List of Figures

1.1 Gauss–Jordan algorithm .............................. 10
2.1 Reflection in a line ................................. 29
2.2 Projection on a line ................................. 30
4.1 Area of triangle $\mathbf{OPQ}$ ......................... 72
5.1 Complex addition and subtraction .................... 96
5.2 Complex conjugate ................................ 97
5.3 Modulus of a complex number ....................... 99
5.4 Apollonius circles .................................. 101
5.5 Argument of a complex number ...................... 104
5.6 Argument examples ................................ 105
5.7 The $n$th roots of unity .............................. 108
5.8 The roots of $z^n = a$ ................................ 109
6.1 Rotating the axes ................................... 116
7.1 An ellipse example .................................. 135
7.2 Ellipse: standard form ............................... 137
7.3 Hyperbola: standard forms ......................... 138
7.4 Parabola: standard forms (i) and (ii) .............. 138
7.5 Parabola: standard forms (iii) and (iv) .......... 139
7.6 1st parabola example ................................ 140
7.7 2nd parabola example ................................ 141
8.1 Equality and addition of vectors .................... 150
8.2 Scalar multiplication of vectors ..................... 151
8.3 Representation of three–dimensional space .......... 155
8.4 The vector $\overrightarrow{AB}$ .......................... 155
8.5 The negative of a vector ............................ 157
8.6 (a) Equality of vectors; (b) Addition and subtraction of vectors. 157
8.7 Position vector as a linear combination of \( \mathbf{i}, \mathbf{j}, \text{ and } \mathbf{k} \). 158
8.8 Representation of a line. 162
8.9 The line \( \mathbf{AB} \). 162
8.10 The cosine rule for a triangle. 167
8.11 Pythagoras’ theorem for a right–angled triangle. 168
8.12 Distance from a point to a line. 169
8.13 Projecting a segment onto a line. 171
8.14 The vector cross–product. 174
8.15 Vector equation for the plane \( \mathbf{ABC} \). 177
8.16 Normal equation of the plane \( \mathbf{ABC} \). 178
8.17 The plane \( ax + by + cz = d \). 179
8.18 Line of intersection of two planes. 182
8.19 Distance from a point to the plane \( ax + by + cz = d \). 184
1.1 Introduction to linear equations

A linear equation in $n$ unknowns $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n, b$ are given real numbers.

For example, with $x$ and $y$ instead of $x_1$ and $x_2$, the linear equation $2x + 3y = 6$ describes the line passing through the points $(3, 0)$ and $(0, 2)$.

Similarly, with $x, y$ and $z$ instead of $x_1, x_2$ and $x_3$, the linear equation $2x + 3y + 4z = 12$ describes the plane passing through the points $(6, 0, 0), (0, 4, 0), (0, 0, 3)$.

A system of $m$ linear equations in $n$ unknowns $x_1, x_2, \ldots, x_n$ is a family of linear equations

$$
\begin{align*}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
& \quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{align*}
$$

We wish to determine if such a system has a solution, that is to find out if there exist numbers $x_1, x_2, \ldots, x_n$ which satisfy each of the equations simultaneously. We say that the system is consistent if it has a solution. Otherwise the system is called inconsistent.
Note that the above system can be written concisely as
\[ \sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m. \]

The matrix
\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]
is called the coefficient matrix of the system, while the matrix
\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
  a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
  \vdots & \vdots & & \vdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
\end{bmatrix}
\]
is called the augmented matrix of the system.

Geometrically, solving a system of linear equations in two (or three) unknowns is equivalent to determining whether or not a family of lines (or planes) has a common point of intersection.

**EXAMPLE 1.1.1** Solve the equation
\[ 2x + 3y = 6. \]

**Solution.** The equation \( 2x + 3y = 6 \) is equivalent to \( 2x = 6 - 3y \) or \( x = 3 - \frac{3}{2}y \), where \( y \) is arbitrary. So there are infinitely many solutions.

**EXAMPLE 1.1.2** Solve the system
\[
\begin{align*}
  x + y + z &= 1 \\
  x - y + z &= 0.
\end{align*}
\]

**Solution.** We subtract the second equation from the first, to get \( 2y = 1 \) and \( y = \frac{1}{2} \). Then \( x = y - z = \frac{1}{2} - z \), where \( z \) is arbitrary. Again there are infinitely many solutions.

**EXAMPLE 1.1.3** Find a polynomial of the form \( y = a_0 + a_1x + a_2x^2 + a_3x^3 \) which passes through the points \((-3, -2), (-1, 2), (1, 5), (2, 1)\).
Solution. When \( x \) has the values \(-3, -1, 1, 2\), then \( y \) takes corresponding values \(-2, 2, 5, 1\) and we get four equations in the unknowns \( a_0, a_1, a_2, a_3 \):

\[
\begin{align*}
a_0 - 3a_1 + 9a_2 - 27a_3 &= -2 \\
a_0 - a_1 + a_2 - a_3 &= 2 \\
a_0 + a_1 + a_2 + a_3 &= 5 \\
a_0 + 2a_1 + 4a_2 + 8a_3 &= 1.
\end{align*}
\]

This system has the unique solution \( a_0 = 93/20, a_1 = 221/120, a_2 = -23/20, a_3 = -41/120 \). So the required polynomial is

\[
y = \frac{93}{20} + \frac{221}{120}x - \frac{23}{20}x^2 - \frac{41}{120}x^3.
\]

In [26, pages 33–35] there are examples of systems of linear equations which arise from simple electrical networks using Kirchhoff’s laws for electrical circuits.

Solving a system consisting of a single linear equation is easy. However if we are dealing with two or more equations, it is desirable to have a systematic method of determining if the system is consistent and to find all solutions.

Instead of restricting ourselves to linear equations with rational or real coefficients, our theory goes over to the more general case where the coefficients belong to an arbitrary field. A field \( F \) is a set \( F \) which possesses operations of addition and multiplication which satisfy the familiar rules of rational arithmetic. There are ten basic properties that a field must have:

**THE FIELD AXIOMS.**

1. \((a + b) + c = a + (b + c)\) for all \( a, b, c \) in \( F \);
2. \((ab)c = a(bc)\) for all \( a, b, c \) in \( F \);
3. \( a + b = b + a \) for all \( a, b \) in \( F \);
4. \( ab = ba \) for all \( a, b \) in \( F \);
5. there exists an element \( 0 \) in \( F \) such that \( 0 + a = a \) for all \( a \) in \( F \);
6. there exists an element \( 1 \) in \( F \) such that \( 1a = a \) for all \( a \) in \( F \);
7. to every \( a \) in \( F \), there corresponds an additive inverse \(-a\) in \( F \), satisfying
\[ a + (-a) = 0; \]

8. to every non-zero \( a \) in \( F \), there corresponds a multiplicative inverse \( a^{-1} \) in \( F \), satisfying
\[ aa^{-1} = 1; \]

9. \( a(b + c) = ab + ac \) for all \( a, b, c \) in \( F \);

10. \( 0 \neq 1 \).

With standard definitions such as \( a - b = a + (-b) \) and \( \frac{a}{b} = ab^{-1} \) for \( b \neq 0 \), we have the following familiar rules:

\[
\begin{align*}
-(a + b) &= (-a) + (-b), \quad (ab)^{-1} = a^{-1}b^{-1}; \\
-(-a) &= a, \quad (a^{-1})^{-1} = a; \\
-(a - b) &= b - a, \quad \left(\frac{a}{b}\right)^{-1} = \frac{b}{a}; \\
\frac{a + c}{b + d} &= \frac{ad + bc}{bd}; \\
\frac{ac}{bd} &= \frac{ac}{bd}; \\
\frac{ab}{ac} &= \frac{b}{c}, \quad \left(\frac{b}{c}\right)^{-1} = \frac{b}{a}; \\
-(ab) &= (-a)b = a(-b); \\
-\left(\frac{a}{b}\right) &= -\frac{a}{b} = \frac{a}{-b}; \\
0a &= 0; \\
(-a)^{-1} &= -(a^{-1}).
\end{align*}
\]

Fields which have only finitely many elements are of great interest in many parts of mathematics and its applications, for example to coding theory. It is easy to construct fields containing exactly \( p \) elements, where \( p \) is a prime number. First we must explain the idea of modular addition and modular multiplication. If \( a \) is an integer, we define \( a \) (mod \( p \)) to be the least remainder on dividing \( a \) by \( p \). That is, if \( a = bp + r \), where \( b \) and \( r \) are integers and \( 0 \leq r < p \), then \( a \) (mod \( p \)) = \( r \).

For example, \(-1 \) (mod 2) = 1, 3 (mod 3) = 0, 5 (mod 3) = 2.
1.1. **INTRODUCTION TO LINEAR EQUATIONS**

Then addition and multiplication mod $p$ are defined by

\[
\begin{align*}
\oplus & : a \oplus b = (a + b) \pmod{p} \\
\otimes & : a \otimes b = (ab) \pmod{p}.
\end{align*}
\]

For example, with $p = 7$, we have $3 \oplus 4 = 7 \pmod{7} = 0$ and $3 \otimes 5 = 15 \pmod{7} = 1$. Here are the complete addition and multiplication tables mod 7:

<table>
<thead>
<tr>
<th>$\oplus$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\otimes$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

If we now let $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, then it can be proved that $\mathbb{Z}_p$ forms a field under the operations of modular addition and multiplication mod $p$. For example, the additive inverse of 3 in $\mathbb{Z}_7$ is 4, so we write $-3 = 4$ when calculating in $\mathbb{Z}_7$. Also the multiplicative inverse of 3 in $\mathbb{Z}_7$ is 5, so we write $3^{-1} = 5$ when calculating in $\mathbb{Z}_7$.

In practice, we write $a \oplus b$ and $a \otimes b$ as $a + b$ and $ab$ or $a \times b$ when dealing with linear equations over $\mathbb{Z}_p$.

The simplest field is $\mathbb{Z}_2$, which consists of two elements 0, 1 with addition satisfying $1 + 1 = 0$. So in $\mathbb{Z}_2$, $-1 = 1$ and the arithmetic involved in solving equations over $\mathbb{Z}_2$ is very simple.

**EXAMPLE 1.1.4** Solve the following system over $\mathbb{Z}_2$:

\[
\begin{align*}
x + y + z &= 0 \\
x + z &= 1.
\end{align*}
\]

**Solution.** We add the first equation to the second to get $y = 1$. Then $x = 1 - z = 1 + z$, with $z$ arbitrary. Hence the solutions are $(x, y, z) = (1, 1, 0)$ and $(0, 1, 1)$.

We use $\mathbb{Q}$ and $\mathbb{R}$ to denote the fields of rational and real numbers, respectively. Unless otherwise stated, the field used will be $\mathbb{Q}$.
1.2 Solving linear equations

We show how to solve any system of linear equations over an arbitrary field, using the *GAUSS–JORDAN* algorithm. We first need to define some terms.

**DEFINITION 1.2.1 (Row–echelon form)** A matrix is in *row–echelon form* if

1. all zero rows (if any) are at the bottom of the matrix and
2. if two successive rows are non–zero, the second row starts with more zeros than the first (moving from left to right).

For example, the matrix

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

is in row–echelon form, whereas the matrix

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

is not in row–echelon form.

The *zero* matrix of any size is always in row–echelon form.

**DEFINITION 1.2.2 (Reduced row–echelon form)** A matrix is in *reduced row–echelon form* if

1. it is in row–echelon form,
2. the leading (leftmost non–zero) entry in each non–zero row is 1,
3. all other elements of the column in which the leading entry 1 occurs are zeros.

For example the matrices

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
0 & 1 & 2 & 0 & 0 & 2 \\
0 & 0 & 1 & 0 & 3 \\
0 & 0 & 0 & 1 & 4 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
are in reduced row–echelon form, whereas the matrices
\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 2
\end{bmatrix}
\quad \text{and} \quad 
\begin{bmatrix}
1 & 2 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]
are not in reduced row–echelon form, but are in row–echelon form.

The zero matrix of any size is always in reduced row–echelon form.

**Notation.** If a matrix is in reduced row–echelon form, it is useful to denote the column numbers in which the leading entries 1 occur, by \(c_1, c_2, \ldots, c_r\), with the remaining column numbers being denoted by \(c_{r+1}, \ldots, c_n\), where \(r\) is the number of non–zero rows. For example, in the 4 \(\times\) 6 matrix above, we have \(r = 3\), \(c_1 = 2\), \(c_2 = 4\), \(c_3 = 5\), \(c_4 = 1\), \(c_5 = 3\), \(c_6 = 6\).

The following operations are the ones used on systems of linear equations and do not change the solutions.

**DEFINITION 1.2.3 (Elementary row operations)** There are three types of elementary row operations that can be performed on matrices:

1. Interchanging two rows:
   \(R_i \leftrightarrow R_j\) interchanges rows \(i\) and \(j\).

2. Multiplying a row by a non–zero scalar:
   \(R_i \rightarrow tR_i\) multiplies row \(i\) by the non–zero scalar \(t\).

3. Adding a multiple of one row to another row:
   \(R_j \rightarrow R_j + tR_i\) adds \(t\) times row \(i\) to row \(j\).

**DEFINITION 1.2.4 (Row equivalence)** Matrix \(A\) is row–equivalent to matrix \(B\) if \(B\) is obtained from \(A\) by a sequence of elementary row operations.

**EXAMPLE 1.2.1** Working from left to right,
\[
A = \begin{bmatrix}
1 & 2 & 0 \\
2 & 1 & 1 \\
1 & -1 & 2
\end{bmatrix}
\quad \Rightarrow \quad 
R_2 \rightarrow R_2 + 2R_3 \quad \begin{bmatrix}
1 & 2 & 0 \\
4 & -1 & 5 \\
1 & -1 & 2
\end{bmatrix} \\
R_2 \leftrightarrow R_3 \quad \begin{bmatrix}
1 & 2 & 0 \\
1 & -1 & 2 \\
4 & -1 & 5
\end{bmatrix}
\quad \Rightarrow \quad 
R_1 \rightarrow 2R_1 \quad \begin{bmatrix}
2 & 4 & 0 \\
1 & -1 & 2 \\
4 & -1 & 5
\end{bmatrix} = B.
\]
Thus \( A \) is row-equivalent to \( B \). Clearly \( B \) is also row-equivalent to \( A \), by performing the inverse row-operations \( R_1 \rightarrow \frac{1}{2} R_1, \ R_2 \leftrightarrow R_3, \ R_2 \rightarrow R_2 - 2R_3 \) on \( B \).

It is not difficult to prove that if \( A \) and \( B \) are row-equivalent augmented matrices of two systems of linear equations, then the two systems have the same solution sets – a solution of the one system is a solution of the other. For example the systems whose augmented matrices are \( A \) and \( B \) in the above example are respectively

\[
\begin{align*}
\begin{cases}
x + 2y &= 0 \\
2x + y &= 1 \\
x - y &= 2
\end{cases}
\quad \text{and} \quad
\begin{cases}
2x + 4y &= 0 \\
x - y &= 2 \\
4x - y &= 5
\end{cases}
\end{align*}
\]

and these systems have precisely the same solutions.

### 1.3 The Gauss–Jordan algorithm

We now describe the **Gauss–Jordan Algorithm**. This is a process which starts with a given matrix \( A \) and produces a matrix \( B \) in reduced row-echelon form, which is row-equivalent to \( A \). If \( A \) is the augmented matrix of a system of linear equations, then \( B \) will be a much simpler matrix than \( A \) from which the consistency or inconsistency of the corresponding system is immediately apparent and in fact the complete solution of the system can be read off.

**STEP 1.**

Find the first non-zero column moving from left to right, (column \( c_1 \)) and select a non-zero entry from this column. By interchanging rows, if necessary, ensure that the first entry in this column is non-zero. Multiply row 1 by the multiplicative inverse of \( a_{1c_1} \), thereby converting \( a_{1c_1} \) to 1. For each non-zero element \( a_{ic_1}, i > 1 \), (if any) in column \( c_1 \), add \(-a_{ic_1}\) times row 1 to row \( i \), thereby ensuring that all elements in column \( c_1 \), apart from the first, are zero.

**STEP 2.** If the matrix obtained at Step 1 has its 2nd, \ldots, \( m \)th rows all zero, the matrix is in reduced row-echelon form. Otherwise suppose that the first column which has a non-zero element in the rows below the first is column \( c_2 \). Then \( c_1 < c_2 \). By interchanging rows below the first, if necessary, ensure that \( a_{2c_2} \) is non-zero. Then convert \( a_{2c_2} \) to 1 and by adding suitable multiples of row 2 to the remaining rows, where necessary, ensure that all remaining elements in column \( c_2 \) are zero.
1.4 SYSTEMATIC SOLUTION OF LINEAR SYSTEMS.

The process is repeated and will eventually stop after \( r \) steps, either because we run out of rows, or because we run out of non–zero columns. In general, the final matrix will be in reduced row–echelon form and will have \( r \) non–zero rows, with leading entries 1 in columns \( c_1, \ldots, c_r \), respectively.

**EXAMPLE 1.3.1**

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

\( R_1 \leftrightarrow R_2 \)

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

\( R_1 \rightarrow \frac{1}{2}R_1 \)

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

\( R_3 \rightarrow R_3 - 5R_1 \)

\[
\begin{bmatrix}
1 & 1 & -1 & \frac{5}{2} \\
0 & 0 & 4 & 0 \\
0 & 0 & 4 & -\frac{15}{2} \\
\end{bmatrix}
\]

\( R_2 \rightarrow \frac{1}{4}R_2 \)

\[
\begin{bmatrix}
1 & 1 & -1 & \frac{5}{2} \\
0 & 0 & 1 & 0 \\
0 & 0 & 4 & -\frac{15}{2} \\
\end{bmatrix}
\]

\( R_1 \rightarrow R_1 + R_2 \)

\[
\begin{bmatrix}
1 & 1 & 0 & \frac{5}{2} \\
0 & 0 & 1 & 0 \\
0 & 0 & 4 & -\frac{15}{2} \\
\end{bmatrix}
\]

\( R_3 \rightarrow R_3 - 4R_2 \)

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

The last matrix is in reduced row–echelon form.

**REMARK 1.3.1** It is possible to show that a given matrix over an arbitrary field is row–equivalent to precisely one matrix which is in reduced row–echelon form.

A flow–chart for the Gauss–Jordan algorithm, based on [1, page 83] is presented in figure 1.1 below.

1.4 Systematic solution of linear systems.

Suppose a system of \( m \) linear equations in \( n \) unknowns \( x_1, \ldots, x_n \) has augmented matrix \( A \) and that \( A \) is row–equivalent to a matrix \( B \) which is in reduced row–echelon form, via the Gauss–Jordan algorithm. Then \( A \) and \( B \) are \( m \times (n + 1) \). Suppose that \( B \) has \( r \) non–zero rows and that the leading entry 1 in row \( i \) occurs in column number \( c_i \), for \( 1 \leq i \leq r \). Then

\[
1 \leq c_1 < c_2 < \cdots < c_r \leq n + 1.
\]
CHAPTER 1. LINEAR EQUATIONS

START

Input A, m, n

i = 1, j = 1

Are the elements in the jth column on and below the ith row all zero?

Yes

No

Let \( a_{pj} \) be the first non-zero element in column j on or below the ith row

Is \( p = i \)?

Yes

No

Interchange the \( p \)th and \( i \)th rows

Divide the \( i \)th row by \( a_{ij} \)

Subtract \( a_{qj} \) times the \( i \)th row from the \( q \)th row for \( q = 1, \ldots, m (q \neq i) \)

Set \( c_i = j \)

\( i = i + 1 \)

\( j = j + 1 \)

Yes

STOP

No

Is \( j = n \)?

Yes

No

Print \( A, c_1, \ldots, c_i \)

Figure 1.1: Gauss–Jordan algorithm.
Also assume that the remaining column numbers are $c_{r+1}, \ldots, c_{n+1}$, where

$$1 \leq c_{r+1} < c_{r+2} < \cdots < c_n \leq n + 1.$$

Case 1: $c_r = n + 1$. The system is inconsistent. For the last non–zero row of $B$ is $[0, 0, \cdots, 1]$ and the corresponding equation is

$$0x_1 + 0x_2 + \cdots + 0x_n = 1,$$

which has no solutions. Consequently the original system has no solutions.

Case 2: $c_r \leq n$. The system of equations corresponding to the non–zero rows of $B$ is consistent. First notice that $r \leq n$ here.

If $r = n$, then $c_1 = 1, c_2 = 2, \cdots, c_n = n$ and

$$B = \begin{bmatrix}
1 & 0 & \cdots & 0 & d_1 \\
0 & 1 & \cdots & 0 & d_2 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 1 & d_n \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}.$$

There is a unique solution $x_1 = d_1, x_2 = d_2, \cdots, x_n = d_n$.

If $r < n$, there will be more than one solution (infinitely many if the field is infinite). For all solutions are obtained by taking the unknowns $x_{c_1}, \cdots, x_{c_r}$ as dependent unknowns and using the $r$ equations corresponding to the non–zero rows of $B$ to express these unknowns in terms of the remaining independent unknowns $x_{c_{r+1}}, \ldots, x_{c_n}$, which can take on arbitrary values:

$$x_{c_1} = b_{1n+1} - b_{1c_{r+1}}x_{c_{r+1}} - \cdots - b_{1c_n}x_{c_n}$$

$$\vdots$$

$$x_{c_r} = b_{rn+1} - b_{rc_{r+1}}x_{c_{r+1}} - \cdots - b_{rc_n}x_{c_n}.$$

In particular, taking $x_{c_{r+1}} = 0, \ldots, x_{c_{n-1}} = 0$ and $x_{c_n} = 0, 1$ respectively, produces at least two solutions.

