not invertible, it may be the case that $A \neq B$. The following construction provides an example illustrating this.

To easily consider the range of $B$, we choose to consider matrices arising from a fairly simple linear operators. Let $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ and $B : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be two linear mappings and choose $B$ to be a linear map defined as below,

- $B(1,0,0,0) = (1,1,1,0) = 1 \cdot (1,0,0,0) - 1 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1)$
- $B(1,1,0,0) = (1,1,1,1) = 0 \cdot (1,0,0,0) + 1 \cdot (1,1,0,0) - 1 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1)$
- $B(1,1,1,0) = (2,1,1,2) = 1 \cdot (1,0,0,0) + 0 \cdot (1,1,0,0) - 1 \cdot (1,1,1,0) + 2 \cdot (1,1,1,1)$
- $B(1,1,1,1) = (0,-1,1,0) = 1 \cdot (1,0,0,0) - 2 \cdot (1,1,0,0) + 1 \cdot (1,1,1,0) + 0 \cdot (1,1,1,1),$

where $\{(1,0,0,0),(1,1,0,0),(1,1,1,0),(1,1,1,1)\}$ forms a basis for $\mathbb{R}^4$. Furthermore,

$$\text{range}(B) = \{a(1,0,1,1) + b(1,1,0,1) : a, b \in \mathbb{R}\}$$

Then the matrix corresponding to the linear mapping $B$ is

$$M(B) = \begin{bmatrix}
1 & 0 & 1 & 1 \\
-1 & 1 & 0 & -2 \\
0 & -1 & -1 & 1 \\
1 & 1 & 2 & 0
\end{bmatrix}.$$
For \( A(B) = B^2 \), it must be the case that \( A \) and \( B \) agree on vectors in the range. For this reason, it is necessary to consider \( B \) as applied to the vectors \((1,0,1,1)\) and \((1,1,0,1)\).

\[
B(1,0,1,1) = (0,-2,2,0) \\
B(1,1,0,1) = (-1,-1,0,-1)
\]

This determines the following values of \( A \).

\[
B(1,0,1,1) = A(1,0,1,1) = (0,-2,2,0) \\
B(1,1,0,1) = A(1,1,0,1) = (-1,-1,0,-1)
\]

As \( A \) is linear, this yields two equations, namely,

\[
A(1,0,1,1) = A(1,0,0,0) - A(1,1,0,0) + A(1,1,1,1) = (0,-2,2,0), \quad \text{and} \\
A(1,1,0,1) = A(1,1,0,0) - A(1,1,1,0) + A(1,1,1,1) = (-1,-1,0,-1).
\]

For this example, we chose \( A(1,1,1,1) = (-1,0,-1,-1) \) arbitrarily. Then, solve for \( A(1,0,0,0), A(1,1,0,0), \) and \( A(1,1,1,0) \), to get,

\[
A(1,0,0,0) = (2,1,2,2) = 1 \cdot (1,0,0,0) - 1 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 2 \cdot (1,1,1,1), \\
A(1,1,0,0) = (1,2,0,1) = -1 \cdot (1,0,0,0) + 2 \cdot (1,1,0,0) - 1 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1), \\
A(1,1,1,0) = (1,2,0,1) = -1 \cdot (1,0,0,0) + 2 \cdot (1,1,0,0) - 1 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1), \\
A(1,1,1,1) = (-1,-1,0,-1)
\]

\[
= 0 \cdot (1,0,0,0) - 1 \cdot (1,1,0,0) + 1 \cdot (1,1,1,0) - 1 \cdot (1,1,1,1).
\]

Thus, the matrix corresponding to the mapping \( A : \mathbb{R}^4 \rightarrow \mathbb{R}^4 \) is given by

\[
\mathcal{M}(A) = \begin{bmatrix}
1 & -1 & -1 & 0 \\
-1 & 2 & 2 & -1 \\
0 & -1 & -1 & 1 \\
2 & 1 & 1 & -1
\end{bmatrix}.
\]

Here, we see that

\[
\mathcal{M}(A)\mathcal{M}(B) = \mathcal{M}(B)^2 = \begin{bmatrix}
2 & 0 & 2 & 2 \\
-4 & -1 & -5 & -3 \\
2 & 1 & 3 & 1 \\
0 & -1 & -1 & 1
\end{bmatrix}.
\]