**EXAMPLE 1.4.1** Solve the system

$$x + y = 0$$

$$x - y = 1$$

$$4x + 2y = 1.$$
CHAPTER 1. LINEAR EQUATIONS

Solution. The augmented matrix of the system is

\[ A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \]

which is row equivalent to

\[ B = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}. \]

We read off the unique solution \( x = \frac{1}{2}, y = -\frac{1}{2}. \)

(Here \( n = 2, r = 2, c_1 = 1, c_2 = 2. \) Also \( c_r = c_2 = 2 < 3 = n + 1 \) and \( r = n. \))

EXAMPLE 1.4.2 Solve the system

\[
\begin{align*}
2x_1 + 2x_2 - 2x_3 &= 5 \\
7x_1 + 7x_2 + x_3 &= 10 \\
5x_1 + 5x_2 - x_3 &= 5.
\end{align*}
\]

Solution. The augmented matrix is

\[ A = \begin{bmatrix} 2 & 2 & -2 & 5 \\ 7 & 7 & 1 & 10 \\ 5 & 5 & -1 & 5 \end{bmatrix} \]

which is row equivalent to

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

We read off inconsistency for the original system.

(Here \( n = 3, r = 3, c_1 = 1, c_2 = 3. \) Also \( c_r = c_3 = 4 = n + 1. \))

EXAMPLE 1.4.3 Solve the system

\[
\begin{align*}
x_1 - x_2 + x_3 &= 1 \\
x_1 + x_2 - x_3 &= 2.
\end{align*}
\]


1.4. SYSTEMATIC SOLUTION OF LINEAR SYSTEMS.

Solution. The augmented matrix is
\[
A = \begin{bmatrix}
1 & -1 & 1 & 1 \\
1 & 1 & -1 & 2 \\
\end{bmatrix}
\]
which is row equivalent to
\[
B = \begin{bmatrix}
1 & 0 & 0 & 3/2 \\
0 & 1 & -1 & 1/2 \\
\end{bmatrix}.
\]

The complete solution is \(x_1 = \frac{3}{2}, x_2 = \frac{1}{2} + x_3\), with \(x_3\) arbitrary.
(Here \(n = 3, r = 2, c_1 = 1, c_2 = 2\). Also \(c_r = c_2 = 2 < 4 = n + 1\) and \(r < n\).)

EXAMPLE 1.4.4 Solve the system
\[
\begin{align*}
6x_3 + 2x_4 - 4x_5 - 8x_6 &= 8 \\
3x_3 + x_4 - 2x_5 - 4x_6 &= 4 \\
2x_1 - 3x_2 + x_3 + 4x_4 - 7x_5 + x_6 &= 2 \\
6x_1 - 9x_2 + 11x_4 - 19x_5 + 3x_6 &= 1.
\end{align*}
\]

Solution. The augmented matrix is
\[
A = \begin{bmatrix}
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
6 & -9 & 0 & 11 & -19 & 3 & 1 \\
\end{bmatrix}
\]
which is row equivalent to
\[
B = \begin{bmatrix}
1 & -\frac{3}{2} & 0 & \frac{11}{6} & -\frac{19}{6} & 0 & \frac{1}{24} \\
0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & 0 & \frac{5}{12} \\
0 & 0 & 0 & 0 & 0 & 1 & \frac{5}{4} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}.
\]

The complete solution is
\[
\begin{align*}
x_1 &= \frac{1}{24} + \frac{3}{2}x_2 - \frac{11}{6}x_4 + \frac{19}{6}x_5, \\
x_3 &= \frac{5}{3} - \frac{1}{3}x_4 + \frac{2}{3}x_5, \\
x_6 &= \frac{1}{4},
\end{align*}
\]
with \(x_2, x_4, x_5\) arbitrary.
(Here \(n = 6, r = 3, c_1 = 1, c_2 = 3, c_3 = 6\); \(c_r = c_3 = 6 < 7 = n + 1\); \(r < n\).)
EXAMPLE 1.4.5 Find the rational number \( t \) for which the following system is consistent and solve the system for this value of \( t \).

\[
\begin{align*}
x + y &= 2 \\
x - y &= 0 \\
3x - y &= t.
\end{align*}
\]

**Solution.** The augmented matrix of the system is

\[
A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 3 & -1 & t \end{bmatrix}
\]

which is row–equivalent to the simpler matrix

\[
B = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & t - 2 \end{bmatrix}.
\]

Hence if \( t \neq 2 \) the system is inconsistent. If \( t = 2 \) the system is consistent and

\[
B = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.
\]

We read off the solution \( x = 1, y = 1 \).

EXAMPLE 1.4.6 For which rationals \( a \) and \( b \) does the following system have (i) no solution, (ii) a unique solution, (iii) infinitely many solutions?

\[
\begin{align*}
x - 2y + 3z &= 4 \\
2x - 3y + az &= 5 \\
3x - 4y + 5z &= b.
\end{align*}
\]

**Solution.** The augmented matrix of the system is

\[
A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 2 & -3 & a & 5 \\ 3 & -4 & 5 & b \end{bmatrix}
\]
1.4. SYSTEMATIC SOLUTION OF LINEAR SYSTEMS.

\[
\begin{align*}
R_2 &\rightarrow R_2 - 2R_1 \\
R_3 &\rightarrow R_3 - 3R_1 \\
R_3 &\rightarrow R_3 - 2R_2
\end{align*}
\]

\[
\begin{bmatrix}
1 & -2 & 3 & 4 \\
0 & 1 & a-6 & -3 \\
0 & 2 & -4 & b-12
\end{bmatrix}
\]

\[
= B.
\]

Case 1. \(a \neq 4\). Then \(-2a + 8 \neq 0\) and we see that \(B\) can be reduced to a matrix of the form

\[
\begin{bmatrix}
1 & 0 & 0 & u \\
0 & 1 & 0 & v \\
0 & 0 & 1 & \frac{b-6}{-2a+8}
\end{bmatrix}
\]

and we have the unique solution \(x = u, y = v, z = (b-6)/(-2a+8)\).

Case 2. \(a = 4\). Then

\[
B = \begin{bmatrix}
1 & -2 & 3 & 4 \\
0 & 1 & -2 & -3 \\
0 & 0 & 0 & b-6
\end{bmatrix}.
\]

If \(b \neq 6\) we get no solution, whereas if \(b = 6\) then

\[
B = \begin{bmatrix}
1 & -2 & 3 & 4 \\
0 & 1 & -2 & -3 \\
0 & 0 & 0 & 0
\end{bmatrix} \quad \text{and} \quad R_1 \rightarrow R_1 + 2R_2 \\
\]

\[
\begin{bmatrix}
1 & 0 & -1 & -2 \\
0 & 1 & -2 & -3 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

We read off the complete solution \(x = -2 + z, y = -3 + 2z,\) with \(z\) arbitrary.

**EXAMPLE 1.4.7** Find the reduced row–echelon form of the following matrix over \(\mathbb{Z}_3\):

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

Hence solve the system

\[
\begin{align*}
2x + y + 2z &= 1 \\
2x + 2y + z &= 0
\end{align*}
\]

over \(\mathbb{Z}_3\).

**Solution.**
\[
\begin{bmatrix}
2 & 1 & 2 & 1 \\
2 & 2 & 1 & 0
\end{bmatrix}
\rightarrow
R_2 \rightarrow R_2 - R_1
\begin{bmatrix}
2 & 1 & 2 & 1 \\
0 & 1 & -1 & -1
\end{bmatrix}
= \begin{bmatrix}
2 & 1 & 2 & 1 \\
0 & 1 & 2 & 2
\end{bmatrix}
\]
\[
R_1 \rightarrow 2R_1
\begin{bmatrix}
1 & 2 & 1 & 2 \\
0 & 1 & 2 & 2
\end{bmatrix}
R_1 \rightarrow R_1 + R_2
\begin{bmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 2 & 2
\end{bmatrix}.
\]

The last matrix is in reduced row–echelon form.

To solve the system of equations whose augmented matrix is the given matrix over \(\mathbb{Z}_3\), we see from the reduced row–echelon form that \(x = 1\) and \(y = 2 - 2z = 2 + z\), where \(z = 0, 1, 2\). Hence there are three solutions to the given system of linear equations: \((x, y, z) = (1, 2, 0), (1, 0, 1)\) and \((1, 1, 2)\).

### 1.5 Homogeneous systems

A system of homogeneous linear equations is a system of the form
\[
\begin{align*}
& a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\
& a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\
& \vdots \\
& a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0.
\end{align*}
\]

Such a system is always consistent as \(x_1 = 0, \cdots, x_n = 0\) is a solution. This solution is called the *trivial* solution. Any other solution is called a *non–trivial* solution.

For example the homogeneous system
\[
\begin{align*}
x - y &= 0 \\
x + y &= 0
\end{align*}
\]
has only the trivial solution, whereas the homogeneous system
\[
\begin{align*}
x - y + z &= 0 \\
x + y + z &= 0
\end{align*}
\]
has the complete solution \(x = -z, y = 0, z\) arbitrary. In particular, taking \(z = 1\) gives the non–trivial solution \(x = -1, y = 0, z = 1\).

There is simple but fundamental theorem concerning homogeneous systems.

**Theorem 1.5.1** A homogeneous system of \(m\) linear equations in \(n\) unknowns always has a non–trivial solution if \(m < n\).
Proof. Suppose that \( m < n \) and that the coefficient matrix of the system is row–equivalent to \( B \), a matrix in reduced row–echelon form. Let \( r \) be the number of non–zero rows in \( B \). Then \( r \leq m < n \) and hence \( n - r > 0 \) and so the number \( n - r \) of arbitrary unknowns is in fact positive. Taking one of these unknowns to be 1 gives a non–trivial solution.

Remark 1.5.1 Let two systems of homogeneous equations in \( n \) unknowns have coefficient matrices \( A \) and \( B \), respectively. If each row of \( B \) is a linear combination of the rows of \( A \) (i.e. a sum of multiples of the rows of \( A \)) and each row of \( A \) is a linear combination of the rows of \( B \), then it is easy to prove that the two systems have identical solutions. The converse is true, but is not easy to prove. Similarly if \( A \) and \( B \) have the same reduced row–echelon form, apart from possibly zero rows, then the two systems have identical solutions and conversely.

There is a similar situation in the case of two systems of linear equations (not necessarily homogeneous), with the proviso that in the statement of the converse, the extra condition that both the systems are consistent, is needed.

1.6 PROBLEMS

1. Which of the following matrices of rationals is in reduced row–echelon form?

   \[
   \begin{pmatrix}
   1 & 0 & 0 & 0 & -3 \\
   0 & 0 & 1 & 0 & 4 \\
   0 & 0 & 0 & 1 & 2 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   0 & 1 & 0 & 0 & 5 \\
   0 & 0 & 1 & 0 & -4 \\
   0 & 0 & 0 & -1 & 3 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   0 & 1 & 0 & 0 \\
   0 & 0 & 1 & 0 \\
   0 & 1 & 0 & -2 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   0 & 1 & 0 & 0 & 0 \\
   0 & 1 & 0 & 0 & 0 \\
   0 & 1 & 0 & 0 & 0 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   1 & 2 & 0 & 0 & 0 \\
   0 & 0 & 1 & 0 & 0 \\
   0 & 0 & 0 & 0 & 0 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   0 & 0 & 0 & 0 \\
   0 & 0 & 0 & 0 \\
   0 & 0 & 0 & 0 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   1 & 0 & 0 & 0 & 1 \\
   0 & 1 & 0 & 0 & 2 \\
   0 & 0 & 0 & 1 & -1 \\
   0 & 0 & 0 & 0 & 0 \\
   \end{pmatrix}
   \]

   [Answers: (a), (e), (g)]

2. Find reduced row–echelon forms which are row–equivalent to the following matrices:

   \[
   \begin{pmatrix}
   0 & 0 & 0 \\
   2 & 4 & 0 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   0 & 1 & 3 \\
   1 & 2 & 4 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   1 & 1 & 1 \\
   1 & 1 & 0 \\
   1 & 0 & 0 \\
   \end{pmatrix}
   \quad \begin{pmatrix}
   2 & 0 & 0 \\
   0 & 0 & 0 \\
   -4 & 0 & 0 \\
   \end{pmatrix}
   .
   \]
3. Solve the following systems of linear equations by reducing the augmented matrix to reduced row–echelon form:

(a) \[
\begin{align*}
x + y + z &= 2 \\
2x + 3y - z &= 8 \\
x - y - z &= -8
\end{align*}
\]
(b) \[
\begin{align*}
x_1 + x_2 - x_3 + 2x_4 &= 10 \\
3x_1 - x_2 + 7x_3 + 4x_4 &= 1 \\
-5x_1 + 3x_2 - 15x_3 - 6x_4 &= 9
\end{align*}
\]
(c) \[
\begin{align*}
3x - y + 7z &= 0 \\
2x - y + 4z &= \frac{1}{2} \\
x - y + z &= 1 \\
6x - 4y + 10z &= 3
\end{align*}
\]
(d) \[
\begin{align*}
2x_2 + 3x_3 - 4x_4 &= 1 \\
2x_3 + 3x_4 &= 4 \\
2x_1 + 2x_2 - 5x_3 + 2x_4 &= 4 \\
2x_1 - 6x_3 + 9x_4 &= 7
\end{align*}
\]

3. Solve the following systems of linear equations by reducing the augmented matrix to reduced row–echelon form:

(a) \[
\begin{align*}
1 & 2 & 0 \\
0 & 0 & 0
\end{align*}
\]
(b) \[
\begin{align*}
1 & 0 & -2 \\
0 & 1 & 3
\end{align*}
\]
(c) \[
\begin{align*}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{align*}
\]
(d) \[
\begin{align*}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{align*}
\]

4. Show that the following system is consistent if and only if \( c = 2a - 3b \) and solve the system in this case.

\[
\begin{align*}
2x - y + 3z &= a \\
3x + y - 5z &= b \\
-5x - 5y + 21z &= c
\end{align*}
\]

5. Find the value of \( t \) for which the following system is consistent and solve the system for this value of \( t \).

\[
\begin{align*}
x + y &= 1 \\
(tx + y) &= t \\
(1 + t)x + 2y &= 3
\end{align*}
\]

[Answer: \( t = 2; x = 1, y = 0. \)]
6. Solve the homogeneous system
\[-3x_1 + x_2 + x_3 + x_4 = 0 \\
x_1 - 3x_2 + x_3 + x_4 = 0 \\
x_1 + x_2 - 3x_3 + x_4 = 0 \\
x_1 + x_2 + x_3 - 3x_4 = 0.\]

[Answer: $x_1 = x_2 = x_3 = x_4$, with $x_4$ arbitrary.]

7. For which rational numbers $\lambda$ does the homogeneous system
\[x + (\lambda - 3)y = 0 \\
(\lambda - 3)x + y = 0\]
have a non-trivial solution?

[Answer: $\lambda = 2, 4$.]

8. Solve the homogeneous system
\[3x_1 + x_2 + x_3 + x_4 = 0 \\
5x_1 - x_2 + x_3 - x_4 = 0.\]

[Answer: $x_1 = -\frac{1}{4}x_3, x_2 = -\frac{1}{4}x_3 - x_4$, with $x_3$ and $x_4$ arbitrary.]

9. Let $A$ be the coefficient matrix of the following homogeneous system of $n$ equations in $n$ unknowns:
\[(1 - n)x_1 + x_2 + \cdots + x_n = 0 \\
x_1 + (1 - n)x_2 + \cdots + x_n = 0 \\
\cdots = 0 \\
x_1 + x_2 + \cdots + (1 - n)x_n = 0.\]

Find the reduced row-echelon form of $A$ and hence, or otherwise, prove that the solution of the above system is $x_1 = x_2 = \cdots = x_n$, with $x_n$ arbitrary.

10. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix over a field $F$. Prove that $A$ is row-equivalent to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ if $ad - bc \neq 0$, but is row-equivalent to a matrix whose second row is zero, if $ad - bc = 0$. 
11. For which rational numbers $a$ does the following system have (i) no solutions (ii) exactly one solution (iii) infinitely many solutions?

\[
\begin{align*}
    x + 2y - 3z &= 4 \\
    3x - y + 5z &= 2 \\
    4x + y + (a^2 - 14)z &= a + 2.
\end{align*}
\]

[Answer: $a = -4$, no solution; $a = 4$, infinitely many solutions; $a \neq \pm 4$, exactly one solution.]

12. Solve the following system of homogeneous equations over $\mathbb{Z}_2$:

\[
\begin{align*}
    x_1 + x_3 + x_5 &= 0 \\
    x_2 + x_4 + x_5 &= 0 \\
    x_1 + x_2 + x_3 + x_4 &= 0 \\
    x_3 + x_4 &= 0.
\end{align*}
\]

[Answer: $x_1 = x_2 = x_4 + x_5, x_3 = x_4$, with $x_4$ and $x_5$ arbitrary elements of $\mathbb{Z}_2$.]

13. Solve the following systems of linear equations over $\mathbb{Z}_5$:

\[
\begin{align*}
    (a) \quad 2x + y + 3z &= 4 & (b) \quad 2x + y + 3z &= 4 \\
    \quad 4x + y + 4z &= 1 & \quad 4x + y + 4z &= 1 \\
    \quad 3x + y + 2z &= 0 & \quad x + y &= 3.
\end{align*}
\]

[Answer: (a) $x = 1$, $y = 2$, $z = 0$; (b) $x = 1 + 2z$, $y = 2 + 3z$, with $z$ an arbitrary element of $\mathbb{Z}_5$.]

14. If $(\alpha_1, \ldots, \alpha_n)$ and $(\beta_1, \ldots, \beta_n)$ are solutions of a system of linear equations, prove that

\[(1 - t)\alpha_1 + t\beta_1, \ldots, (1 - t)\alpha_n + t\beta_n\]

is also a solution.

15. If $(\alpha_1, \ldots, \alpha_n)$ is a solution of a system of linear equations, prove that the complete solution is given by $x_1 = \alpha_1 + y_1, \ldots, x_n = \alpha_n + y_n$, where $(y_1, \ldots, y_n)$ is the general solution of the associated homogeneous system.
16. Find the values of $a$ and $b$ for which the following system is consistent. Also find the complete solution when $a = b = 2$.

\[
\begin{align*}
x + y - z + w &= 1 \\
a x + y + z + w &= b \\
3x + 2y + aw &= 1 + a.
\end{align*}
\]

[Answer: $a \neq 2$ or $a = 2 = b$; $x = 1 - 2z$, $y = 3z - w$, with $z$, $w$ arbitrary.]

17. Let $F = \{0, 1, a, b\}$ be a field consisting of 4 elements.

(a) Determine the addition and multiplication tables of $F$. (Hint: Prove that the elements $1 + 0$, $1 + 1$, $1 + a$, $1 + b$ are distinct and deduce that $1 + 1 + 1 + 1 = 0$; then deduce that $1 + 1 = 0$.)

(b) A matrix $A$, whose elements belong to $F$, is defined by

\[
A = \begin{bmatrix}
1 & a & b & a \\
1 & b & b & 1 \\
a & b & b & 1 \\
1 & 1 & 1 & a
\end{bmatrix}.
\]

prove that the reduced row–echelon form of $A$ is given by the matrix

\[
B = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & b \\
0 & 0 & 1 & 1
\end{bmatrix}.
\]
Chapter 2

MATRICES

2.1 Matrix arithmetic

A matrix over a field $F$ is a rectangular array of elements from $F$. The symbol $M_{m \times n}(F)$ denotes the collection of all $m \times n$ matrices over $F$. Matrices will usually be denoted by capital letters and the equation $A = [a_{ij}]$ means that the element in the $i$-th row and $j$-th column of the matrix $A$ equals $a_{ij}$. It is also occasionally convenient to write $a_{ij} = (A)_{ij}$. For the present, all matrices will have rational entries, unless otherwise stated.

**EXAMPLE 2.1.1** The formula $a_{ij} = 1/(i + j)$ for $1 \leq i \leq 3, 1 \leq j \leq 4$ defines a $3 \times 4$ matrix $A = [a_{ij}]$, namely

$$A = \begin{bmatrix}
\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\
\frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\
\frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7}
\end{bmatrix}.$$  

**DEFINITION 2.1.1 (Equality of matrices)** Matrices $A$ and $B$ are said to be equal if $A$ and $B$ have the same size and corresponding elements are equal; that is $A$ and $B \in M_{m \times n}(F)$ and $A = [a_{ij}], B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \leq i \leq m, 1 \leq j \leq n$.