We note that the matrix \( \mathcal{M}(A) \) is not unique, as the choice of \( A(1,1,1,1) \) was arbitrary. For any noninvertible \( n \times n \) matrix \( B \) over a field \( \mathbb{F} \), the process used to construct this example can easily be applied to find a matrix \( A \) such that \( A \neq B \) but \( A^2 = B^2 \). Again, find the range of \( B \) (with dimension less than \( n \)), and fix \( A = B \) on a basis for this space. We can extend this basis to a basis of \( \mathbb{F}^n \), and choose \( A \) arbitrarily on the additional basis vectors.
3.2 An Example for $AB = A^2$, with $A \neq B$

Now we consider the case when the relation $AB = A^2$ holds where $A$ and $B$ are two matrices of the same dimensions. If the matrix $A$ is invertible, it is quite obvious to see that $A = B$. Therefore, it is necessary to seek a solution of the above relation where $A$ is singular. If the matrix $A$ is a nilpotent matrix of index 2, $A^2 = 0$. If $B$ is zero matrix, then $AB = A^2$. Also, if the matrix $A$ is an idempotent matrix, i.e., $A^2 = A$ and $B$ is an identity matrix, $AB = A^2$.

Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be two linear mappings. Suppose $u \in \text{Nullspace}(A)$, then,

$$A^2(u) = 0 = A(B(u)).$$

Thus, $B(u) \in \text{Nullspace}(A)$. Therefore, $B$ is invariant on the null space of $A$.

To easily consider the null space of $A$, we consider matrices arising from fairly simple linear operators. Let $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ and $B : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ be two linear operators defined as

$$A(1,0,0,0) = (0,0,0,0) = 0 \cdot (1,0,0,0) + 0 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 0 \cdot (1,1,1,1)$$
$$A(1,1,0,0) = (1,0,1,1) = 0 \cdot (1,0,0,0) - 1 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1)$$
$$A(1,1,1,0) = (0,0,0,0) = 0 \cdot (1,0,0,0) + 0 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 0 \cdot (1,1,1,1)$$
$$A(1,1,1,1) = (1,2,1,3) = -1 \cdot (1,0,0,0) + 1 \cdot (1,1,0,0) - 2 \cdot (1,1,1,0) + 3 \cdot (1,1,1,1)$$

and

$$B(1,0,0,0) = (5,3,3,0) = 2 \cdot (1,0,0,0) + 0 \cdot (1,1,0,0) + 3 \cdot (1,1,1,0) + 0 \cdot (1,1,1,1)$$
$$B(1,1,0,0) = (1,0,1,1) = 0 \cdot (1,0,0,0) - 1 \cdot (1,1,0,0) + 0 \cdot (1,1,1,0) + 1 \cdot (1,1,1,1)$$
$$B(1,1,1,0) = (10,4,4,0) = 6 \cdot (1,0,0,0) + 0 \cdot (1,1,0,0) + 4 \cdot (1,1,1,0) + 0 \cdot (1,1,1,1)$$
$$B(1,1,1,1) = (1,2,1,3) = -1 \cdot (1,0,0,0) + 1 \cdot (1,1,0,0) - 2 \cdot (1,1,1,0) + 3 \cdot (1,1,1,1),$$

so that $B$ is invariant on the null space of $A$. Then the matrices associated with the linear operators $A$ and $B$ are given by

$$\mathcal{M}(A) = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & 3 \end{bmatrix}$$

and

$$\mathcal{M}(B) = \begin{bmatrix} 2 & 0 & 6 & -1 \\ 0 & -1 & 0 & 1 \\ 3 & 0 & 4 & -2 \\ 0 & 1 & 0 & 3 \end{bmatrix}.$$
Multiplying, it is the case that
\[
M(A)M(B) = M(A^2) = \begin{bmatrix} 0 & -1 & 0 & -3 \\ 0 & 2 & 0 & 2 \\ 0 & -2 & 0 & -6 \\ 0 & 2 & 0 & 10 \end{bmatrix}.
\]

### 3.3 Matrix Commutators

Let \( A \) and \( B \) be two matrices of size \( n \times n \) with elements from a field \( \mathbb{F} \). The matrix commutator of the matrices \( A \) and \( B \) is denoted by \([A,B]\) and is defined as \([A,B] = AB - BA\).

As discussed previously, \( Tr(AB) = Tr(BA) \) for any matrices \( A \) and \( B \), the trace of the matrix commutator is always zero. Also, if \( Tr(A) = 0 \), the matrix \( A \) must be a commutator of two matrices.

If the matrix commutator of the matrices \( A \) and \( B \) is a zero matrix i.e., \([A,B] = 0\), the matrices \( A \) and \( B \) commute. A generalization of commutative matrices is quasi-commutative matrices. Two matrices \( A \) and \( B \) are said to be quasi-commutative matrices if \([A,B] \neq 0\) and both of these matrices commute with the commutator \([A,B]\). For quasi-commutative matrices, \([A,B]\) is necessarily nilpotent.

To define the concept of \( k \)-commutativity, write the commutator \([A,B]\) of the matrices \( A \) and \( B \) as \([A,B]_1\). Then, define \([A,B]_i = [A,[A,B]_{i-1}]\). The matrix \( B \) is said to be \( k \)-commutative with respect to \( A \) if \([A,B]_k = 0\).

The matrix \( B_k = A^kB - C(k,1)A^{k-1}BA + C(k,2)A^{k-2}BA^2 - \ldots + (-1)^kBA^k \) is the \( k \)-th commute of \( A \) with respect to \( B \). Write \( B \) as \( B_0 \). Then, in general, \( B_{i+1} = AB_i - B_iA \) for \( i = 0, 1, 2, \ldots \) [10]. Two matrices are said to be mutually \( k \)-commutative if \( A \) is \( k \)-commutative with respect to \( B \) and \( B \) is \( k \)-commutative with respect to \( A \). If the matrices \( A \) and \( B \) are commutative in the usual sense, i.e., \( AB = BA \), then they are mutually one-commutative.

Regarding quasi-commutative matrices, McCoy [9] was able to use the following theorem to show that quasi-commutative matrices have a common eigenvector.

**Theorem 3.3.1.** For every polynomial \( p(x,y) \) of the non-commutative variables \( x,y \), the matrix \( p(A,B)[A,B] \) is nilpotent if and only if the matrices \( A \) and \( B \) can be simultaneously triangularised by a non-singular matrix. Consequently, if the McCoy condition holds, then \( A \) and \( B \) have a common eigenvector [11].

Theorem 3.3.1 is used in the next section to imply common eigenvectors when \( AB = A^2 \).
3.4 Common Eigenvectors

According to McCoy, the matrices $A$ and $B$ can be simultaneously triangularised if and only if the matrix $p(A,B)[A,B]$ is nilpotent for any $p(A,B)$ where $p(A,B)$ is a non-commuting polynomial in the matrices $A$ and $B$. If the matrix $p(A,B)[A,B]$ is nilpotent for every $p[A,B]$, then the matrices $A$ and $B$ have a common eigenvector.

It can be shown that this must be the case when $AB = A^2$.

**Theorem 3.4.1.** If $AB = A^2$, then $A$ and $B$ have a common eigenvector.

**Proof.** Suppose $AB = A^2$. Then,

$$AB^i = A^iB = A^{i+1}.$$ (3.1)

Let $p(A,B)$ be any polynomial in non-commutative variables. Using Equation 3.1, $p(A,B)$ can be simplified substantially to the form $p(A,B) = \sum a_{ij}B^iA^j$. As a result,

$$p(A,B)[A,B] = (\sum a_{ij}B^iA^j)(AB - BA),$$

$$= (\sum a_{ij}B^iA^j)(A^2 - BA),$$

$$= (\sum a_{ij}B^iA^j)(A - B)A,$n

$$= (\sum a_{ij}B^iA^jA - \sum a_{ij}B^iA^jB)A,$n

$$= (\sum_{j>1} a_{ij}(B^iA^{j+1} - B^iA^{j+1}) + \sum_{j=0} a_{i0}(B^iA - B^{i+1})A,$n

$$= \sum a_i(B^iA - B^{i+1})A,$n

$$= (\sum a_iB^i)(A - B)A.$n

Thus, $p(A,B)[A,B] = f(B)[A,B]$ for any polynomial $p(A,B)$. To show that $p(A,B)[A,B]$ is nilpotent, consider the square.

$$(p(A,B)[A,B])^2 = (p(A,B)(AB - BA))^2$$

$$= (f(B)(A^2 - BA))^2$$

$$= (f(B)(A - B)(A))(f(B)(A - B)A)$$

$$= (f(B)(A - B))(A \cdot f(B)(A - B))A$$

Thus,

$$p(A,B)[A,B]^2 = (f(B)(A - B))(A \cdot f(B)(A - B))A.$$ (3.2)