**DEFINITION 2.1.2 (Addition of matrices)** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be of the same size. Then $A + B$ is the matrix obtained by adding corresponding elements of $A$ and $B$; that is

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$  

23
DEFINITION 2.1.3 (Scalar multiple of a matrix) Let \( A = [a_{ij}] \) and \( t \in F \) (that is \( t \) is a scalar). Then \( tA \) is the matrix obtained by multiplying all elements of \( A \) by \( t \); that is
\[
tA = t[a_{ij}] = [ta_{ij}].
\]

DEFINITION 2.1.4 (Additive inverse of a matrix) Let \( A = [a_{ij}] \). Then \(-A\) is the matrix obtained by replacing the elements of \( A \) by their additive inverses; that is
\[
-A = -[a_{ij}] = [-a_{ij}].
\]

DEFINITION 2.1.5 (Subtraction of matrices) Matrix subtraction is defined for two matrices \( A = [a_{ij}] \) and \( B = [b_{ij}] \) of the same size, in the usual way; that is
\[
A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}].
\]

DEFINITION 2.1.6 (The zero matrix) For each \( m, n \) the matrix in \( M_{m \times n}(F) \), all of whose elements are zero, is called the zero matrix (of size \( m \times n \)) and is denoted by the symbol 0.

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, \( s \) and \( t \) will be arbitrary scalars and \( A, B, C \) are matrices of the same size.)

1. \((A + B) + C = A + (B + C)\);
2. \(A + B = B + A\);
3. \(0 + A = A\);
4. \(A + (-A) = 0\);
5. \((s + t)A = sA + tA, \quad (s - t)A = sA - tA\);
6. \(t(A + B) = tA + tB, \quad t(A - B) = tA - tB\);
7. \(s(tA) = (st)A\);
8. \(1A = A, \quad 0A = 0, \quad (-1)A = -A\);
9. \(tA = 0 \Rightarrow t = 0 \) or \( A = 0 \).

Other similar properties will be used when needed.
DEFINITION 2.1.7 (Matrix product) Let $A = [a_{ij}]$ be a matrix of size $m \times n$ and $B = [b_{jk}]$ be a matrix of size $n \times p$; (that is the number of columns of $A$ equals the number of rows of $B$). Then $AB$ is the $m \times p$ matrix $C = [c_{ik}]$ whose $(i, k)$–th element is defined by the formula

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}.$$

EXAMPLE 2.1.2

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$;
2. $\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$;
3. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$;
4. $\begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 11 \end{bmatrix}$;
5. $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Matrix multiplication obeys many of the familiar laws of arithmetic apart from the commutative law.

1. $(AB)C = A(BC)$ if $A$, $B$, $C$ are $m \times n$, $n \times p$, $p \times q$, respectively;
2. $t(AB) = (tA)B = A(tB)$, $A(-B) = (-A)B = -(AB)$;
3. $(A + B)C = AC + BC$ if $A$ and $B$ are $m \times n$ and $C$ is $n \times p$;
4. $D(A + B) = DA + DB$ if $A$ and $B$ are $m \times n$ and $D$ is $p \times m$.

We prove the associative law only:
First observe that $(AB)C$ and $A(BC)$ are both of size $m \times q$.
Let $A = [a_{ij}]$, $B = [b_{jk}]$, $C = [c_{kl}]$. Then

$$((AB)C)_{il} = \sum_{k=1}^{p} (AB)_{ik}c_{kl} = \sum_{k=1}^{p} \left( \sum_{j=1}^{n} a_{ij}b_{jk} \right) c_{kl} = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij}b_{jk}c_{kl}.$$
Similarly
\[(A(BC))_{il} = \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ij} b_{jk} c_{kl}.\]

However the double summations are equal. For sums of the form
\[\sum_{j=1}^{n} \sum_{k=1}^{p} d_{jk}\]
and
\[\sum_{p}^{n} \sum_{k=1}^{n} d_{jk}\]
represent the sum of the \(np\) elements of the rectangular array \([d_{jk}]\), by rows and by columns, respectively. Consequently
\[((AB)C)_{il} = (A(BC))_{il}\]
for \(1 \leq i \leq m, 1 \leq l \leq q\). Hence \((AB)C = A(BC)\).

The system of \(m\) linear equations in \(n\) unknowns
\[
\begin{align*}
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
  \vdots & \\
  a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]
is equivalent to a single matrix equation
\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
= \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix},
\]
that is \(AX = B\), where \(A = [a_{ij}]\) is the coefficient matrix of the system,
\[
X = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]
is the vector of unknowns and \(B = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}\) is the vector of constants.

Another useful matrix equation equivalent to the above system of linear equations is
\[
\begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix}
x_1 + \begin{bmatrix}
  a_{12} \\
  a_{22} \\
  \vdots \\
  a_{m2}
\end{bmatrix}
x_2 + \cdots + \begin{bmatrix}
  a_{1n} \\
  a_{2n} \\
  \vdots \\
  a_{mn}
\end{bmatrix}
x_n = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}.
\]
EXAMPLE 2.1.3 The system
\[
\begin{align*}
  x + y + z &= 1 \\
  x - y + z &= 0.
\end{align*}
\]
is equivalent to the matrix equation
\[
\begin{bmatrix}
  1 & 1 & 1 \\
  1 & -1 & 1
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} =
\begin{bmatrix}
  1 \\
  0
\end{bmatrix}
\]
and to the equation
\[
x \begin{bmatrix}
  1 \\
  1
\end{bmatrix} + y \begin{bmatrix}
  1 \\
  -1
\end{bmatrix} + z \begin{bmatrix}
  1 \\
  1
\end{bmatrix} =
\begin{bmatrix}
  1 \\
  0
\end{bmatrix}.
\]

2.2 Linear transformations

An \textit{n–dimensional column vector} is an \(n \times 1\) matrix over \(F\). The collection of all \(n\–dimensional column vectors is denoted by \(F^n\).

Every matrix is associated with an important type of function called a \textit{linear transformation}.

**DEFINITION 2.2.1 (Linear transformation)** With \(A \in M_{m \times n}(F)\), we associate the function \(T_A : F^n \rightarrow F^m\) defined by \(T_A(X) = AX\) for all \(X \in F^n\). More explicitly, using components, the above function takes the form

\[
\begin{align*}
y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n,
\end{align*}
\]

where \(y_1, y_2, \cdots, y_m\) are the components of the column vector \(T_A(X)\).

The function just defined has the property that

\[
T_A(sX + tY) = sT_A(X) + tT_A(Y)
\]
for all \(s, t \in F\) and all \(n\–dimensional column vectors \(X, Y\). For

\[
T_A(sX + tY) = A(sX + tY) = s(AX) + t(AY) = sT_A(X) + tT_A(Y).
\]
REMARK 2.2.1 It is easy to prove that if $T : F^n \to F^m$ is a function satisfying equation 2.1, then $T = T_A$, where $A$ is the $m \times n$ matrix whose columns are $T(E_1), \ldots, T(E_n)$, respectively, where $E_1, \ldots, E_n$ are the $n$-dimensional unit vectors defined by

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

One well-known example of a linear transformation arises from rotating the $(x, y)$-plane in 2-dimensional Euclidean space, anticlockwise through $\theta$ radians. Here a point $(x, y)$ will be transformed into the point $(x_1, y_1)$, where

$$x_1 = x \cos \theta - y \sin \theta, \quad y_1 = x \sin \theta + y \cos \theta.$$

In 3-dimensional Euclidean space, the equations

$$x_1 = x \cos \theta - y \sin \theta, \quad y_1 = x \sin \theta + y \cos \theta, \quad z_1 = z;$$

$$x_1 = x, \quad y_1 = y \cos \phi - z \sin \phi, \quad z_1 = y \sin \phi + z \cos \phi;$$

$$x_1 = x \cos \psi - z \sin \psi, \quad y_1 = y, \quad z_1 = x \sin \psi + z \cos \psi;$$

correspond to rotations about the positive $z, x, y$-axes, anticlockwise through $\theta, \phi, \psi$ radians, respectively.

The product of two matrices is related to the product of the corresponding linear transformations:

If $A$ is $m \times n$ and $B$ is $n \times p$, then the function $T_AT_B : F^p \to F^m$, obtained by first performing $T_B$, then $T_A$ is in fact equal to the linear transformation $T_{AB}$. For if $X \in F^p$, we have

$$T_AT_B(X) = A(BX) = (AB)X = T_{AB}(X).$$

The following example is useful for producing rotations in 3-dimensional animated design. (See [27, pages 97–112].)

EXAMPLE 2.2.1 The linear transformation resulting from successively rotating 3-dimensional space about the positive $z, x, y$-axes, anticlockwise through $\theta, \phi, \psi$ radians respectively, is equal to $T_{ABC}$, where
2.2. LINEAR TRANSFORMATIONS

\begin{align*}
\mathbf{C} &= \begin{bmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix}, \quad 
\mathbf{B} &= \begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \phi & -\sin \phi \\
0 & \sin \phi & \cos \phi
\end{bmatrix}.
\end{align*}

\begin{align*}
\mathbf{A} &= \begin{bmatrix}
\cos \psi & 0 & -\sin \psi \\
0 & 1 & 0 \\
\sin \psi & 0 & \cos \psi
\end{bmatrix}.
\end{align*}

The matrix \( ABC \) is quite complicated:

\begin{align*}
A(BC) &= \begin{bmatrix}
\cos \psi & 0 & -\sin \psi \\
0 & 1 & 0 \\
\sin \psi & 0 & \cos \psi
\end{bmatrix}
\begin{bmatrix}
\cos \theta & -\sin \theta & 0 \\
\cos \phi \sin \theta & \cos \phi \cos \theta & -\sin \phi \\
\sin \phi \sin \theta & \sin \phi \cos \theta & \cos \phi
\end{bmatrix}

&= \begin{bmatrix}
\cos \psi \cos \theta - \sin \psi \sin \phi \sin \theta & -\cos \psi \sin \theta - \sin \psi \sin \phi \sin \theta & - \sin \psi \cos \phi \\
\cos \phi \sin \theta & \cos \phi \cos \theta & -\sin \phi \\
\sin \phi \sin \theta + \cos \psi \sin \phi \sin \theta & -\sin \psi \sin \theta + \cos \psi \sin \phi \cos \theta & \cos \psi \cos \phi
\end{bmatrix}.
\end{align*}

**EXAMPLE 2.2.2** Another example of a linear transformation arising from geometry is reflection of the plane in a line \( l \) inclined at an angle \( \theta \) to the positive \( x \)-axis.

We reduce the problem to the simpler case \( \theta = 0 \), where the equations of transformation are \( x_1 = x, y_1 = -y \). First rotate the plane clockwise through \( \theta \) radians, thereby taking \( l \) into the \( x \)-axis; next reflect the plane in the \( x \)-axis; then rotate the plane anticlockwise through \( \theta \) radians, thereby restoring \( l \) to its original position.
In terms of matrices, we get transformation equations
\[
\begin{bmatrix}
  x_1 \\
  y_1
\end{bmatrix} = \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  1 & 0 \\
  0 & -1
\end{bmatrix} \begin{bmatrix}
  \cos (-\theta) & -\sin (-\theta) \\
  \sin (-\theta) & \cos (-\theta)
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  \cos \theta & \sin \theta \\
  \sin \theta & -\cos \theta
\end{bmatrix} \begin{bmatrix}
  \cos \theta & \sin \theta \\
  -\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  \cos 2\theta & \sin 2\theta \\
  \sin 2\theta & -\cos 2\theta
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix}.
\]

The more general transformation
\[
\begin{bmatrix}
  x_1 \\
  y_1
\end{bmatrix} = a \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix} + \begin{bmatrix}
  u \\
  v
\end{bmatrix}, \quad a > 0,
\]

represents a rotation, followed by a scaling and then by a translation. Such transformations are important in computer graphics. See [23, 24].

**EXAMPLE 2.2.3** Our last example of a geometrical linear transformation arises from projecting the plane onto a line \( l \) through the origin, inclined at angle \( \theta \) to the positive \( x \)-axis. Again we reduce that problem to the simpler case where \( l \) is the \( x \)-axis and the equations of transformation are \( x_1 = x, y_1 = 0 \).

In terms of matrices, we get transformation equations
\[
\begin{bmatrix}
  x_1 \\
  y_1
\end{bmatrix} = \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
  1 & 0 \\
  0 & 0
\end{bmatrix} \begin{bmatrix}
  \cos (-\theta) & -\sin (-\theta) \\
  \sin (-\theta) & \cos (-\theta)
\end{bmatrix} \begin{bmatrix}
  x \\
  y
\end{bmatrix}
\]
\[
\begin{align*}
= & \begin{bmatrix}
\cos \theta & 0 \\
\sin \theta & 0
\end{bmatrix}
\begin{bmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix} \\
= & \begin{bmatrix}
\cos^2 \theta & \cos \theta \sin \theta \\
\sin \theta \cos \theta & \sin^2 \theta
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}.
\end{align*}
\]

2.3 Recurrence relations

**DEFINITION 2.3.1 (The identity matrix)** The \( n \times n \) matrix \( I_n = [\delta_{ij}] \), defined by \( \delta_{ij} = 1 \) if \( i = j \), \( \delta_{ij} = 0 \) if \( i \neq j \), is called the \( n \times n \) identity matrix of order \( n \). In other words, the columns of the identity matrix of order \( n \) are the unit vectors \( E_1, \ldots, E_n \), respectively.

For example, \( I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \).

**THEOREM 2.3.1** If \( A \) is \( m \times n \), then \( I_mA = A = AI_n \).

**DEFINITION 2.3.2 (k–th power of a matrix)** If \( A \) is an \( n \times n \) matrix, we define \( A^k \) recursively as follows: \( A^0 = I_n \) and \( A^{k+1} = A^kA \) for \( k \geq 0 \).

For example \( A^1 = A^0A = I_nA = A \) and hence \( A^2 = A^1A = AA \).

The usual index laws hold provided \( AB = BA \):

1. \( A^m A^n = A^{m+n} \), \( (A^m)^n = A^{mn} \);
2. \( (AB)^n = A^nB^n \);
3. \( A^m B^n = B^n A^m \);
4. \( (A + B)^2 = A^2 + 2AB + B^2 \);
5. \( (A + B)^n = \sum_{i=0}^{n} \binom{n}{i} A^i B^{n-i} \);
6. \( (A + B)(A - B) = A^2 - B^2 \).

We now state a basic property of the natural numbers.

**AXIOM 2.3.1 (PRINCIPLE OF MATHEMATICAL INDUCTION)**

If for each \( n \geq 1 \), \( P_n \) denotes a mathematical statement and

(i) \( P_1 \) is true,
(ii) the truth of $P_n$ implies that of $P_{n+1}$ for each $n \geq 1$.

then $P_n$ is true for all $n \geq 1$.

**EXAMPLE 2.3.1** Let $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$. Prove that

$$A^n = \begin{bmatrix} 1 + 6n & 4n \\ -9n & 1 - 6n \end{bmatrix} \text{ if } n \geq 1.$$

**Solution.** We use the principle of mathematical induction.

Take $P_n$ to be the statement

$$A^n = \begin{bmatrix} 1 + 6n & 4n \\ -9n & 1 - 6n \end{bmatrix}.$$  

Then $P_1$ asserts that

$$A^1 = \begin{bmatrix} 1 + 6 \times 1 & 4 \times 1 \\ -9 \times 1 & 1 - 6 \times 1 \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix},$$

which is true. Now let $n \geq 1$ and assume that $P_n$ is true. We have to deduce that

$$A^{n+1} = \begin{bmatrix} 1 + 6(n+1) & 4(n+1) \\ -9(n+1) & 1 - 6(n+1) \end{bmatrix} = \begin{bmatrix} 7 + 6n & 4n + 4 \\ -9n - 9 & -5 - 6n \end{bmatrix}.$$  

Now

$$A^{n+1} = A^n A = \begin{bmatrix} 1 + 6n & 4n \\ -9n & 1 - 6n \end{bmatrix} \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix} = \begin{bmatrix} (1 + 6n)7 + (4n)(-9) & (1 + 6n)4 + (4n)(-5) \\ (-9n)7 + (1 - 6n)(-9) & (-9n)4 + (1 - 6n)(-5) \end{bmatrix} = \begin{bmatrix} 7 + 6n & 4n + 4 \\ -9n - 9 & -5 - 6n \end{bmatrix},$$

and “the induction goes through”.

The last example has an application to the solution of a system of recurrence relations:
EXAMPLE 2.3.2 The following system of recurrence relations holds for all $n \geq 0$:

$$
\begin{align*}
    x_{n+1} &= 7x_n + 4y_n \\
y_{n+1} &= -9x_n - 5y_n.
\end{align*}
$$

Solve the system for $x_n$ and $y_n$ in terms of $x_0$ and $y_0$.

**Solution.** Combine the above equations into a single matrix equation

$$
\begin{bmatrix}
    x_{n+1} \\
y_{n+1}
\end{bmatrix} =
\begin{bmatrix}
    7 & 4 \\
-9 & -5
\end{bmatrix}
\begin{bmatrix}
    x_n \\
y_n
\end{bmatrix},
$$

or $X_{n+1} = AX_n$, where $A = \begin{bmatrix}
    7 & 4 \\
-9 & -5
\end{bmatrix}$ and $X_n = \begin{bmatrix}
    x_n \\
y_n
\end{bmatrix}$.

We see that

$$
\begin{align*}
    X_1 &= AX_0 \\
X_2 &= AX_1 = A(AX_0) = A^2X_0 \\
&\vdots \\
X_n &= A^nX_0.
\end{align*}
$$

(The truth of the equation $X_n = A^nX_0$ for $n \geq 1$, strictly speaking, follows by mathematical induction; however for simple cases such as the above, it is customary to omit the strict proof and supply instead a few lines of motivation for the inductive statement.)

Hence the previous example gives

$$
\begin{align*}
    \begin{bmatrix}
        x_n \\
y_n
    \end{bmatrix} = X_n &=
    \begin{bmatrix}
        1 + 6n & 4n \\
-9n & 1 - 6n
    \end{bmatrix}
    \begin{bmatrix}
        x_0 \\
y_0
    \end{bmatrix} \\
    &=
    \begin{bmatrix}
        (1 + 6n)x_0 + (4n)y_0 \\
(-9n)x_0 + (1 - 6n)y_0
    \end{bmatrix},
\end{align*}
$$

and hence $x_n = (1 + 6n)x_0 + 4ny_0$ and $y_n = (-9n)x_0 + (1 - 6n)y_0$, for $n \geq 1$.

2.4 PROBLEMS

1. Let $A, B, C, D$ be matrices defined by

$$
A = \begin{bmatrix}
    3 & 0 \\
-1 & 2 \\
1 & 1
\end{bmatrix}, \quad B = \begin{bmatrix}
    1 & 5 & 2 \\
-1 & 1 & 0 \\
-4 & 1 & 3
\end{bmatrix}.
$$
\[ C = \begin{bmatrix} -3 & -1 \\ 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & -1 \\ 2 & 0 \end{bmatrix}. \]

Which of the following matrices are defined? Compute those matrices which are defined.

\( A + B, A + C, AB, BA, CD, DC, D^2. \)

[Answers: \( A + C, BA, CD, D^2; \)
\[ \begin{bmatrix} 0 & -1 \\ 1 & 3 \\ 5 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 12 \\ -4 & 2 \\ -10 & 5 \end{bmatrix}, \quad \begin{bmatrix} -14 & 3 \\ 10 & -2 \\ 22 & -4 \end{bmatrix}, \quad \begin{bmatrix} 14 & -4 \\ 8 & -2 \end{bmatrix}. \]

2. Let \( A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}. \) Show that if \( B \) is a \( 3 \times 2 \) such that \( AB = I_2, \) then

\[ B = \begin{bmatrix} a & b \\ -a - 1 & 1 - b \\ a + 1 & b \end{bmatrix} \]

for suitable numbers \( a \) and \( b. \) Use the associative law to show that \( (BA)^2 B = B. \)

3. If \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \) prove that \( A^2 - (a + d)A + (ad - bc)I_2 = 0.\)

4. If \( A = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}, \) use the fact \( A^2 = 4A - 3I_2 \) and mathematical induction, to prove that

\[ A^n = \frac{(3^n - 1)}{2} A + \frac{3 - 3^n}{2} I_2 \quad \text{if} \quad n \geq 1. \]

5. A sequence of numbers \( x_1, x_2, \ldots, x_n, \ldots \) satisfies the recurrence relation \( x_{n+1} = ax_n + bx_{n-1} \) for \( n \geq 1, \) where \( a \) and \( b \) are constants. Prove that

\[ \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = A \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}, \]
where \( A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \) and hence express \( \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} \) in terms of \( \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} \).

If \( a = 4 \) and \( b = -3 \), use the previous question to find a formula for \( x_n \) in terms of \( x_1 \) and \( x_0 \).

[Answer:
\[
x_n = \frac{3^n - 1}{2} x_1 + \frac{3 - 3^n}{2} x_0.
\]
]

6. Let \( A = \begin{bmatrix} 2a & -a^2 \\ 1 & 0 \end{bmatrix} \).

(a) Prove that

\[
A^n = \begin{bmatrix} (n+1)a^n & -na^{n+1} \\ na^{n-1} & (1-n)a^n \end{bmatrix} \quad \text{if } n \geq 1.
\]

(b) A sequence \( x_0, x_1, \ldots, x_n, \ldots \) satisfies the recurrence relation \( x_{n+1} = 2ax_n - a^2x_{n-1} \) for \( n \geq 1 \). Use part (a) and the previous question to prove that \( x_n = na^{n-1}x_1 + (1-n)a^n x_0 \) for \( n \geq 1 \).

7. Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) and suppose that \( \lambda_1 \) and \( \lambda_2 \) are the roots of the quadratic polynomial \( x^2 - (a+d)x + ad - bc \). (\( \lambda_1 \) and \( \lambda_2 \) may be equal.) Let \( k_n \) be defined by \( k_0 = 0, k_1 = 1 \) and for \( n \geq 2 \)
\[
k_n = \sum_{i=1}^{n} \lambda_1^{n-i} \lambda_2^{i-1}.
\]

Prove that
\[
k_{n+1} = (\lambda_1 + \lambda_2)k_n - \lambda_1\lambda_2 k_{n-1},
\]
if \( n \geq 1 \). Also prove that
\[
k_n = \begin{cases} 
(\lambda_1^n - \lambda_2^n)/(\lambda_1 - \lambda_2) & \text{if } \lambda_1 \neq \lambda_2, \\
n\lambda_1^{n-1} & \text{if } \lambda_1 = \lambda_2.
\end{cases}
\]

Use mathematical induction to prove that if \( n \geq 1 \),
\[
A^n = k_n A - \lambda_1\lambda_2 k_{n-1} I_2,
\]
[Hint: Use the equation \( A^2 = (a+d)A - (ad - bc)I_2 \).]
8. Use Question 6 to prove that if \( A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \), then

\[
A^n = \frac{3^n}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{(-1)^{n-1}}{2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}
\]

if \( n \geq 1 \).

9. The Fibonacci numbers are defined by the equations \( F_0 = 0, F_1 = 1 \) and \( F_{n+1} = F_n + F_{n-1} \) if \( n \geq 1 \). Prove that

\[
F_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right)
\]

if \( n \geq 0 \).

10. Let \( r > 1 \) be an integer. Let \( a \) and \( b \) be arbitrary positive integers. Sequences \( x_n \) and \( y_n \) of positive integers are defined in terms of \( a \) and \( b \) by the recurrence relations

\[
\begin{align*}
x_{n+1} & = x_n + ry_n \\
y_{n+1} & = x_n + y_n,
\end{align*}
\]

for \( n \geq 0 \), where \( x_0 = a \) and \( y_0 = b \).

Use Question 6 to prove that

\[
\frac{x_n}{y_n} \to \sqrt{r} \quad \text{as} \quad n \to \infty.
\]

2.5 Non–singular matrices

**DEFINITION 2.5.1 (Non–singular matrix)**

A square matrix \( A \in M_{n \times n}(F) \) is called non–singular or invertible if there exists a matrix \( B \in M_{n \times n}(F) \) such that

\[ AB = I_n = BA. \]

Any matrix \( B \) with the above property is called an inverse of \( A \). If \( A \) does not have an inverse, \( A \) is called singular.
THEOREM 2.5.1 (Inverses are unique)

If $A$ has inverses $B$ and $C$, then $B = C$.

**Proof.** Let $B$ and $C$ be inverses of $A$. Then $AB = I_n = BA$ and $AC = I_n = CA$. Then $B(AC) = BI_n = B$ and $(BA)C = I_nC = C$. Hence because $B(AC) = (BA)C$, we deduce that $B = C$.

**REMARK 2.5.1** If $A$ has an inverse, it is denoted by $A^{-1}$. So

$$AA^{-1} = I_n = A^{-1}A.$$  

Also if $A$ is non–singular, it follows that $A^{-1}$ is also non–singular and

$$(A^{-1})^{-1} = A.$$  

**THEOREM 2.5.2** If $A$ and $B$ are non–singular matrices of the same size, then so is $AB$. Moreover

$$(AB)^{-1} = B^{-1}A^{-1}.$$  

**Proof.**

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$  

Similarly

$$(B^{-1}A^{-1})(AB) = I_n.$$  

**REMARK 2.5.2** The above result generalizes to a product of $m$ non–singular matrices: If $A_1, \ldots, A_m$ are non–singular $n \times n$ matrices, then the product $A_1 \ldots A_m$ is also non–singular. Moreover

$$(A_1 \ldots A_m)^{-1} = A_m^{-1} \ldots A_1^{-1}.$$  

(Thus the inverse of the product equals the product of the inverses in the reverse order.)

**EXAMPLE 2.5.1** If $A$ and $B$ are $n \times n$ matrices satisfying $A^2 = B^2 = (AB)^2 = I_n$, prove that $AB = BA$.

**Solution.** Assume $A^2 = B^2 = (AB)^2 = I_n$. Then $A, B, AB$ are non–singular and $A^{-1} = A, B^{-1} = B, (AB)^{-1} = AB$.

But $(AB)^{-1} = B^{-1}A^{-1}$ and hence $AB = BA$. 
EXAMPLE 2.5.2 $A = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$ is singular. For suppose $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is an inverse of $A$. Then the equation $AB = I_2$ gives

$$\begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and equating the corresponding elements of column 1 of both sides gives the system

$$a + 2c = 1$$
$$4a + 8c = 0$$

which is clearly inconsistent.

THEOREM 2.5.3 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\Delta = ad - bc \neq 0$. Then $A$ is non–singular. Also

$$A^{-1} = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

REMARK 2.5.3 The expression $ad - bc$ is called the determinant of $A$ and is denoted by the symbols $\det A$ or $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Proof. Verify that the matrix $B = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ satisfies the equation $AB = I_2 = BA$.

EXAMPLE 2.5.3 Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 5 & 0 & 0 \end{bmatrix}.$$ 

Verify that $A^3 = 5I_3$, deduce that $A$ is non–singular and find $A^{-1}$.

Solution. After verifying that $A^3 = 5I_3$, we notice that

$$A \left( \frac{1}{5} A^2 \right) = I_3 = \left( \frac{1}{5} A^2 \right) A.$$ 

Hence $A$ is non–singular and $A^{-1} = \frac{1}{5} A^2$. 
THEOREM 2.5.4 If the coefficient matrix $A$ of a system of $n$ equations in $n$ unknowns is non–singular, then the system $AX = B$ has the unique solution $X = A^{-1}B$.

Proof. Assume that $A^{-1}$ exists.

1. (Uniqueness.) Assume that $AX = B$. Then

\[
(A^{-1}A)X = A^{-1}B, \\
I_nX = A^{-1}B, \\
X = A^{-1}B.
\]

2. (Existence.) Let $X = A^{-1}B$. Then

\[
AX = A(A^{-1}B) = (AA^{-1})B = I_n B = B.
\]

THEOREM 2.5.5 (Cramer’s rule for 2 equations in 2 unknowns)

The system

\[
ax + by = e \\
 cx + dy = f
\]

has a unique solution if $\Delta = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, namely

\[
x = \Delta_1 \Delta, \\
y = \Delta_2 \Delta,
\]

where

\[
\Delta_1 = \begin{vmatrix} e & b \\ f & d \end{vmatrix} \quad \text{and} \quad \Delta_2 = \begin{vmatrix} a & e \\ c & f \end{vmatrix}.
\]

Proof. Suppose $\Delta \neq 0$. Then $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has inverse

\[
A^{-1} = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}
\]

and we know that the system

\[
A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix}
\]
has the unique solution

\[
\begin{bmatrix}
  x \\
  y
\end{bmatrix} = A^{-1} \begin{bmatrix}
  e \\
  f
\end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix}
  d & -b \\
  -c & a
\end{bmatrix} \begin{bmatrix}
  e \\
  f
\end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix}
  \Delta_1 \\
  \Delta_2
\end{bmatrix} = \begin{bmatrix}
  \Delta_1/\Delta \\
  \Delta_2/\Delta
\end{bmatrix}.
\]

Hence \( x = \Delta_1/\Delta, \ y = \Delta_2/\Delta \).

**COROLLARY 2.5.1** The homogeneous system

\[
\begin{align*}
ax + by &= 0 \\
 cx + dy &= 0
\end{align*}
\]

has only the trivial solution if \( \Delta = \begin{vmatrix}
  a & b \\
  c & d
\end{vmatrix} \neq 0 \).

**EXAMPLE 2.5.4** The system

\[
\begin{align*}
7x + 8y &= 100 \\
2x - 9y &= 10
\end{align*}
\]

has the unique solution \( x = \Delta_1/\Delta, \ y = \Delta_2/\Delta \), where

\[
\Delta = \begin{vmatrix}
  7 & 8 \\
  2 & -9
\end{vmatrix} = -79, \ \Delta_1 = \begin{vmatrix}
  100 & 8 \\
  10 & -9
\end{vmatrix} = -980, \ \Delta_2 = \begin{vmatrix}
  7 & 100 \\
  2 & 10
\end{vmatrix} = -130.
\]

So \( x = \frac{980}{79} \) and \( y = \frac{130}{79} \).

**THEOREM 2.5.6** Let \( A \) be a square matrix. If \( A \) is non–singular, the homogeneous system \( AX = 0 \) has only the trivial solution. Equivalently, if the homogenous system \( AX = 0 \) has a non–trivial solution, then \( A \) is singular.

**Proof.** If \( A \) is non–singular and \( AX = 0 \), then \( X = A^{-1}0 = 0 \).

**REMARK 2.5.4** If \( A_{s1}, \ldots, A_{sn} \) denote the columns of \( A \), then the equation

\[
AX = x_1A_{s1} + \ldots + x_nA_{sn}
\]

holds. Consequently theorem 2.5.6 tells us that if there exist scalars \( x_1, \ldots, x_n \), not all zero, such that

\[
x_1A_{s1} + \ldots + x_nA_{sn} = 0,
\]
that is, if the columns of $A$ are linearly dependent, then $A$ is singular. An equivalent way of saying that the columns of $A$ are linearly dependent is that one of the columns of $A$ is expressible as a sum of certain scalar multiples of the remaining columns of $A$; that is one column is a linear combination of the remaining columns.

**EXAMPLE 2.5.5**

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

is singular. For it can be verified that $A$ has reduced row-echelon form

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and consequently $AX = 0$ has a non-trivial solution $x = -1$, $y = -1$, $z = 1$.

**REMARK 2.5.5** More generally, if $A$ is row-equivalent to a matrix containing a zero row, then $A$ is singular. For then the homogeneous system $AX = 0$ has a non-trivial solution.

An important class of non-singular matrices is that of the elementary row matrices.

**DEFINITION 2.5.2 (Elementary row matrices)** There are three types, $E_{ij}$, $E_i(t)$, $E_{ij}(t)$, corresponding to the three kinds of elementary row operation:

1. $E_{ij}$, $(i \neq j)$ is obtained from the identity matrix $I_n$ by interchanging rows $i$ and $j$.
2. $E_i(t)$, $(t \neq 0)$ is obtained by multiplying the $i$-th row of $I_n$ by $t$.
3. $E_{ij}(t)$, $(i \neq j)$ is obtained from $I_n$ by adding $t$ times the $j$-th row of $I_n$ to the $i$-th row.

**EXAMPLE 2.5.6** $(n = 3)$

$$E_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \ E_2(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ E_{23}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$
The elementary row matrices have the following distinguishing property:

**Theorem 2.5.7** If a matrix $A$ is pre-multiplied by an elementary row-matrix, the resulting matrix is the one obtained by performing the corresponding elementary row-operation on $A$.

**Example 2.5.7**

\[
E_{23} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ e & f \\ c & d \end{bmatrix}.
\]

**Corollary 2.5.2** The three types of elementary row-matrices are non-singular. Indeed

1. $E_{ij}^{-1} = E_{ij}$;
2. $E_i^{-1}(t) = E_i(t^{-1})$;
3. $(E_{ij}(t))^{-1} = E_{ij}(-t)$.

**Proof.** Taking $A = I_n$ in the above theorem, we deduce the following equations:

\[
E_{ij}E_{ij} = I_n, \\
E_i(t)E_i(t^{-1}) = I_n = E_i(t^{-1})E_i(t) \quad \text{if } t \neq 0, \\
E_{ij}(t)E_{ij}(-t) = I_n = E_{ij}(-t)E_{ij}(t).
\]

**Example 2.5.8** Find the $3 \times 3$ matrix $A = E_3(5)E_{23}(2)E_{12}$ explicitly. Also find $A^{-1}$.

**Solution.**

\[
A = E_3(5)E_{23}(2) \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_3(5) \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 5 \end{bmatrix}.
\]

To find $A^{-1}$, we have

\[
A^{-1} = (E_3(5)E_{23}(2)E_{12})^{-1} = E_{12}^{-1}(E_{23}(2))^{-1}(E_3(5))^{-1} = E_{12}E_{23}(-2)E_3(5^{-1}).
\]
2.5. NON–SINGULAR MATRICES

\[
= E_{12}E_{23}(-2) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix} \\
= E_{12} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{-2}{5} \\ 0 & 0 & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 0 & 1 & \frac{-2}{5} \\ 1 & 0 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}.
\]

**REMARK 2.5.6** Recall that \( A \) and \( B \) are row–equivalent if \( B \) is obtained from \( A \) by a sequence of elementary row operations. If \( E_1, \ldots, E_r \) are the respective corresponding elementary row matrices, then

\[
B = E_r \cdots (E_2(E_1A)) \cdots = (E_r \cdots E_1)A = PA,
\]

where \( P = E_r \cdots E_1 \) is non–singular. Conversely if \( B = PA \), where \( P \) is non–singular, then \( A \) is row–equivalent to \( B \). For as we shall now see, \( P \) is in fact a product of elementary row matrices.

**THEOREM 2.5.8** Let \( A \) be non–singular \( n \times n \) matrix. Then

(i) \( A \) is row–equivalent to \( I_n \),

(ii) \( A \) is a product of elementary row matrices.

**Proof.** Assume that \( A \) is non–singular and let \( B \) be the reduced row–echelon form of \( A \). Then \( B \) has no zero rows, for otherwise the equation \( AX = 0 \) would have a non–trivial solution. Consequently \( B = I_n \).

It follows that there exist elementary row matrices \( E_1, \ldots, E_r \) such that \( E_r \cdots (E_1A) \cdots = B = I_n \) and hence \( A = E_r^{-1} \cdots E_1^{-1} \), a product of elementary row matrices.

**THEOREM 2.5.9** Let \( A \) be \( n \times n \) and suppose that \( A \) is row–equivalent to \( I_n \). Then \( A \) is non–singular and \( A^{-1} \) can be found by performing the same sequence of elementary row operations on \( I_n \) as were used to convert \( A \) to \( I_n \).

**Proof.** Suppose that \( E_r \cdots E_1A = I_n \). In other words \( BA = I_n \), where \( B = E_r \cdots E_1 \) is non–singular. Then \( B^{-1}(BA) = B^{-1}I_n \) and so \( A = B^{-1} \), which is non–singular.

Also \( A^{-1} = (B^{-1})^{-1} = B = E_r \cdots (E_1I_n) \cdots \), which shows that \( A^{-1} \) is obtained from \( I_n \) by performing the same sequence of elementary row operations as were used to convert \( A \) to \( I_n \).
**REMARK 2.5.7** It follows from theorem 2.5.9 that if $A$ is singular, then $A$ is row–equivalent to a matrix whose last row is zero.

**EXAMPLE 2.5.9** Show that $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ is non–singular, find $A^{-1}$ and express $A$ as a product of elementary row matrices.

**Solution.** We form the *partitioned* matrix $[A|I_2]$ which consists of $A$ followed by $I_2$. Then any sequence of elementary row operations which reduces $A$ to $I_2$ will reduce $I_2$ to $A^{-1}$. Here

$$[A|I_2] = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

$$R_2 \rightarrow R_2 - R_1 \quad \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix}$$

$$R_2 \rightarrow (-1)R_2 \quad \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

$$R_1 \rightarrow R_1 - 2R_2 \quad \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{bmatrix}.$$  

Hence $A$ is row–equivalent to $I_2$ and $A$ is non–singular. Also

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

We also observe that

$$E_{12}(-2)E_2(-1)E_{21}(-1)A = I_2.$$  

Hence

$$A^{-1} = E_{12}(-2)E_2(-1)E_{21}(-1)$$

$$A = E_{21}(1)E_2(-1)E_{12}(2).$$

The next result is the converse of Theorem 2.5.6 and is useful for proving the non–singularity of certain types of matrices.

**THEOREM 2.5.10** Let $A$ be an $n \times n$ matrix with the property that the homogeneous system $AX = 0$ has only the trivial solution. Then $A$ is non–singular. Equivalently, if $A$ is singular, then the homogeneous system $AX = 0$ has a non–trivial solution.
2.5. NON–SINGULAR MATRICES

Proof. If $A$ is $n \times n$ and the homogeneous system $AX = 0$ has only the trivial solution, then it follows that the reduced row–echelon form $B$ of $A$ cannot have zero rows and must therefore be $I_n$. Hence $A$ is non–singular.

**COROLLARY 2.5.3** Suppose that $A$ and $B$ are $n \times n$ and $AB = I_n$. Then $BA = I_n$.

**Proof.** Let $AB = I_n$, where $A$ and $B$ are $n \times n$. We first show that $B$ is non–singular. Assume $BX = 0$. Then $A(BX) = A0 = 0$, so $(AB)X = 0$, $I_nX = 0$ and hence $X = 0$.

Then from $AB = I_n$ we deduce $(AB)B^{-1} = I_nB^{-1}$ and hence $A = B^{-1}$. The equation $BB^{-1} = I_n$ then gives $BA = I_n$.

Before we give the next example of the above criterion for non-singularity, we introduce an important matrix operation.

**DEFINITION 2.5.3 (The transpose of a matrix)** Let $A$ be an $m \times n$ matrix. Then $A^t$, the transpose of $A$, is the matrix obtained by interchanging the rows and columns of $A$. In other words if $A = [a_{ij}]$, then $(A^t)_{ji} = a_{ij}$. Consequently $A^t$ is $n \times m$.

The transpose operation has the following properties:

1. $(A^t)^t = A$;
2. $(A \pm B)^t = A^t \pm B^t$ if $A$ and $B$ are $m \times n$;
3. $(sA)^t = sA^t$ if $s$ is a scalar;
4. $(AB)^t = B^tA^t$ if $A$ is $m \times n$ and $B$ is $n \times p$;
5. If $A$ is non–singular, then $A^t$ is also non–singular and 
   $$(A^t)^{-1} = (A^{-1})^t;$$
6. $X^tX = x_1^2 + \ldots + x_n^2$ if $X = [x_1, \ldots, x_n]^t$ is a column vector.

We prove only the fourth property. First check that both $(AB)^t$ and $B^tA^t$ have the same size $(p \times m)$. Moreover, corresponding elements of both matrices are equal. For if $A = [a_{ij}]$ and $B = [b_{jk}]$, we have

$$((AB)^t)_{ki} = (AB)_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$$
There are two important classes of matrices that can be defined concisely in terms of the transpose operation.

**DEFINITION 2.5.4 (Symmetric matrix)** A real matrix \( A \) is called symmetric if \( A^t = A \). In other words \( A \) is square (\( n \times n \) say) and \( a_{ji} = a_{ij} \) for all \( 1 \leq i \leq n, 1 \leq j \leq n \). Hence

\[
A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}
\]

is a general \( 2 \times 2 \) symmetric matrix.

**DEFINITION 2.5.5 (Skew-symmetric matrix)** A real matrix \( A \) is called skew-symmetric if \( A^t = -A \). In other words \( A \) is square (\( n \times n \) say) and \( a_{ji} = -a_{ij} \) for all \( 1 \leq i \leq n, 1 \leq j \leq n \).

**REMARK 2.5.8** Taking \( i = j \) in the definition of skew-symmetric matrix gives \( a_{ii} = -a_{ii} \) and so \( a_{ii} = 0 \). Hence

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

is a general \( 2 \times 2 \) skew-symmetric matrix.

We can now state a second application of the above criterion for non-singularity.

**COROLLARY 2.5.4** Let \( B \) be an \( n \times n \) skew-symmetric matrix. Then \( A = I_n - B \) is non-singular.

**Proof.** Let \( A = I_n - B \), where \( B^t = -B \). By Theorem 2.5.10 it suffices to show that \( AX = 0 \) implies \( X = 0 \).

We have \((I_n - B)X = 0\), so \( X = BX \). Hence \( X^tX = X^tBX \).

Taking transposes of both sides gives

\[
(X^tBX)^t = (X^tX)^t \\
X^tB^t(X^t)^t = X^t(X^t)^t \\
X^t(-B)X = X^tX \\
-X^tBX = X^tX = X^tBX.
\]

Hence \( X^tX = -X^tX \) and \( X^tX = 0 \). But if \( X = [x_1, \ldots, x_n]^t \), then \( X^tX = x_1^2 + \ldots + x_n^2 = 0 \) and hence \( x_1 = 0, \ldots, x_n = 0 \).
2.6 Least squares solution of equations

Suppose $AX = B$ represents a system of linear equations with real coefficients which may be inconsistent, because of the possibility of experimental errors in determining $A$ or $B$. For example, the system
\[
\begin{align*}
x &= 1 \\
y &= 2 \\
x + y &= 3.001
\end{align*}
\]
is inconsistent.

It can be proved that the associated system $A^t AX = A^t B$ is always consistent and that any solution of this system minimizes the sum $r_1^2 + \ldots + r_m^2$, where $r_1, \ldots, r_m$ (the residuals) are defined by
\[
r_i = a_{i1}x_1 + \ldots + a_{in}x_n - b_i,
\]
for $i = 1, \ldots, m$. The equations represented by $A^t AX = A^t B$ are called the normal equations corresponding to the system $AX = B$ and any solution of the system of normal equations is called a least squares solution of the original system.

**EXAMPLE 2.6.1** Find a least squares solution of the above inconsistent system.

**Solution.** Here $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 2 \\ 3.001 \end{bmatrix}$.

Then $A^t A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

Also $A^t B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3.001 \end{bmatrix} = \begin{bmatrix} 4.001 \\ 5.001 \end{bmatrix}$.