Let $f(B) = a_0 + a_1B + a_2B^2 + ... + a_nB^n$. Then, we have,

$$A \cdot f(B) = A(a_0 + a_1B + a_2B^2 + ... + a_nB^n),$$

$$= a_0A + a_1AB + a_2AB^2 + ... + a_nAB^n,$n

$$= a_0A + a_1A^2 + a_2A^3 + ... + a_nA^{n+1}.$$
Thus,

\[ A \cdot f(B)(A - B) = (a_0A + a_1A^2 + \ldots + a_nA^{n+1})A - (a_0A + a_1A^2 + \ldots + a_nA^{n+1})B, \]

\[ = (a_0A^2 + a_1A^3 + \ldots + a_nA^{n+2}) - (a_0AB + a_1A^2B + \ldots + a_nA^{n+1}B), \]

\[ = (a_0A^2 + a_1A^3 + \ldots + a_nA^{n+2}) - (a_0A^2 + a_1A^3 + \ldots + a_nA^{n+2}), \]

\[ = 0. \]

Now, the relation Equation 3.2 becomes

\[ (p(A,B)[A,B])^2 = (f(B)(A - B))(A \cdot f(B)(A - B))A, \]

\[ = (f(B)(A - B)) \cdot 0 \cdot A, \]

\[ = 0. \]

Therefore, the matrix \( p(A,B)[A,B] \) is nilpotent with index 2 if \( A^2 = AB \). According to the McCoy condition, if \( p(A,B)[A,B] \) is nilpotent for any polynomial, the matrices \( A \) and \( B \) have a common eigenvector. Therefore, if \( A^2 = AB \), then the matrices \( A \) and \( B \) have a common eigenvector. \( \square \)
Chapter 4

CONCLUSION

4.1 Results

In this thesis, we attempted to find properties of solutions of matrix equations of two matrices taken two at a time.

In the first chapter, we introduced the topic by providing motivation in the form of matrices that commute and the Yang-Baxter equation. Regarding matrices that commute, we noted some special families of matrices that satisfy the commutative property. We also found the solution of the matrix equation \( AB = BA \) by direct method when the size of the matrix \( A \) is \( 2 \times 2 \). Furthermore, we considered some theorems which describe the properties of commuting matrices, especially regarding common eigenvectors. The Yang-Baxter equation provided motivation as our results deal with two matrices taken two at a time, a simplification of two matrices taken three at a time, of which the Yang-Baxter equation is a special case.

In the second chapter, we noted that \( A^2 = B^2 \) does not imply \( A = B \) by the use of an example. We considered the relation \( A^2 = B^2 \) as related to adjacency matrices in graphs and determine conditions for the unique solution of \( A^2 = B^2 \). In the general case, we were able to show that two \( n \times n \) matrices \( A \) and \( B \) share a common eigenvector, if there are more than \( n/2 \) distinct eigenvalues whose squares are also distinct. If fact, the bound of \( n/2 \) provided in Theorem 2.2.2 was shown to be sharp by using the matrices \( A_k \) and \( B_k \). These matrices have \( n/2 \) distinct eigenvectors whose squares are also distinct, however, they have no common eigenvectors.

In the third chapter, we considered the matrix equation \( AB = B^2 \) and \( AB = A^2 \). We provided one example of two distinct matrices \( A \) and \( B \) such that \( AB = B^2 \) and \( AB = A^2 \). Also, we considered the matrix commutator of two matrices and used the theorem of McCoy to show that the matrices satisfying \( AB = A^2 \) have a common eigenvector.

4.2 Future Work

One possible avenue for future work is to consider matrix equations from two matrices \( (A \text{ or } B) \) taken three at a time. We can find many matrix equations if we take two matrices \( A \) and \( B \) taken three at a time (see Table 4.2). But solutions of such equations are not
\[ AAA = BBB \quad AAB = BBA \]
\[ ABA = BAB \quad ABB = BAA \]
\[ AAA = AAB \quad AAA = ABA \]
\[ AAA = ABB \quad AAA = BAA \]
\[ AAA = BAB \quad AAA = BBA \]
\[ AAB = ABA \quad AAB = ABB \]
\[ AAB = BAA \quad AAB = BAB \]
\[ ABA = ABB \quad ABA = BAA \]

Table 4.1: The unique matrix equations of two matrices taken three at a time

necessarily easy to find or describe. One such equation is \( ABA = BAB \) which is the matrix form of the Yang-Baxter equation.

Another possible avenue for future inquiry is to reconsider the equation \( A^m = B^m \) for two \( n \times n \) matrices \( A \) and \( B \). We determined that more than \( n/2 \) eigenvalues such that \( \lambda_i \neq -\lambda_j \) will force a common eigenvector in \( A \) and \( B \) if \( A^m = B^m \). This is sharp when \( m = 2 \), but this may not be the case for \( m > 2 \). In fact, there is reason to believe that the value of \( n/2 \) is not sharp for \( m > 2 \).