So the normal equations are
\[
\begin{align*}
2x + y &= 4.001 \\
x + 2y &= 5.001
\end{align*}
\]
which have the unique solution
\[
x = \frac{3.001}{3}, \quad y = \frac{6.001}{3}.
\]
EXAMPLE 2.6.2 Points \((x_1, y_1), \ldots, (x_n, y_n)\) are experimentally determined and should lie on a line \(y = mx + c\). Find a least squares solution to the problem.

Solution. The points have to satisfy

\[
mx_1 + c = y_1 \\
\vdots \\
mx_n + c = y_n,
\]

or \(Ax = B\), where

\[
A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad X = \begin{bmatrix} m \\ c \end{bmatrix}, \quad B = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.
\]

The normal equations are given by \((A^tA)X = A^tB\). Here

\[
A^tA = \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} x_1^2 + \cdots + x_n^2 & x_1 + \cdots + x_n \\ x_1 + \cdots + x_n & n \end{bmatrix}
\]

Also

\[
A^tB = \begin{bmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ y_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 + \cdots + x_ny_n \\ y_1 + \cdots + y_n \end{bmatrix}.
\]

It is not difficult to prove that

\[
\Delta = \det(A^tA) = \sum_{1\leq i<j\leq n} (x_i - x_j)^2,
\]

which is positive unless \(x_1 = \ldots = x_n\). Hence if not all of \(x_1, \ldots, x_n\) are equal, \(A^tA\) is non-singular and the normal equations have a unique solution. This can be shown to be

\[
m = \frac{1}{\Delta} \sum_{1\leq i<j\leq n} (x_i - x_j)(y_i - y_j), \quad c = \frac{1}{\Delta} \sum_{1\leq i<j\leq n} (x_iy_j - x_jy_i)(x_i - x_j).
\]

REMARK 2.6.1 The matrix \(A^tA\) is symmetric.
2.7 PROBLEMS

1. Let \( A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix} \). Prove that \( A \) is non–singular, find \( A^{-1} \) and express \( A \) as a product of elementary row matrices.

[Answer: \( A^{-1} = \begin{bmatrix} \frac{1}{13} & -\frac{4}{13} \\ \frac{3}{13} & \frac{1}{13} \end{bmatrix} \),

\( A = E_{21}(-3)E_{2}(13)E_{12}(4) \) is one such decomposition.]

2. A square matrix \( D = [d_{ij}] \) is called diagonal if \( d_{ij} = 0 \) for \( i \neq j \). (That is the off–diagonal elements are zero.) Prove that pre–multiplication of a matrix \( A \) by a diagonal matrix \( D \) results in matrix \( DA \) whose rows are the rows of \( A \) multiplied by the respective diagonal elements of \( D \). State and prove a similar result for post–multiplication by a diagonal matrix.

Let diag \( (a_1, \ldots, a_n) \) denote the diagonal matrix whose diagonal elements \( d_{ii} \) are \( a_1, \ldots, a_n \), respectively. Show that

\[
\text{diag}(a_1, \ldots, a_n) \text{diag}(b_1, \ldots, b_n) = \text{diag}(a_1 b_1, \ldots, a_n b_n)
\]

and deduce that if \( a_1 \ldots a_n \neq 0 \), then \( \text{diag}(a_1, \ldots, a_n) \) is non–singular and

\[
(\text{diag}(a_1, \ldots, a_n))^{-1} = \text{diag}(a_1^{-1}, \ldots, a_n^{-1}).
\]

Also prove that \( \text{diag}(a_1, \ldots, a_n) \) is singular if \( a_i = 0 \) for some \( i \).

3. Let \( A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 2 & 6 \\ 3 & 7 & 9 \end{bmatrix} \). Prove that \( A \) is non–singular, find \( A^{-1} \) and express \( A \) as a product of elementary row matrices.

[Answers: \( A^{-1} = \begin{bmatrix} -12 & 7 & -2 \\ 9 & -3 & 1 \\ 2 & 0 & 0 \end{bmatrix} \),

\( A = E_{12}E_{31}(3)E_{23}E_{3}(2)E_{12}(2)E_{13}(24)E_{23}(-9) \) is one such decomposition.]
4. Find the rational number \( k \) for which the matrix
\[
A = \begin{bmatrix}
1 & 2 & k \\
3 & -1 & 1 \\
5 & 3 & -5
\end{bmatrix}
\]
is singular. [Answer: \( k = -3 \).]

5. Prove that \( A = \begin{bmatrix} 1 & 2 \\ -2 & 2 \end{bmatrix} \) is singular and find a non–singular matrix \( P \) such that \( PA \) has last row zero.

6. If \( A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \end{bmatrix} \), verify that \( A^2 - 2A + 13I_2 = 0 \) and deduce that \( A^{-1} = -\frac{1}{13}(A - 2I_2) \).

7. Let \( A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix} \).
   
   (i) Verify that \( A^3 = 3A^2 - 3A + I_3 \).

   (ii) Express \( A^4 \) in terms of \( A^2 \), \( A \) and \( I_3 \) and hence calculate \( A^4 \) explicitly.

   (iii) Use (i) to prove that \( A \) is non–singular and find \( A^{-1} \) explicitly.

   [Answers: (ii) \( A^4 = 6A^2 - 8A + 3I_3 = \begin{bmatrix} -11 & -8 & -4 \\ 12 & 9 & 4 \\ 20 & 16 & 5 \end{bmatrix} \);

   (iii) \( A^{-1} = A^2 - 3A + 3I_3 = \begin{bmatrix} -1 & -3 & 1 \\ 2 & 4 & -1 \\ 0 & 1 & 0 \end{bmatrix} \).]

8. (i) Let \( B \) be an \( n \times n \) matrix such that \( B^3 = 0 \). If \( A = I_n - B \), prove that \( A \) is non–singular and \( A^{-1} = I_n + B + B^2 \).

   Show that the system of linear equations \( AX = b \) has the solution

   \[
   X = b + Bb + B^2b.
   \]

   (ii) If \( B = \begin{bmatrix} 0 & r & s \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix} \), verify that \( B^3 = 0 \) and use (i) to determine \((I_3 - B)^{-1}\) explicitly.
2.7. PROBLEMS

[Answer: \[
\begin{bmatrix}
1 & r & s + rt \\
0 & 1 & t \\
0 & 0 & 1
\end{bmatrix}.
\]

9. Let \(A\) be \(n \times n\).

(i) If \(A^2 = 0\), prove that \(A\) is singular.

(ii) If \(A^2 = A\) and \(A \neq I_n\), prove that \(A\) is singular.

10. Use Question 7 to solve the system of equations

\[
\begin{align*}
x + y - z &= a \\
z &= b \\
2x + y + 2z &= c
\end{align*}
\]

where \(a, b, c\) are given rationals. Check your answer using the Gauss–Jordan algorithm.

[Answer: \(x = -a - 3b + c, y = 2a + 4b - c, z = b\).]

11. Determine explicitly the following products of \(3 \times 3\) elementary row matrices.

(i) \(E_{12}E_{23}\)  (ii) \(E_1(5)E_{12}\)  (iii) \(E_{12}(3)E_{21}(-3)\)  (iv) \((E_1(100))^{-1}\)

(v) \(E_{12}^{-1}\)  (vi) \((E_{12}(7))^{-1}\)  (vii) \((E_{12}(7)E_{31}(1))^{-1}\).

[Answers: (i) \[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\]

(ii) \[
\begin{bmatrix}
0 & 5 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(iii) \[
\begin{bmatrix}
-8 & 3 & 0 \\
-3 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(iv) \[
\begin{bmatrix}
\frac{1}{100} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(v) \[
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(vi) \[
\begin{bmatrix}
1 & -7 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(vii) \[
\begin{bmatrix}
1 & -7 & 0 \\
0 & 1 & 0 \\
-1 & 7 & 1
\end{bmatrix}
\]

12. Let \(A\) be the following product of \(4 \times 4\) elementary row matrices:

\[A = E_{32}(2)E_{14}E_{42}(3)\].

Find \(A\) and \(A^{-1}\) explicitly.

[Answers: \[
A = \begin{bmatrix}
0 & 3 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 2 & 0 \\
1 & 0 & 0 & 0
\end{bmatrix}, \quad A^{-1} = \begin{bmatrix}
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 \\
1 & -3 & 0 & 0
\end{bmatrix}.
\]
13. Determine which of the following matrices over $\mathbb{Z}_2$ are non–singular and find the inverse, where possible.

(a) $$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

(b) $$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

[Answer: (a) $$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$]

14. Determine which of the following matrices are non–singular and find the inverse, where possible.

(a) $$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

(b) $$\begin{bmatrix} 2 & 2 & 4 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

(c) $$\begin{bmatrix} 4 & 6 & -3 \\ 0 & 0 & 7 \\ 0 & 0 & 5 \end{bmatrix}$$

(d) $$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

(e) $$\begin{bmatrix} 1 & 2 & 4 & 6 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

(f) $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

[Answers: (a) $$\begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 1 & -1 & -1 \end{bmatrix}$$

(b) $$\begin{bmatrix} -\frac{1}{2} & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

(d) $$\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \end{bmatrix}$$

(e) $$\begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$]

15. Let $A$ be a non–singular $n \times n$ matrix. Prove that $A^t$ is non–singular and that $(A^t)^{-1} = (A^{-1})^t$.

16. Prove that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has no inverse if $ad - bc = 0$.

[Hint: Use the equation $A^2 - (a + d)A + (ad - bc)I_2 = 0$.]
17. Prove that the real matrix  \( A = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix} \) is non-singular by proving that \( A \) is row-equivalent to \( I_3 \).

18. If \( P^{-1} AP = B \), prove that \( P^{-1} A^n P = B^n \) for \( n \geq 1 \).

19. Let \( A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \), \( P = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix} \). Verify that \( P^{-1} AP = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \) and deduce that

\[
A^n = \frac{1}{7} \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix} + \frac{1}{7} \left( \frac{5}{12} \right)^n \begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix}.
\]

20. Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) be a Markov matrix; that is a matrix whose elements are non-negative and satisfy \( a+c = 1 = b+d \). Also let \( P = \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix} \). Prove that if \( A \neq I_2 \) then

(i) \( P \) is non-singular and \( P^{-1} AP = \begin{bmatrix} 1 & 0 \\ 0 & a+d-1 \end{bmatrix} \),

(ii) \( A^n \rightarrow \frac{1}{b+c} \begin{bmatrix} b & b \\ c & c \end{bmatrix} \) as \( n \rightarrow \infty \), if \( A \neq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \).

21. If \( X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \) and \( Y = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} \), find \( XX^t, X^tX, YY^t, Y^tY \).

[Answers: \( \begin{bmatrix} 5 & 11 & 17 \\ 11 & 25 & 39 \\ 17 & 39 & 61 \end{bmatrix}, \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}, \begin{bmatrix} 1 & -3 & -4 \\ -3 & 9 & 12 \\ -4 & 12 & 16 \end{bmatrix}, 26. \)]

22. Prove that the system of linear equations

\[
\begin{align*}
  x + 2y &= 4 \\
  x + y &= 5 \\
  3x + 5y &= 12
\end{align*}
\]

is inconsistent and find a least squares solution of the system.

[Answer: \( x = 6, y = -7/6. \)]
23. The points (0, 0), (1, 0), (2, −1), (3, 4), (4, 8) are required to lie on a parabola $y = a + bx + cx^2$. Find a least squares solution for $a$, $b$, $c$. Also prove that no parabola passes through these points.

[Answer: $a = \frac{1}{5}$, $b = -2$, $c = 1$.]

24. If $A$ is a symmetric $n \times n$ real matrix and $B$ is $n \times m$, prove that $B^T A B$ is a symmetric $m \times m$ matrix.

25. If $A$ is $m \times n$ and $B$ is $n \times m$, prove that $A B$ is singular if $m > n$.

26. Let $A$ and $B$ be $n \times n$. If $A$ or $B$ is singular, prove that $A B$ is also singular.
Chapter 3

SUBSPACES

3.1 Introduction

Throughout this chapter, we will be studying $\mathbb{F}^n$, the set of all $n$-dimensional column vectors with components from a field $F$. We continue our study of matrices by considering an important class of subsets of $\mathbb{F}^n$ called subspaces. These arise naturally for example, when we solve a system of $m$ linear homogeneous equations in $n$ unknowns.

We also study the concept of linear dependence of a family of vectors. This was introduced briefly in Chapter 2, Remark 2.5.4. Other topics discussed are the row space, column space and null space of a matrix over $F$, the dimension of a subspace, particular examples of the latter being the rank and nullity of a matrix.

3.2 Subspaces of $\mathbb{F}^n$

**DEFINITION 3.2.1** A subset $S$ of $\mathbb{F}^n$ is called a subspace of $\mathbb{F}^n$ if

1. The zero vector belongs to $S$; (that is, $0 \in S$);
2. If $u \in S$ and $v \in S$, then $u + v \in S$; ($S$ is said to be closed under vector addition);
3. If $u \in S$ and $t \in F$, then $tu \in S$; ($S$ is said to be closed under scalar multiplication).

**EXAMPLE 3.2.1** Let $A \in M_{m \times n}(F)$. Then the set of vectors $X \in \mathbb{F}^n$ satisfying $AX = 0$ is a subspace of $\mathbb{F}^n$ called the null space of $A$ and is denoted here by $N(A)$. (It is sometimes called the solution space of $A$.)
CHAPTER 3. SUBSPACES

Proof. (1) \( A0 = 0 \), so \( 0 \in N(A) \); (2) If \( X, Y \in N(A) \), then \( AX = 0 \) and \( AY = 0 \), so \( A(X + Y) = AX + AY = 0 + 0 = 0 \) and so \( X + Y \in N(A) \); (3) If \( X \in N(A) \) and \( t \in F \), then \( A(tX) = t(AX) = t0 = 0 \), so \( tX \in N(A) \).

For example, if \( A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \), then \( N(A) = \{0\} \), the set consisting of just the zero vector. If \( A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \), then \( N(A) \) is the set of all scalar multiples of \([-2, 1]^t\).

**EXAMPLE 3.2.2** Let \( X_1, \ldots, X_m \in F^n \). Then the set consisting of all linear combinations \( x_1X_1 + \cdots + x_mX_m \), where \( x_1, \ldots, x_m \in F \), is a subspace of \( F^m \). This subspace is called the subspace spanned or generated by \( X_1, \ldots, X_m \) and is denoted here by \( \langle X_1, \ldots, X_m \rangle \). We also call \( X_1, \ldots, X_m \) a spanning family for \( S = \langle X_1, \ldots, X_m \rangle \).

Proof. (1) \( 0 = 0X_1 + \cdots + 0X_m \), so \( 0 \in \langle X_1, \ldots, X_m \rangle \); (2) If \( X, Y \in \langle X_1, \ldots, X_m \rangle \), then \( X = x_1X_1 + \cdots + x_mX_m \) and \( Y = y_1X_1 + \cdots + y_mX_m \), so

\[
X + Y = (x_1X_1 + \cdots + x_mX_m) + (y_1X_1 + \cdots + y_mX_m) \\
= (x_1 + y_1)X_1 + \cdots + (x_m + y_m)X_m \in \langle X_1, \ldots, X_m \rangle.
\]

(3) If \( X \in \langle X_1, \ldots, X_m \rangle \) and \( t \in F \), then

\[
X = x_1X_1 + \cdots + x_mX_m \\
tX = t(x_1X_1 + \cdots + x_mX_m) \\
= (tx_1)X_1 + \cdots + (tx_m)X_m \in \langle X_1, \ldots, X_m \rangle.
\]

For example, if \( A \in M_{m \times n}(F) \), the subspace generated by the columns of \( A \) is an important subspace of \( F^m \) and is called the column space of \( A \). The column space of \( A \) is denoted here by \( C(A) \). Also the subspace generated by the rows of \( A \) is a subspace of \( F^n \) and is called the row space of \( A \) and is denoted by \( R(A) \).

**EXAMPLE 3.2.3** For example \( F^n = \langle E_1, \ldots, E_m \rangle \), where \( E_1, \ldots, E_n \) are the \( n \)-dimensional unit vectors. For if \( X = [x_1, \ldots, x_n]^t \in F^n \), then \( X = x_1E_1 + \cdots + x_nE_n \).

**EXAMPLE 3.2.4** Find a spanning family for the subspace \( S \) of \( \mathbb{R}^3 \) defined by the equation \( 2x - 3y + 5z = 0 \).
3.2. SUBSPACES OF $F^N$

Solution. ($S$ is in fact the null space of $[2, -3, 5]$, so $S$ is indeed a subspace of $\mathbb{R}^3$.)

If $[x, y, z]^t \in S$, then $x = \frac{3}{2}y - \frac{5}{2}z$. Then

$$
\begin{bmatrix}
  x \\
  y \\
  z 
\end{bmatrix} = \begin{bmatrix}
  \frac{3}{2}y - \frac{5}{2}z \\
  y \\
  z 
\end{bmatrix} = \begin{bmatrix}
  \frac{3}{2} \\
  1 \\
  0 
\end{bmatrix} + z \begin{bmatrix}
  -\frac{5}{2} \\
  0 \\
  1 
\end{bmatrix}
$$

and conversely. Hence $[\frac{3}{2}, 1, 0]^t$ and $[-\frac{5}{2}, 0, 1]^t$ form a spanning family for $S$.

The following result is easy to prove:

**Lemma 3.2.1** Suppose each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$. Then any linear combination of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$.

As a corollary we have

**Theorem 3.2.1** Subspaces $\langle X_1, \ldots, X_r \rangle$ and $\langle Y_1, \ldots, Y_s \rangle$ are equal if each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$ and each of $Y_1, \ldots, Y_s$ is a linear combination of $X_1, \ldots, X_r$.

**Corollary 3.2.1** Subspaces $\langle X_1, \ldots, X_r, Z_1, \ldots, Z_t \rangle$ and $\langle X_1, \ldots, X_r \rangle$ are equal if each of $Z_1, \ldots, Z_t$ is a linear combination of $X_1, \ldots, X_r$.

**Example 3.2.5** If $X$ and $Y$ are vectors in $\mathbb{R}^n$, then

$$
\langle X, Y \rangle = \langle X + Y, X - Y \rangle.
$$

Solution. Each of $X + Y$ and $X - Y$ is a linear combination of $X$ and $Y$. Also

$$
X = \frac{1}{2}(X + Y) + \frac{1}{2}(X - Y) \quad \text{and} \quad Y = \frac{1}{2}(X + Y) - \frac{1}{2}(X - Y),
$$

so each of $X$ and $Y$ is a linear combination of $X + Y$ and $X - Y$.

There is an important application of Theorem 3.2.1 to row equivalent matrices (see Definition 1.2.4):

**Theorem 3.2.2** If $A$ is row equivalent to $B$, then $R(A) = R(B)$.

Proof. Suppose that $B$ is obtained from $A$ by a sequence of elementary row operations. Then it is easy to see that each row of $B$ is a linear combination of the rows of $A$. But $A$ can be obtained from $B$ by a sequence of elementary operations, so each row of $A$ is a linear combination of the rows of $B$. Hence by Theorem 3.2.1, $R(A) = R(B)$. 
CHAPTER 3. SUBSPACES

REMARK 3.2.1 If $A$ is row equivalent to $B$, it is not always true that $C(A) = C(B)$.

For example, if $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, then $B$ is in fact the reduced row–echelon form of $A$. However we see that

$$C(A) = \langle \left[ \begin{array}{c} 1 \\ 1 \end{array} \right], \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] \rangle = \langle \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] \rangle$$

and similarly $C(B) = \langle \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] \rangle$.

Consequently $C(A) \neq C(B)$, as $\left[ \begin{array}{c} 1 \\ 1 \end{array} \right] \in C(A)$ but $\left[ \begin{array}{c} 1 \\ 1 \end{array} \right] \notin C(B)$.

3.3 Linear dependence

We now recall the definition of linear dependence and independence of a family of vectors in $F^n$ given in Chapter 2.

DEFINITION 3.3.1 Vectors $X_1, \ldots, X_m$ in $F^n$ are said to be linearly dependent if there exist scalars $x_1, \ldots, x_m$, not all zero, such that

$$x_1X_1 + \cdots + x_mX_m = 0.$$ 

In other words, $X_1, \ldots, X_m$ are linearly dependent if some $X_i$ is expressible as a linear combination of the remaining vectors.

$X_1, \ldots, X_m$ are called linearly independent if they are not linearly dependent. Hence $X_1, \ldots, X_m$ are linearly independent if and only if the equation

$$x_1X_1 + \cdots + x_mX_m = 0$$

has only the trivial solution $x_1 = 0, \ldots, x_m = 0$.

EXAMPLE 3.3.1 The following three vectors in $\mathbb{R}^3$

$$X_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ 7 \\ 12 \end{bmatrix}$$

are linearly dependent, as $2X_1 + 3X_2 + (-1)X_3 = 0$. 
3.3. LINEAR DEPENDENCE

REMARK 3.3.1 If $X_1, \ldots, X_m$ are linearly independent and

$$x_1X_1 + \cdots + x_mX_m = y_1X_1 + \cdots + y_mX_m,$$

then $x_1 = y_1, \ldots, x_m = y_m$. For the equation can be rewritten as

$$(x_1 - y_1)X_1 + \cdots + (x_m - y_m)X_m = 0$$

and so $x_1 - y_1 = 0, \ldots, x_m - y_m = 0$.

THEOREM 3.3.1 A family of $m$ vectors in $F^n$ will be linearly dependent if $m > n$. Equivalently, any linearly independent family of $m$ vectors in $F^n$ must satisfy $m \leq n$.

Proof. The equation

$$x_1X_1 + \cdots + x_mX_m = 0$$

is equivalent to $n$ homogeneous equations in $m$ unknowns. By Theorem 1.5.1, such a system has a non–trivial solution if $m > n$.

The following theorem is an important generalization of the last result and is left as an exercise for the interested student:

THEOREM 3.3.2 A family of $s$ vectors in $\langle X_1, \ldots, X_r \rangle$ will be linearly dependent if $s > r$. Equivalently, a linearly independent family of $s$ vectors in $\langle X_1, \ldots, X_r \rangle$ must have $s \leq r$.

Here is a useful criterion for linear independence which is sometimes called the left–to–right test:

THEOREM 3.3.3 Vectors $X_1, \ldots, X_m$ in $F^n$ are linearly independent if

(a) $X_1 \neq 0$;

(b) For each $k$ with $1 < k \leq m$, $X_k$ is not a linear combination of $X_1, \ldots, X_{k-1}$.

One application of this criterion is the following result:

THEOREM 3.3.4 Every subspace $S$ of $F^n$ can be represented in the form $S = \langle X_1, \ldots, X_m \rangle$, where $m \leq n$. 
Proof. If $S = \{0\}$, there is nothing to prove – we take $X_1 = 0$ and $m = 1$.

So we assume $S$ contains a non–zero vector $X_1$; then $\langle X_1 \rangle \subseteq S$ as $S$ is a subspace. If $S = \langle X_1 \rangle$, we are finished. If not, $S$ will contain a vector $X_2$, not a linear combination of $X_1$; then $\langle X_1, X_2 \rangle \subseteq S$ as $S$ is a subspace. If $S = \langle X_1, X_2 \rangle$, we are finished. If not, $S$ will contain a vector $X_3$ which is not a linear combination of $X_1$ and $X_2$. This process must eventually stop, for at stage $k$ we have constructed a family of $k$ linearly independent vectors $X_1, \ldots, X_k$, all lying in $F^n$ and hence $k \leq n$.

There is an important relationship between the columns of $A$ and $B$, if $A$ is row–equivalent to $B$.

**THEOREM 3.3.5** Suppose that $A$ is row equivalent to $B$ and let $c_1, \ldots, c_r$ be distinct integers satisfying $1 \leq c_i \leq n$. Then

(a) Columns $A_{sc_1}, \ldots, A_{sc_r}$ of $A$ are linearly dependent if and only if the corresponding columns of $B$ are linearly dependent; indeed more is true:

$$x_1A_{sc_1} + \cdots + x_rA_{sc_r} = 0 \iff x_1B_{sc_1} + \cdots + x_rB_{sc_r} = 0.$$ 

(b) Columns $A_{sc_1}, \ldots, A_{sc_r}$ of $A$ are linearly independent if and only if the corresponding columns of $B$ are linearly independent.

(c) If $1 \leq c_{r+1} \leq n$ and $c_{r+1}$ is distinct from $c_1, \ldots, c_r$, then

$$A_{sc_{r+1}} = z_1A_{sc_1} + \cdots + z_rA_{sc_r} \iff B_{sc_{r+1}} = z_1B_{sc_1} + \cdots + z_rB_{sc_r}.$$ 

Proof. First observe that if $Y = [y_1, \ldots, y_n]^t$ is an $n$–dimensional column vector and $A$ is $m \times n$, then

$$AY = y_1A_{sc_1} + \cdots + y_nA_{sc_n}.$$ 

Also $AY = 0 \iff BY = 0$, if $B$ is row equivalent to $A$. Then (a) follows by taking $y_i = x_{cj}$ if $i = c_j$ and $y_i = 0$ otherwise.

(b) is logically equivalent to (a), while (c) follows from (a) as

$$A_{sc_{r+1}} = z_1A_{sc_1} + \cdots + z_rA_{sc_r} \iff z_1A_{sc_1} + \cdots + z_rA_{sc_r} + (-1)A_{sc_{r+1}} = 0$$

$$\iff z_1B_{sc_1} + \cdots + z_rB_{sc_r} + (-1)B_{sc_{r+1}} = 0$$

$$\iff B_{sc_{r+1}} = z_1B_{sc_1} + \cdots + z_rB_{sc_r}.$$
EXAMPLE 3.3.2 The matrix
\[
A = \begin{bmatrix}
1 & 1 & 5 & 1 & 4 \\
2 & -1 & 1 & 2 & 2 \\
3 & 0 & 6 & 0 & -3
\end{bmatrix}
\]
has reduced row–echelon form equal to
\[
B = \begin{bmatrix}
1 & 0 & 2 & 0 & -1 \\
0 & 1 & 3 & 0 & 2 \\
0 & 0 & 0 & 1 & 3
\end{bmatrix}.
\]
We notice that \( B_{*1}, B_{*2} \) and \( B_{*4} \) are linearly independent and hence so are \( A_{*1}, A_{*2} \) and \( A_{*4} \). Also
\[
\begin{align*}
B_{*3} &= 2B_{*1} + 3B_{*2} \\
B_{*5} &= (-1)B_{*1} + 2B_{*2} + 3B_{*4},
\end{align*}
\]
so consequently
\[
\begin{align*}
A_{*3} &= 2A_{*1} + 3A_{*2} \\
A_{*5} &= (-1)A_{*1} + 2A_{*2} + 3A_{*4}.
\end{align*}
\]

3.4 Basis of a subspace

We now come to the important concept of basis of a vector subspace.

DEFINITION 3.4.1 Vectors \( X_1, \ldots, X_m \) belonging to a subspace \( S \) are said to form a basis of \( S \) if

(a) Every vector in \( S \) is a linear combination of \( X_1, \ldots, X_m \);

(b) \( X_1, \ldots, X_m \) are linearly independent.

Note that (a) is equivalent to the statement that \( S = \langle X_1, \ldots, X_m \rangle \) as we automatically have \( \langle X_1, \ldots, X_m \rangle \subseteq S \). Also, in view of Remark 3.3.1 above, (a) and (b) are equivalent to the statement that every vector in \( S \) is uniquely expressible as a linear combination of \( X_1, \ldots, X_m \).

EXAMPLE 3.4.1 The unit vectors \( E_1, \ldots, E_n \) form a basis for \( F^n \).
REMARK 3.4.1 The subspace \{0\}, consisting of the zero vector alone, does not have a basis. For every vector in a linearly independent family must necessarily be non–zero. (For example, if \(X_1 = 0\), then we have the non–trivial linear relation
\[
1X_1 + 0X_2 + \cdots + 0X_m = 0
\]
and \(X_1, \ldots, X_m\) would be linearly independent.)

However if we exclude this case, every other subspace of \(F^n\) has a basis:

THEOREM 3.4.1 A subspace of the form \(\langle X_1, \ldots, X_m \rangle\), where at least one of \(X_1, \ldots, X_m\) is non–zero, has a basis \(X_{c_1}, \ldots, X_{c_r}\), where \(1 \leq c_1 < \cdots < c_r \leq m\).

Proof. (The left–to–right algorithm). Let \(c_1\) be the least index \(k\) for which \(X_k\) is non–zero. If \(c_1 = m\) or if all the vectors \(X_k\) with \(k > c_1\) are linear combinations of \(X_{c_1}\), terminate the algorithm and let \(r = 1\). Otherwise let \(c_2\) be the least integer \(k > c_1\) such that \(X_k\) is not a linear combination of \(X_{c_1}\).

If \(c_2 = m\) or if all the vectors \(X_k\) with \(k > c_2\) are linear combinations of \(X_{c_1}\) and \(X_{c_2}\), terminate the algorithm and let \(r = 2\). Eventually the algorithm will terminate at the \(r\)–th stage, either because \(c_r = m\), or because all vectors \(X_k\) with \(k > c_r\) are linear combinations of \(X_{c_1}, \ldots, X_{c_r}\).

Then it is clear by the construction of \(X_{c_1}, \ldots, X_{c_r}\), using Corollary 3.2.1 that

(a) \(\langle X_{c_1}, \ldots, X_{c_r} \rangle = \langle X_1, \ldots, X_m \rangle\);  
(b) the vectors \(X_{c_1}, \ldots, X_{c_r}\) are linearly independent by the left–to–right test.

Consequently \(X_{c_1}, \ldots, X_{c_r}\) form a basis (called the left–to–right basis) for the subspace \(\langle X_1, \ldots, X_m \rangle\).

EXAMPLE 3.4.2 Let \(X\) and \(Y\) be linearly independent vectors in \(\mathbb{R}^n\). Then the subspace \(\langle 0, 2X, X, -Y, X + Y \rangle\) has left–to–right basis consisting of \(2X, -Y\).

A subspace \(S\) will in general have more than one basis. For example, any permutation of the vectors in a basis will yield another basis. Given one particular basis, one can determine all bases for \(S\) using a simple formula. This is left as one of the problems at the end of this chapter.

We settle for the following important fact about bases:
Theorem 3.4.2 Any two bases for a subspace $S$ must contain the same number of elements.

Proof. For if $X_1,\ldots,X_r$ and $Y_1,\ldots,Y_s$ are bases for $S$, then $Y_1,\ldots,Y_s$ form a linearly independent family in $S = \langle X_1,\ldots,X_r \rangle$ and hence $s \leq r$ by Theorem 3.3.2. Similarly $r \leq s$ and hence $r = s$.

Definition 3.4.2 This number is called the dimension of $S$ and is written $\dim S$. Naturally we define $\dim \{0\} = 0$.

It follows from Theorem 3.3.1 that for any subspace $S$ of $F^n$, we must have $\dim S \leq n$.

Example 3.4.3 If $E_1,\ldots,E_n$ denote the $n$-dimensional unit vectors in $F^n$, then $\dim \langle E_1,\ldots,E_i \rangle = i$ for $1 \leq i \leq n$.

The following result gives a useful way of exhibiting a basis.

Theorem 3.4.3 A linearly independent family of $m$ vectors in a subspace $S$, with $\dim S = m$, must be a basis for $S$.

Proof. Let $X_1,\ldots,X_m$ be a linearly independent family of vectors in a subspace $S$, where $\dim S = m$. We have to show that every vector $X \in S$ is expressible as a linear combination of $X_1,\ldots,X_m$. We consider the following family of vectors in $S$: $X_1,\ldots,X_m, X$. This family contains $m+1$ elements and is consequently linearly dependent by Theorem 3.3.2. Hence we have

$$x_1X_1 + \cdots + x_mX_m + x_{m+1}X = 0,$$

where not all of $x_1,\ldots,x_{m+1}$ are zero. Now if $x_{m+1} = 0$, we would have

$$x_1X_1 + \cdots + x_mX_m = 0,$$

with not all of $x_1,\ldots,x_m$ zero, contradicting the assumption that $X_1,\ldots,X_m$ are linearly independent. Hence $x_{m+1} \neq 0$ and we can use equation 3.1 to express $X$ as a linear combination of $X_1,\ldots,X_m$:

$$X = \frac{-x_1}{x_{m+1}}X_1 + \cdots + \frac{-x_m}{x_{m+1}}X_m.$$
3.5 Rank and nullity of a matrix

We can now define three important integers associated with a matrix.

**Definition 3.5.1** Let \( A \in M_{m \times n}(F) \). Then

(a) column rank \( A = \dim C(A) \);
(b) row rank \( A = \dim R(A) \);
(c) nullity \( A = \dim N(A) \).

We will now see that the reduced row–echelon form \( B \) of a matrix \( A \) allows us to exhibit bases for the row space, column space and null space of \( A \). Moreover, an examination of the number of elements in each of these bases will immediately result in the following theorem:

**Theorem 3.5.1** Let \( A \in M_{m \times n}(F) \). Then

(a) column rank \( A = \) row rank \( A \);
(b) column rank \( A + \) nullity \( A = n \).

Finding a basis for \( R(A) \): The \( r \) non–zero rows of \( B \) form a basis for \( R(A) \) and hence row rank \( A = r \).

For we have seen earlier that \( R(A) = R(B) \). Also

\[
R(B) = \langle B_{1*}, \ldots, B_{m*} \rangle \\
    = \langle B_{1*}, \ldots, B_{r*}, 0 \ldots, 0 \rangle \\
    = \langle B_{1*}, \ldots, B_{r*} \rangle.
\]

The linear independence of the non–zero rows of \( B \) is proved as follows: Let the leading entries of rows 1, \ldots, \( r \) of \( B \) occur in columns \( c_1, \ldots, c_r \). Suppose that

\[
x_1 B_{1*} + \cdots + x_r B_{r*} = 0.
\]

Then equating components \( c_1, \ldots, c_r \) of both sides of the last equation, gives \( x_1 = 0, \ldots, x_r = 0 \), in view of the fact that \( B \) is in reduced row–echelon form.

Finding a basis for \( C(A) \): The \( r \) columns \( A_{*c_1}, \ldots, A_{*c_r} \) form a basis for \( C(A) \) and hence column rank \( A = r \). For it is clear that columns \( c_1, \ldots, c_r \) of \( B \) form the left–to–right basis for \( C(B) \) and consequently from parts (b) and (c) of Theorem 3.3.5, it follows that columns \( c_1, \ldots, c_r \) of \( A \) form the left–to–right basis for \( C(A) \).
Finding a basis for $N(A)$: For notational simplicity, let us suppose that $c_1 = 1, \ldots, c_r = r$. Then $B$ has the form
\[
B = \begin{bmatrix}
1 & 0 & \cdots & 0 & b_{1r+1} & \cdots & b_{1n} \\
0 & 1 & \cdots & 0 & b_{2r+1} & \cdots & b_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & b_{rr+1} & \cdots & b_{rn} \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\end{bmatrix}.
\]

Then $N(B)$ and hence $N(A)$ are determined by the equations
\[
\begin{align*}
x_1 &= (-b_{1r+1})x_{r+1} + \cdots + (-b_{1n})x_n \\
&\vdots \\
x_r &= (-b_{rr+1})x_{r+1} + \cdots + (-b_{rn})x_n,
\end{align*}
\]
where $x_{r+1}, \ldots, x_n$ are arbitrary elements of $F$. Hence the general vector $X$ in $N(A)$ is given by
\[
\begin{bmatrix}
x_1 \\
x_r \\
x_{r+1} \\
x_n
\end{bmatrix} = x_{r+1} \begin{bmatrix}
-b_{1r+1} \\
\vdots \\
-b_{rr+1} \\
1
\end{bmatrix} + \cdots + x_n \begin{bmatrix}
-b_n \\
\vdots \\
-b_{rn} \\
0
\end{bmatrix} + \begin{bmatrix}
0 \\
\vdots \\
0 \\
1
\end{bmatrix}.
\]

(3.2)

Hence $N(A)$ is spanned by $X_1, \ldots, X_{n-r}$, as $x_{r+1}, \ldots, x_n$ are arbitrary. Also $X_1, \ldots, X_{n-r}$ are linearly independent. For equating the right hand side of equation 3.2 to 0 and then equating components $r+1, \ldots, n$ of both sides of the resulting equation, gives $x_{r+1} = 0, \ldots, x_n = 0$.

Consequently $X_1, \ldots, X_{n-r}$ form a basis for $N(A)$.

Theorem 3.5.1 now follows. For we have
\[
\text{row rank } A = \dim R(A) = r \\
\text{column rank } A = \dim C(A) = r.
\]

Hence
\[
\text{row rank } A = \text{column rank } A.
\]
Also
\[
\text{column rank } A + \text{ nullity } A = r + \dim N(A) = r + (n - r) = n.
\]

**DEFINITION 3.5.2** The common value of column rank \( A \) and row rank \( A \) is called the rank of \( A \) and is denoted by \( \text{rank } A \).

**EXAMPLE 3.5.1** Given that the reduced row–echelon form of
\[
A = \begin{bmatrix}
1 & 1 & 5 & 1 & 4 \\
2 & -1 & 1 & 2 & 2 \\
3 & 0 & 6 & 0 & -3
\end{bmatrix}
\]

equal to
\[
B = \begin{bmatrix}
1 & 0 & 2 & 0 & -1 \\
0 & 1 & 3 & 0 & 2 \\
0 & 0 & 0 & 1 & 3
\end{bmatrix},
\]

find bases for \( R(A), C(A) \) and \( N(A) \).

Solution. \([1, 0, 2, 0, -1], [0, 1, 3, 0, 2] \) and \([0, 0, 1, 3] \) form a basis for \( R(A) \). Also
\[
A_{*1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad A_{*2} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad A_{*4} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}
\]
form a basis for \( C(A) \).

Finally \( N(A) \) is given by
\[
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5
\end{bmatrix} = \begin{bmatrix}
-2x_3 + x_5 \\
-3x_3 - 2x_5 \\
x_3 \\
-3x_5 \\
x_5
\end{bmatrix} = x_3 \begin{bmatrix}
-2 \\
-3 \\
1 \\
0 \\
0
\end{bmatrix} + x_5 \begin{bmatrix}
1 \\
-2 \\
0 \\
0 \\
1
\end{bmatrix} = x_3X_1 + x_5X_2,
\]
where \( x_3 \) and \( x_5 \) are arbitrary. Hence \( X_1 \) and \( X_2 \) form a basis for \( N(A) \).

Here rank \( A = 3 \) and nullity \( A = 2 \).

**EXAMPLE 3.5.2** Let \( A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \). Then \( B = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \) is the reduced row–echelon form of \( A \).
3.6. PROBLEMS

Hence $[1, 2]$ is a basis for $R(A)$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a basis for $C(A)$. Also $N(A)$ is given by the equation $x_1 = -2x_2$, where $x_2$ is arbitrary. Then

$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

and hence $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ is a basis for $N(A)$.

Here rank $A = 1$ and nullity $A = 1$.

**EXAMPLE 3.5.3** Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the reduced row–echelon form of $A$.

Hence $[1, 0]$, $[0, 1]$ form a basis for $R(A)$ while $[1, 3]$, $[2, 4]$ form a basis for $C(A)$. Also $N(A) = \{0\}$.

Here rank $A = 2$ and nullity $A = 0$.

We conclude this introduction to vector spaces with a result of great theoretical importance.

**THEOREM 3.5.2** Every linearly independent family of vectors in a subspace $S$ can be extended to a basis of $S$.

Proof. Suppose $S$ has basis $X_1, \ldots, X_m$ and that $Y_1, \ldots, Y_r$ is a linearly independent family of vectors in $S$. Then

$S = \langle X_1, \ldots, X_m \rangle = \langle Y_1, \ldots, Y_r, X_1, \ldots, X_m \rangle$,

as each of $Y_1, \ldots, Y_r$ is a linear combination of $X_1, \ldots, X_m$.

Then applying the left–to–right algorithm to the second spanning family for $S$ will yield a basis for $S$ which includes $Y_1, \ldots, Y_r$.

3.6 PROBLEMS

1. Which of the following subsets of $\mathbb{R}^2$ are subspaces?

   (a) $[x, y]$ satisfying $x = 2y$;
   (b) $[x, y]$ satisfying $x = 2y$ and $2x = y$;
   (c) $[x, y]$ satisfying $x = 2y + 1$;
   (d) $[x, y]$ satisfying $xy = 0$;
(e) \([x, y]\) satisfying \(x \geq 0\) and \(y \geq 0\).

[Answer: (a) and (b).]

2. If \(X, Y, Z\) are vectors in \(\mathbb{R}^n\), prove that

\[\langle X, Y, Z \rangle = \langle X + Y, X + Z, Y + Z \rangle.\]

3. Determine if \(X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}\), \(X_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}\) and \(X_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \end{bmatrix}\) are linearly independent in \(\mathbb{R}^4\).

4. For which real numbers \(\lambda\) are the following vectors linearly independent in \(\mathbb{R}^3\)?

\[X_1 = \begin{bmatrix} \lambda \\ -1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ \lambda \\ -1 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.\]

5. Find bases for the row, column and null spaces of the following matrix over \(\mathbb{Q}\):

\[A = \begin{bmatrix} 1 & 1 & 2 & 0 & 1 \\ 2 & 2 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 8 & 11 & 19 & 0 & 11 \end{bmatrix}.\]

6. Find bases for the row, column and null spaces of the following matrix over \(\mathbb{Z}_2\):

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

7. Find bases for the row, column and null spaces of the following matrix over \(\mathbb{Z}_5\):

\[A = \begin{bmatrix} 1 & 1 & 2 & 0 & 1 & 3 \\ 2 & 1 & 4 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 3 & 0 \\ 3 & 0 & 2 & 4 & 3 & 2 \end{bmatrix}.\]
8. Find bases for the row, column and null spaces of the matrix \( A \) defined in section 1.6, Problem 17. (Note: In this question, \( F \) is a field of four elements.)

9. If \( X_1, \ldots, X_m \) form a basis for a subspace \( S \), prove that
\[
X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_m
\]
also form a basis for \( S \).

10. Let \( A = \begin{bmatrix} a & b & c \\ 1 & 1 & 1 \end{bmatrix} \). Find conditions on \( a, b, c \) such that (a) rank \( A = 1 \); (b) rank \( A = 2 \).

[Answer: (a) \( a = b = c \); (b) at least two of \( a, b, c \) are distinct.]

11. Let \( S \) be a subspace of \( F^n \) with dim \( S = m \). If \( X_1, \ldots, X_m \) are vectors in \( S \) with the property that \( S = \langle X_1, \ldots, X_m \rangle \), prove that \( X_1, \ldots, X_m \) form a basis for \( S \).

12. Find a basis for the subspace \( S \) of \( \mathbb{R}^3 \) defined by the equation
\[
x + 2y + 3z = 0.
\]
Verify that \( Y_1 = [-1, -1, 1]^t \in S \) and find a basis for \( S \) which includes \( Y_1 \).

13. Let \( X_1, \ldots, X_m \) be vectors in \( F^n \). If \( X_i = X_j \), where \( i < j \), prove that \( X_1, \ldots, X_m \) are linearly dependent.

14. Let \( X_1, \ldots, X_{m+1} \) be vectors in \( F^n \). Prove that
\[
\dim \langle X_1, \ldots, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle
\]
if \( X_{m+1} \) is a linear combination of \( X_1, \ldots, X_m \), but
\[
\dim \langle X_1, \ldots, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle + 1
\]
if \( X_{m+1} \) is not a linear combination of \( X_1, \ldots, X_m \).
Deduce that the system of linear equations \( AX = B \) is consistent, if and only if
\[
\text{rank} [A|B] = \text{rank} A.
\]
15. Let $a_1, \ldots, a_n$ be elements of $F$, not all zero. Prove that the set of vectors $[x_1, \ldots, x_n]^t$ where $x_1, \ldots, x_n$ satisfy

$$a_1x_1 + \cdots + a_nx_n = 0$$

is a subspace of $F^n$ with dimension equal to $n - 1$.

16. Prove Lemma 3.2.1, Theorem 3.2.1, Corollary 3.2.1 and Theorem 3.3.2.

17. Let $R$ and $S$ be subspaces of $F^n$, with $R \subseteq S$. Prove that

$$\dim R \leq \dim S$$

and that equality implies $R = S$. (This is a very useful way of proving equality of subspaces.)

18. Let $R$ and $S$ be subspaces of $F^n$. If $R \cup S$ is a subspace of $F^n$, prove that $R \subseteq S$ or $S \subseteq R$.

19. Let $X_1, \ldots, X_r$ be a basis for a subspace $S$. Prove that all bases for $S$ are given by the family $Y_1, \ldots, Y_r$, where

$$Y_i = \sum_{j=1}^{r} a_{ij}X_j,$$

and where $A = [a_{ij}] \in \text{M}_{r \times r}(F)$ is a non-singular matrix.
DEFINITION 4.0.1 If \( A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \), we define the determinant of \( A \), (also denoted by \( \det A \),) to be the scalar
\[
\det A = a_{11}a_{22} - a_{12}a_{21}.
\]
The notation \( \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \) is also used for the determinant of \( A \).

If \( A \) is a real matrix, there is a geometrical interpretation of \( \det A \). If \( P = (x_1, y_1) \) and \( Q = (x_2, y_2) \) are points in the plane, forming a triangle with the origin \( O = (0, 0) \), then apart from sign, \( \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \) is the area of the triangle \( OPQ \). For, using polar coordinates, let \( x_1 = r_1 \cos \theta_1 \) and \( y_1 = r_1 \sin \theta_1 \), where \( r_1 = OP \) and \( \theta_1 \) is the angle made by the ray \( \overrightarrow{OP} \) with the positive \( x \)-axis. Then triangle \( OPQ \) has area \( \frac{1}{2} OP \cdot OQ \sin \alpha \), where \( \alpha = \angle POQ \). If triangle \( OPQ \) has anti-clockwise orientation, then the ray \( \overrightarrow{OQ} \) makes angle \( \theta_2 = \theta_1 + \alpha \) with the positive \( x \)-axis. (See Figure 4.1.)

Also \( x_2 = r_2 \cos \theta_2 \) and \( y_2 = r_2 \sin \theta_2 \). Hence
\[
\text{Area } OPQ = \frac{1}{2} OP \cdot OQ \sin \alpha
= \frac{1}{2} OP \cdot OQ \sin (\theta_2 - \theta_1)
= \frac{1}{2} OP \cdot OQ (\sin \theta_2 \cos \theta_1 - \cos \theta_2 \sin \theta_1)
= \frac{1}{2} (OQ \sin \theta_2 \cdot OP \cos \theta_1 - OQ \cos \theta_2 \cdot OP \sin \theta_1)
\]
Figure 4.1: Area of triangle $OPQ$.

\[
= \frac{1}{2}(y_2x_1 - x_2y_1)
\]

\[
= \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.
\]

Similarly, if triangle $OPQ$ has clockwise orientation, then its area equals

\[-\frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.
\]

For a general triangle $P_1P_2P_3$, with $P_i = (x_i, y_i), i = 1, 2, 3$, we can take $P_1$ as the origin. Then the above formula gives

\[
\frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} \quad \text{or} \quad -\frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix},
\]

according as vertices $P_1P_2P_3$ are anti-clockwise or clockwise oriented.

We now give a recursive definition of the determinant of an $n \times n$ matrix $A = [a_{ij}], n \geq 3$.

**DEFINITION 4.0.2 (Minor)** Let $M_{ij}(A)$ (or simply $M_{ij}$ if there is no ambiguity) denote the determinant of the $(n-1) \times (n-1)$ submatrix of $A$ formed by deleting the $i$–th row and $j$–th column of $A$. ($M_{ij}(A)$ is called the $(i, j)$ minor of $A$.)

Assume that the determinant function has been defined for matrices of size $(n-1) \times (n-1)$. Then $\det A$ is defined by the so-called *first-row Laplace*
expansion:

$$\det A = a_{11} M_{11}(A) - a_{12} M_{12}(A) + \ldots + (-1)^{1+n} M_{1n}(A)$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} M_{1j}(A).$$

For example, if $A = [a_{ij}]$ is a $3 \times 3$ matrix, the Laplace expansion gives

$$\det A = a_{11} M_{11}(A) - a_{12} M_{12}(A) + a_{13} M_{13}(A)$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. $$

(The recursive definition also works for $2 \times 2$ determinants, if we define the determinant of a $1 \times 1$ matrix $[t]$ to be the scalar $t$:

$$\det A = a_{11} M_{11}(A) - a_{12} M_{12}(A) = a_{11}a_{22} - a_{12}a_{21}.$$  

**EXAMPLE 4.0.1** If $P_1P_2P_3$ is a triangle with $P_i = (x_i, y_i), i = 1, 2, 3,$ then the area of triangle $P_1P_2P_3$ is

$$\frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \quad \text{or} \quad -\frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix},$$

according as the orientation of $P_1P_2P_3$ is anti–clockwise or clockwise.

For from the definition of $3 \times 3$ determinants, we have

$$\frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} x_1 & y_2 & 1 \\ x_2 & y_3 & 1 \\ x_3 & y_1 & 1 \end{vmatrix} - \frac{1}{2} \begin{vmatrix} x_2 & y_1 & 1 \\ x_3 & y_2 & 1 \\ x_1 & y_3 & 1 \end{vmatrix} + \frac{1}{2} \begin{vmatrix} x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ x_1 & y_1 & 1 \end{vmatrix}$$

$$= \frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}.$$

One property of determinants that follows immediately from the definition is the following:

**THEOREM 4.0.1** If a row of a matrix is zero, then the value of the determinant is zero.
(The corresponding result for columns also holds, but here a simple proof by induction is needed.)

One of the simplest determinants to evaluate is that of a lower triangular matrix.

**THEOREM 4.0.2** Let \( A = [a_{ij}] \), where \( a_{ij} = 0 \) if \( i < j \). Then

\[
\text{det} A = a_{11}a_{22}\ldots a_{nn}. \tag{4.1}
\]

An important special case is when \( A \) is a diagonal matrix. If \( A = \text{diag}(a_1, \ldots, a_n) \) then \( \text{det} A = a_1\ldots a_n \). In particular, for a scalar matrix \( tI_n \), we have \( \text{det}(tI_n) = t^n \).

**Proof.** Use induction on the size \( n \) of the matrix.

The result is true for \( n = 2 \). Now let \( n > 2 \) and assume the result true for matrices of size \( n - 1 \). If \( A \) is \( n \times n \), then expanding \( \text{det} A \) along row 1 gives

\[
\text{det} A = a_{11} \begin{vmatrix} a_{22} & 0 & \ldots & 0 \\
   a_{32} & a_{33} & \ldots & 0 \\
   \vdots & & & \\
   a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}
= a_{11}(a_{22}\ldots a_{nn})
\]

by the induction hypothesis.

If \( A \) is upper triangular, equation 4.1 remains true and the proof is again an exercise in induction, with the slight difference that the column version of theorem 4.0.1 is needed.

**REMARK 4.0.1** It can be shown that the expanded form of the determinant of an \( n \times n \) matrix \( A \) consists of \( n! \) signed products \( \pm a_{1i_1}a_{2i_2}\ldots a_{ni_n} \), where \( (i_1, i_2, \ldots, i_n) \) is a permutation of \( (1, 2, \ldots, n) \), the sign being 1 or \(-1\), according as the number of inversions of \( (i_1, i_2, \ldots, i_n) \) is even or odd. An inversion occurs when \( i_r > i_s \) but \( r < s \). (The proof is not easy and is omitted.)

The definition of the determinant of an \( n \times n \) matrix was given in terms of the first–row expansion. The next theorem says that we can expand the determinant along any row or column. (The proof is not easy and is omitted.)
**THEOREM 4.0.3**

\[
\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A)
\]

for \(i = 1, \ldots, n\) (the so-called \(i\)-th row expansion) and

\[
\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A)
\]

for \(j = 1, \ldots, n\) (the so-called \(j\)-th column expansion).

**REMARK 4.0.2** The expression \((-1)^{i+j}\) obeys the chess-board pattern of signs:

\[
\begin{bmatrix}
+ & - & + & \cdots \\
- & + & - & \cdots \\
+ & - & + & \cdots \\
\vdots
\end{bmatrix}
\]

The following theorems can be proved by straightforward inductions on the size of the matrix:

**THEOREM 4.0.4** A matrix and its transpose have equal determinants; that is

\[
\det A^t = \det A.
\]

**THEOREM 4.0.5** If two rows of a matrix are equal, the determinant is zero. Similarly for columns.

**THEOREM 4.0.6** If two rows of a matrix are interchanged, the determinant changes sign.

**EXAMPLE 4.0.2** If \(P_1 = (x_1, y_1)\) and \(P_2 = (x_2, y_2)\) are distinct points, then the line through \(P_1\) and \(P_2\) has equation

\[
\begin{bmatrix}
{x} & {y} & 1 \\
{x_1} & {y_1} & 1 \\
{x_2} & {y_2} & 1
\end{bmatrix} = 0.
\]
For, expanding the determinant along row 1, the equation becomes

\[ ax + by + c = 0, \]

where

\[
\begin{aligned}
a &= \begin{vmatrix} y_1 & 1 \\ y_2 & 1 \end{vmatrix} = y_1 - y_2 \\
b &= -\begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} = x_2 - x_1.
\end{aligned}
\]

This represents a line, as not both \( a \) and \( b \) can be zero. Also this line passes through \( P_i \), \( i = 1, 2 \). For the determinant has its first and \( i \)-th rows equal if \( x = x_i \) and \( y = y_i \) and is consequently zero.

There is a corresponding formula in three-dimensional geometry. If \( P_1, P_2, P_3 \) are non-collinear points in three-dimensional space, with \( P_i = (x_i, y_i, z_i), i = 1, 2, 3 \), then the equation

\[
\begin{vmatrix}
  x & y & z & 1 \\
  x_1 & y_1 & z_1 & 1 \\
  x_2 & y_2 & z_2 & 1 \\
  x_3 & y_3 & z_3 & 1
\end{vmatrix} = 0
\]

represents the plane through \( P_1, P_2, P_3 \). For, expanding the determinant along row 1, the equation becomes \( ax + by + cz + d = 0 \), where

\[
\begin{aligned}
a &= \begin{vmatrix} y_1 & z_1 & 1 \\ y_2 & z_2 & 1 \\ y_3 & z_3 & 1 \end{vmatrix}, \\
b &= -\begin{vmatrix} x_1 & z_1 & 1 \\ x_2 & z_2 & 1 \\ x_3 & z_3 & 1 \end{vmatrix}, \\
c &= \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.
\end{aligned}
\]

As we shall see in chapter 6, this represents a plane if at least one of \( a, b, c \) is non-zero. However, apart from sign and a factor \( \frac{1}{2} \), the determinant expressions for \( a, b, c \) give the values of the areas of projections of triangle \( P_1P_2P_3 \) on the \((y, z), (x, z)\) and \((x, y)\) planes, respectively. Geometrically, it is then clear that at least one of \( a, b, c \) is non-zero. It is also possible to give an algebraic proof of this fact.

Finally, the plane passes through \( P_i \), \( i = 1, 2, 3 \) as the determinant has its first and \( i \)-th rows equal if \( x = x_i \), \( y = y_i \), \( z = z_i \) and is consequently zero. We now work towards proving that a matrix is non-singular if its determinant is non-zero.

**DEFINITION 4.0.3 (Cofactor)** The \((i, j)\) cofactor of \( A \), denoted by \( C_{ij}(A) \) (or \( C_{ij} \) if there is no ambiguity) is defined by

\[ C_{ij}(A) = (-1)^{i+j}M_{ij}(A). \]
REMARK 4.0.3 It is important to notice that $C_{ij}(A)$, like $M_{ij}(A)$, does not depend on $a_{ij}$. Use will be made of this observation presently.

In terms of the cofactor notation, Theorem 3.0.2 takes the form

**THEOREM 4.0.7**

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}(A)$$

for $i = 1, \ldots, n$ and

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij}(A)$$

for $j = 1, \ldots, n$.

Another result involving cofactors is

**THEOREM 4.0.8** Let $A$ be an $n \times n$ matrix. Then

(a) $\sum_{j=1}^{n} a_{ij} C_{kj}(A) = 0$ if $i \neq k$.

Also

(b) $\sum_{i=1}^{n} a_{ij} C_{ik}(A) = 0$ if $j \neq k$.

**Proof.**

If $A$ is $n \times n$ and $i \neq k$, let $B$ be the matrix obtained from $A$ by replacing row $k$ by row $i$. Then $\det B = 0$ as $B$ has two identical rows.

Now expand $\det B$ along row $k$. We get

$$0 = \det B = \sum_{j=1}^{n} b_{kj} C_{kj}(B) = \sum_{j=1}^{n} a_{ij} C_{kj}(A),$$

in view of Remark 4.0.3.
DEFINITION 4.0.4 (Adjoint) If $A = [a_{ij}]$ is an $n \times n$ matrix, the adjoint of $A$, denoted by $\text{adj} \, A$, is the transpose of the matrix of cofactors. Hence

$$\text{adj} \, A = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$ 

Theorems 4.0.7 and 4.0.8 may be combined to give

**THEOREM 4.0.9** Let $A$ be an $n \times n$ matrix. Then

$$A(\text{adj} \, A) = (\det A) I_n = (\text{adj} \, A) A.$$ 

**Proof.**

$$(A \text{adj} \, A)_{ik} = \sum_{j=1}^{n} a_{ij} (\text{adj} \, A)_{jk}$$

$$= \sum_{j=1}^{n} a_{ij} C_{kj}(A)$$

$$= \delta_{ik} \det A$$

$$= (\det A) I_n)_{ik}.$$ 

Hence $A(\text{adj} \, A) = (\det A) I_n$. The other equation is proved similarly.

**COROLLARY 4.0.1 (Formula for the inverse)** If $\det A \neq 0$, then $A$ is non–singular and

$$A^{-1} = \frac{1}{\det A} \text{adj} \, A.$$ 

**EXAMPLE 4.0.3** The matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 8 & 9 \end{bmatrix}$$

is non–singular. For

$$\det A = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 8 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 8 & 8 \end{vmatrix}$$

$$= -3 + 24 - 24$$

$$= -3 \neq 0.$$
Also

\[ A^{-1} = \frac{1}{3} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} \begin{bmatrix} 5 & 6 & -2 & 3 \\ 8 & 9 & 8 & 9 \\ -2 & 3 & 2 & 3 \end{bmatrix} \]

\[ = \frac{1}{3} \begin{bmatrix} 5 & 6 & -2 & 3 \\ 8 & 9 & 8 & 9 \\ -2 & 3 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 & 5 & -1 & 2 \\ 8 & 8 & 8 & 4 \end{bmatrix} \]

\[ = \frac{1}{3} \begin{bmatrix} -3 & 6 & -3 \\ 12 & -15 & 6 \\ -8 & 8 & -3 \end{bmatrix}. \]

The following theorem is useful for simplifying and numerically evaluating a determinant. Proofs are obtained by expanding along the corresponding row or column.

**THEOREM 4.0.10** The determinant is a linear function of each row and column.

For example

\[
\begin{vmatrix}
  a_{11} + a'_{11} & a_{12} + a'_{12} & a_{13} + a'_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
= \begin{vmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
+ \begin{vmatrix}
  a'_{11} & a'_{12} & a'_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
\]

\[
\begin{vmatrix}
  t a_{11} & t a_{12} & t a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
= t \begin{vmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}.
\]

**COROLLARY 4.0.2** If a multiple of a row is added to another row, the value of the determinant is unchanged. Similarly for columns.

**Proof.** We illustrate with a 3 × 3 example, but the proof is really quite general.

\[
\begin{vmatrix}
  a_{11} + t a_{21} & a_{12} + t a_{22} & a_{13} + t a_{23} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
= \begin{vmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
+ \begin{vmatrix}
  t a_{21} & t a_{22} & t a_{23} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{vmatrix}
\]
To evaluate a determinant numerically, it is advisable to reduce the matrix to row–echelon form, recording any sign changes caused by row interchanges, together with any factors taken out of a row, as in the following examples.

**EXAMPLE 4.0.4** Evaluate the determinant

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
8 & 8 & 9
\end{vmatrix}
\]

**Solution.** Using row operations \( R_2 \rightarrow R_2 - 4R_1 \) and \( R_3 \rightarrow R_3 - 8R_1 \) and then expanding along the first column, gives

\[
\begin{vmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
8 & 8 & 9
\end{vmatrix}
= \begin{vmatrix}
1 & 2 & 3 \\
0 & -3 & -6 \\
0 & -8 & -15
\end{vmatrix}
= \begin{vmatrix}
-3 & -6 \\
-8 & -15
\end{vmatrix}
= -3 \begin{vmatrix}
1 & 2 \\
-8 & -15
\end{vmatrix}
= -3 \begin{vmatrix}
1 & 2 \\
0 & 1
\end{vmatrix}
= -3.
\]

**EXAMPLE 4.0.5** Evaluate the determinant

\[
\begin{vmatrix}
1 & 1 & 2 & 1 \\
3 & 1 & 4 & 5 \\
7 & 6 & 1 & 2 \\
1 & 1 & 3 & 4
\end{vmatrix}
\]

**Solution.**

\[
\begin{vmatrix}
1 & 1 & 2 & 1 \\
3 & 1 & 4 & 5 \\
7 & 6 & 1 & 2 \\
1 & 1 & 3 & 4
\end{vmatrix}
= \begin{vmatrix}
1 & 1 & 2 & 1 \\
0 & -2 & -2 & 2 \\
0 & -1 & -13 & -5 \\
0 & 0 & 1 & 3
\end{vmatrix}
\]
EXAMPLE 4.0.6 (Vandermonde determinant) Prove that

\[
\begin{vmatrix}
1 & 1 & 1 \\
a & b & c \\
a^2 & b^2 & c^2
\end{vmatrix}
= (b - a)(c - a)(c - b).
\]

Solution. Subtracting column 1 from columns 2 and 3, then expanding along row 1, gives

\[
\begin{vmatrix}
1 & 1 & 1 \\
a & b & c \\
a^2 & b^2 & c^2
\end{vmatrix}
= \begin{vmatrix}
1 & 0 & 0 \\
a & b - a & c - a \\
a^2 & b^2 - a^2 & c^2 - a^2
\end{vmatrix}
= \begin{vmatrix}
b - a & c - a \\
b^2 - a^2 & c^2 - a^2
\end{vmatrix}
= (b - a)(c - a) \begin{vmatrix} 1 & 1 \\ b + a & c + a \end{vmatrix}
= (b - a)(c - a)(c - b).
\]

REMARK 4.0.4 From theorems 4.0.6, 4.0.10 and corollary 4.0.2, we deduce

(a) \( \det (E_{ij}A) = -\det A, \)

(b) \( \det (E_i(t)A) = t \det A, \) if \( t \neq 0, \)
(c) \( \det (E_{ij}(t)A) = \det A \).

It follows that if \( A \) is row–equivalent to \( B \), then \( \det B = c \det A \), where \( c \neq 0 \).

Hence \( \det B \neq 0 \Leftrightarrow \det A \neq 0 \) and \( \det B = 0 \Leftrightarrow \det A = 0 \). Consequently from theorem 2.5.8 and remark 2.5.7, we have the following important result:

**THEOREM 4.0.11** Let \( A \) be an \( n \times n \) matrix. Then

(i) \( A \) is non–singular if and only if \( \det A \neq 0 \);

(ii) \( A \) is singular if and only if \( \det A = 0 \);

(iii) the homogeneous system \( AX = 0 \) has a non–trivial solution if and only if \( \det A = 0 \).

**EXAMPLE 4.0.7** Find the rational numbers \( a \) for which the following homogeneous system has a non–trivial solution and solve the system for these values of \( a \):

\[
\begin{align*}
    x - 2y + 3z &= 0 \\
    ax + 3y + 2z &= 0 \\
    6x + y + az &= 0.
\end{align*}
\]

**Solution.** The coefficient determinant of the system is

\[
\Delta = \begin{vmatrix} 1 & -2 & 3 \\ a & 3 & 2 \\ 6 & 1 & a \end{vmatrix} = \begin{vmatrix} 1 & -2 & 3 \\ 0 & 3+2a & 2-3a \\ 0 & 13 & a-18 \end{vmatrix} = \begin{vmatrix} 3+2a & 2-3a \\ 13 & a-18 \end{vmatrix} = (3+2a)(a-18) - 13(2-3a) = 2a^2 + 6a - 80 = 2(a + 8)(a - 5).
\]

So \( \Delta = 0 \Leftrightarrow a = -8 \) or \( a = 5 \) and these values of \( a \) are the only values for which the given homogeneous system has a non–trivial solution.

If \( a = -8 \), the coefficient matrix has reduced row–echelon form equal to

\[
\begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & -2 \\
0 & 0 & 0
\end{bmatrix}
\]
and so the complete solution is $x = z$, $y = 2z$, with $z$ arbitrary. If $a = 5$, the coefficient matrix has reduced row–echelon form equal to
\[
\begin{bmatrix}
  1 & 0 & 1 \\
  0 & 1 & -1 \\
  0 & 0 & 0 \\
\end{bmatrix}
\]
and so the complete solution is $x = -z$, $y = z$, with $z$ arbitrary.

**EXAMPLE 4.0.8** Find the values of $t$ for which the following system is consistent and solve the system in each case:
\[
\begin{align*}
x + y &= 1 \\
tx + y &= t \\
(1 + t)x + 2y &= 3.
\end{align*}
\]

**Solution.** Suppose that the given system has a solution $(x_0, y_0)$. Then the following homogeneous system
\[
\begin{align*}
x + y + z &= 0 \\
tx + y + tz &= 0 \\
(1 + t)x + 2y + 3z &= 0
\end{align*}
\]
will have a non–trivial solution
\[
x = x_0, \quad y = y_0, \quad z = -1.
\]

Hence the coefficient determinant $\Delta$ is zero. However
\[
\Delta = \begin{vmatrix} 1 & 1 & 1 \\ t & 1 & t \\ 1 + t & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ t & 1 - t & 0 \\ 1 + t & 1 - t & 2 - t \end{vmatrix} = \begin{vmatrix} 1 - t & 0 \\ 1 - t & 2 - t \end{vmatrix} = (1-t)(2-t).
\]

Hence $t = 1$ or $t = 2$. If $t = 1$, the given system becomes
\[
\begin{align*}
x + y &= 1 \\
x + y &= 1 \\
2x + 2y &= 3
\end{align*}
\]
which is clearly inconsistent. If $t = 2$, the given system becomes
\[
\begin{align*}
x + y &= 1 \\
2x + y &= 2 \\
3x + 2y &= 3
\end{align*}
\]
which has the unique solution $x = 1$, $y = 0$.

To finish this section, we present an old (1750) method of solving a system of $n$ equations in $n$ unknowns called Cramer’s rule. The method is not used in practice. However it has a theoretical use as it reveals explicitly how the solution depends on the coefficients of the augmented matrix.

**THEOREM 4.0.12 (Cramer’s rule)** The system of $n$ linear equations in $n$ unknowns $x_1, \ldots, x_n$

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
    \vdots \\
    a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{align*}
\]

has a unique solution if $\Delta = \det[a_{ij}] \neq 0$, namely

\[
    x_1 = \frac{\Delta_1}{\Delta},
    x_2 = \frac{\Delta_2}{\Delta}, \ldots,
    x_n = \frac{\Delta_n}{\Delta},
\]

where $\Delta_i$ is the determinant of the matrix formed by replacing the $i$–th column of the coefficient matrix $A$ by the entries $b_1, b_2, \ldots, b_n$.

**Proof.** Suppose the coefficient determinant $\Delta \neq 0$. Then by corollary 4.0.1, $A^{-1}$ exists and is given by $A^{-1} = \frac{1}{\Delta} \text{adj } A$ and the system has the unique solution

\[
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
= A^{-1}
\begin{bmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_n
\end{bmatrix}
= \frac{1}{\Delta}
\begin{bmatrix}
    C_{11} & C_{21} & \cdots & C_{n1} \\
    C_{12} & C_{22} & \cdots & C_{n2} \\
    \vdots & \vdots & & \vdots \\
    C_{1n} & C_{2n} & \cdots & C_{nn}
\end{bmatrix}
\begin{bmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_n
\end{bmatrix}
= \frac{1}{\Delta}
\begin{bmatrix}
    b_1C_{11} + b_2C_{21} + \cdots + b_nC_{n1} \\
    b_2C_{12} + b_2C_{22} + \cdots + b_nC_{n2} \\
    \vdots \\
    b_nC_{1n} + b_2C_{2n} + \cdots + b_nC_{nn}
\end{bmatrix}.
\]

However the $i$–th component of the last vector is the expansion of $\Delta_i$ along column $i$. Hence

\[
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix}
= \frac{1}{\Delta}
\begin{bmatrix}
    \Delta_1 \\
    \Delta_2 \\
    \vdots \\
    \Delta_n
\end{bmatrix}
= \begin{bmatrix}
    \Delta_1/\Delta \\
    \Delta_2/\Delta \\
    \vdots \\
    \Delta_n/\Delta
\end{bmatrix}.
\]
4.1 PROBLEMS

1. If the points \( P_i = (x_i, y_i), i = 1, 2, 3, 4 \) form a quadrilateral with vertices in anti-clockwise orientation, prove that the area of the quadrilateral equals
\[
\frac{1}{2} \left( \begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 & x_4 & x_1 \\ y_2 & y_3 & y_4 & y_1 \end{vmatrix} + \begin{vmatrix} x_3 & x_4 & x_1 & x_2 \\ y_3 & y_4 & y_1 & y_2 \end{vmatrix} \right).
\]
(This formula generalizes to a simple polygon and is known as the Surveyor’s formula.)

2. Prove that the following identity holds by expressing the left-hand side as the sum of 8 determinants:
\[
\begin{vmatrix} a + x & b + y & c + z \\ x + u & y + v & z + w \\ u + a & v + b & w + c \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ x & y & z \\ u & v & w \end{vmatrix}.
\]

3. Prove that
\[
\begin{vmatrix} n^2 & (n + 1)^2 & (n + 2)^2 \\ (n + 1)^2 & (n + 2)^2 & (n + 3)^2 \\ (n + 2)^2 & (n + 3)^2 & (n + 4)^2 \end{vmatrix} = -8.
\]

4. Evaluate the following determinants:

(a) \[
\begin{vmatrix} 246 & 427 & 327 \\ 1014 & 543 & 443 \\ -342 & 721 & 621 \end{vmatrix}
\]

(b) \[
\begin{vmatrix} 1 & 2 & 3 & 4 \\ -2 & 1 & -4 & 3 \\ 3 & -4 & -1 & 2 \\ 4 & 3 & -2 & -1 \end{vmatrix}
\]

[Answers: (a) -29400000; (b) 900.]

5. Compute the inverse of the matrix
\[
A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 4 \\ 5 & 2 & -3 \end{bmatrix}
\]
by first computing the adjoint matrix.

[Answer: \( A^{-1} = \frac{1}{13} \begin{bmatrix} -4 & 2 \\ 29 & -10 \\ 1 & -2 \end{bmatrix} \).]
6. Prove that the following identities hold:

\[
\begin{vmatrix}
2a & 2b & b - c \\
2b & 2a & a + c \\
\ a + b & \ a + b & \ b
\end{vmatrix} = -2(a - b)^2(a + b),
\]

\[
\begin{vmatrix}
b + c & b & c \\
c & c + a & a \\
\ b & \ a & \ a + b
\end{vmatrix} = 2a(b^2 + c^2).
\]

7. Let \( P_i = (x_i, y_i), i = 1, 2, 3. \) If \( x_1, x_2, x_3 \) are distinct, prove that there is precisely one curve of the form \( y = ax^2 + bx + c \) passing through \( P_1, P_2 \) and \( P_3 \).

8. Let

\[
A = \begin{bmatrix}
1 & 1 & -1 \\
2 & 3 & k \\
1 & k & 3
\end{bmatrix}.
\]

Find the values of \( k \) for which \( \det A = 0 \) and hence, or otherwise, determine the value of \( k \) for which the following system has more than one solution:

\[
\begin{align*}
x + y - z &= 1 \\
2x + 3y + kz &= 3 \\
x + ky + 3z &= 2.
\end{align*}
\]

Solve the system for this value of \( k \) and determine the solution for which \( x^2 + y^2 + z^2 \) has least value.

[Answer: \( k = 2; \ x = 10/21, \ y = 13/21, \ z = 2/21 \).]

9. By considering the coefficient determinant, find all rational numbers \( a \) and \( b \) for which the following system has (i) no solutions, (ii) exactly one solution, (iii) infinitely many solutions:

\[
\begin{align*}
x - 2y + bz &= 3 \\
ax + 2z &= 2 \\
5x + 2y &= 1.
\end{align*}
\]

Solve the system in case (iii).

[Answer: (i) \( ab = 12 \) and \( a \neq 3 \), no solution; \( ab \neq 12 \), unique solution; \( a = 3, \ b = 4 \), infinitely many solutions; \( x = -\frac{2}{3}z + \frac{2}{5}, \ y = \frac{5}{3}z - \frac{7}{5}, \) with \( z \) arbitrary.]
10. Express the determinant of the matrix
\[
B = \begin{bmatrix}
1 & 1 & 2 & 1 \\
1 & 2 & 3 & 4 \\
2 & 4 & 7 & 2t + 6 \\
2 & 2 & 6 - t & t
\end{bmatrix}
\]
as as polynomial in \( t \) and hence determine the rational values of \( t \) for which \( B^{-1} \) exists.

[Answer: \( \det B = (t - 2)(2t - 1); t \neq 2 \) and \( t \neq \frac{1}{2} \).]

11. If \( A \) is a \( 3 \times 3 \) matrix over a field and \( \det A \neq 0 \), prove that
   (i) \( \det (\text{adj} \ A) = (\det A)^2 \),
   (ii) \( (\text{adj} \ A)^{-1} = \frac{1}{\det A} A = \text{adj} (A^{-1}) \).

12. Suppose that \( A \) is a real \( 3 \times 3 \) matrix such that \( A^t A = I_3 \).
   (i) Prove that \( A^t (A - I_3) = -(A - I_3)^t \).
   (ii) Prove that \( \det A = \pm 1 \).
   (iii) Use (i) to prove that if \( \det A = 1 \), then \( (A - I_3) = 0 \).

13. If \( A \) is a square matrix such that one column is a linear combination of the remaining columns, prove that \( \det A = 0 \). Prove that the converse also holds.

14. Use Cramer’s rule to solve the system
\[
\begin{align*}
-2x + 3y - z &= 1 \\
x + 2y - z &= 4 \\
-2x - y + z &= -3.
\end{align*}
\]
[Answer: \( x = 2, \ y = 3, \ z = 4 \).]

15. Use remark 4.0.4 to deduce that
\[
\det E_{ij} = -1, \quad \det E_i(t) = t, \quad \det E_{ij}(t) = 1
\]
and use theorem 2.5.8 and induction, to prove that
\[
\det (BA) = \det B \det A,
\]
if \( B \) is non–singular. Also prove that the formula holds when \( B \) is singular.
16. Prove that
\[
\begin{vmatrix}
a + b + c & a + b & a & a \\
 a + b & a + b + c & a & a \\
 a & a & a + b + c & a + b \\
 a & a & a + b & a + b + c
\end{vmatrix} = c^2(2b+c)(4a+2b+c).
\]

17. Prove that
\[
\begin{vmatrix}
1 + u_1 & u_1 & u_1 & u_1 \\
 u_2 & 1 + u_2 & u_2 & u_2 \\
 u_3 & u_3 & 1 + u_3 & u_3 \\
 u_4 & u_4 & u_4 & 1 + u_4
\end{vmatrix} = 1 + u_1 + u_2 + u_3 + u_4.
\]

18. Let \( A \in M_{n \times n}(F) \). If \( A^t = -A \), prove that \( \det A = 0 \) if \( n \) is odd and \( 1 + 1 \neq 0 \) in \( F \).

19. Prove that
\[
\begin{vmatrix}
1 & 1 & 1 & 1 \\
 r & 1 & 1 & 1 \\
 r & r & 1 & 1 \\
 r & r & r & 1
\end{vmatrix} = (1 - r)^3.
\]

20. Express the determinant
\[
\begin{vmatrix}
1 & a^2 - bc & a^4 \\
1 & b^2 - ca & b^4 \\
1 & c^2 - ab & c^4
\end{vmatrix}
\]
as the product of one quadratic and four linear factors.

[Answer: \((b-a)(c-a)(c-b)(a+b+c)(b^2 + bc + c^2 + ac + ab + a^2)\).]
Chapter 5

COMPLEX NUMBERS

5.1 Constructing the complex numbers

One way of introducing the field \( \mathbb{C} \) of complex numbers is via the arithmetic of \( 2 \times 2 \) matrices.

**DEFINITION 5.1.1** A complex number is a matrix of the form

\[
\begin{bmatrix}
x & -y \\
y & x
\end{bmatrix},
\]

where \( x \) and \( y \) are real numbers.

Complex numbers of the form \( \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \) are scalar matrices and are called *real* complex numbers and are denoted by the symbol \( \{ x \} \).

The real complex numbers \( \{ x \} \) and \( \{ y \} \) are respectively called the *real* part and *imaginary part* of the complex number \( \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \).

The complex number \( \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \) is denoted by the symbol \( i \).

We have the identities

\[
\begin{bmatrix} x & -y \\ y & x \end{bmatrix} = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} + \begin{bmatrix} 0 & -y \\ y & 0 \end{bmatrix} = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y & 0 \\ 0 & y \end{bmatrix}
\]

\[= \{ x \} + i \{ y \}, \]

\[i^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \{-1\}.\]
CHAPTER 5. COMPLEX NUMBERS

Complex numbers of the form $i\{y\}$, where $y$ is a non–zero real number, are called imaginary numbers.

If two complex numbers are equal, we can equate their real and imaginary parts:

$$\{x_1\} + i\{y_1\} = \{x_2\} + i\{y_2\} \Rightarrow x_1 = x_2 \text{ and } y_1 = y_2,$$

if $x_1, x_2, y_1, y_2$ are real numbers. Noting that $\{0\} + i\{0\} = \{0\}$, gives the useful special case is

$$\{x\} + i\{y\} = \{0\} \Rightarrow x = 0 \text{ and } y = 0,$$

if $x$ and $y$ are real numbers.

The sum and product of two real complex numbers are also real complex numbers:

$$\{x\} + \{y\} = \{x + y\}, \quad \{x\}\{y\} = \{xy\}.$$ 

Also, as real complex numbers are scalar matrices, their arithmetic is very simple. They form a field under the operations of matrix addition and multiplication. The additive identity is $\{0\}$, the additive inverse of $\{x\}$ is $\{-x\}$, the multiplicative identity is $\{1\}$ and the multiplicative inverse of $\{x\}$ is $\{x^{-1}\}$. Consequently

$$\{x\} - \{y\} = \{x\} + (-\{y\}) = \{x\} + \{-y\} = \{x - y\},$$

$$\frac{\{x\}}{\{y\}} = \{x\}\{y^{-1}\} = \{x\}\{y^{-1}\} = \{xy^{-1}\} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

It is customary to blur the distinction between the real complex number $\{x\}$ and the real number $x$ and write $\{x\}$ as $x$. Thus we write the complex number $\{x\} + i\{y\}$ simply as $x + iy$.

More generally, the sum of two complex numbers is a complex number:

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2); \quad (5.1)$$

and (using the fact that scalar matrices commute with all matrices under matrix multiplication and $\{-1\}A = -A$ if $A$ is a matrix), the product of two complex numbers is a complex number:

$$(x_1 + iy_1)(x_2 + iy_2) = x_1(x_2 + iy_2) + (iy_1)(x_2 + iy_2)$$

$$= x_1x_2 + x_1(iy_2) + (iy_1)x_2 + (iy_1)(iy_2)$$

$$= x_1x_2 + ix_1y_2 + ixy_1x_2 + i^2y_1y_2$$

$$= (x_1x_2 + \{-1\}y_1y_2) + i(x_1y_2 + y_1x_2)$$

$$= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2), \quad (5.2)$$
5.2 Calculating with complex numbers

The set \( \mathbb{C} \) of complex numbers forms a field under the operations of matrix addition and multiplication. The additive identity is 0, the additive inverse of \( x + iy \) is the complex number \( (-x) + i(-y) \), the multiplicative identity is 1 and the multiplicative inverse of the non–zero complex number \( x + iy \) is the complex number \( u + iv \), where

\[
u = \frac{-y}{x^2 + y^2} \quad \text{and} \quad v = \frac{x}{x^2 + y^2}.
\]

(If \( x + iy \neq 0 \), then \( x \neq 0 \) or \( y \neq 0 \), so \( x^2 + y^2 \neq 0 \).)

From equations 5.1 and 5.2, we observe that addition and multiplication of complex numbers is performed just as for real numbers, replacing \( i^2 \) by \(-1\), whenever it occurs.

A useful identity satisfied by complex numbers is

\[
r^2 + s^2 = (r + is)(r - is).
\]

This leads to a method of expressing the ratio of two complex numbers in the form \( x + iy \), where \( x \) and \( y \) are real complex numbers.

\[
\frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} = \frac{(x_1x_2 + y_1y_2) + i(-x_1y_2 + y_1x_2)}{x_2^2 + y_2^2}.
\]

The process is known as rationalization of the denominator.

5.2 Calculating with complex numbers

We can now do all the standard linear algebra calculations over the field of complex numbers – find the reduced row–echelon form of an matrix whose elements are complex numbers, solve systems of linear equations, find inverses and calculate determinants.

For example,

\[
\begin{vmatrix}
1 + i & 2 - i \\
7 & 8 - 2i
\end{vmatrix} = (1 + i)(8 - 2i) - 7(2 - i)
\]

\[
= (8 - 2i) + i(8 - 2i) - 14 + 7i
\]

\[
= -4 + 13i \neq 0.
\]
Then by Cramer’s rule, the linear system

\[
\begin{align*}
(1 + i)z + (2 - i)w &= 2 + 7i \\
7z + (8 - 2i)w &= 4 - 9i
\end{align*}
\]

has the unique solution

\[
z = \frac{\begin{vmatrix}
2 + 7i & 2 - i \\
4 - 9i & 8 - 2i
\end{vmatrix}}{-4 + 13i}
\]

\[
= \frac{(2 + 7i)(8 - 2i) - (4 - 9i)(2 - i)}{-4 + 13i}
\]

\[
= \frac{2(8 - 2i) + (7i)(8 - 2i) - \{(4(2 - i) - 9i(2 - i)\}}{-4 + 13i}
\]

\[
= \frac{16 - 4i + 56i - 14i^2 - \{8 - 4i - 18i + 9i^2\}}{-4 + 13i}
\]

\[
= \frac{31 + 74i}{-4 + 13i}
\]

\[
= \frac{(31 + 74i)(-4 + 13i)}{(-4)^2 + 13^2}
\]

\[
= \frac{838 - 699i}{185}
\]

\[
= \frac{838}{185} - \frac{699}{185}i
\]

and similarly \(w = \frac{698}{185} + \frac{229}{185}i\).

An important property enjoyed by complex numbers is that every complex number has a square root:

**Theorem 5.2.1**

If \(w\) is a non-zero complex number, then the equation \(z^2 = w\) has precisely two solutions \(z \in \mathbb{C}\).

**Proof.** Case 1. Suppose \(b = 0\). Then if \(a > 0\), \(z = \sqrt{a}\) is a solution, while if \(a < 0\), \(i\sqrt{-a}\) is a solution.

Case 2. Suppose \(b \neq 0\). Let \(z = x + iy\), \(w = a + ib\), \(x, y, a, b \in \mathbb{R}\). Then the equation \(z^2 = w\) becomes

\[
(x + iy)^2 = x^2 - y^2 + 2xyi = a + ib,
\]
so equating real and imaginary parts gives

\( x^2 - y^2 = a \) and \( 2xy = b \).

Hence \( x \neq 0 \) and \( y = b/(2x) \). Consequently

\[
 x^2 - \left( \frac{b}{2x} \right)^2 = a,
\]

so \( 4x^4 - 4ax^2 - b^2 = 0 \) and \( 4(x^2)^2 - 4a(x^2) - b^2 = 0 \). Hence

\[
 x^2 = \frac{4a \pm \sqrt{16a^2 + 16b^2}}{8} = \frac{a \pm \sqrt{a^2 + b^2}}{2}.
\]

However \( x^2 > 0 \), so we must take the + sign, as \( a - \sqrt{a^2 + b^2} < 0 \). Hence

\[
 x^2 = \frac{a + \sqrt{a^2 + b^2}}{2}, \quad x = \pm \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}.
\]

Then \( y \) is determined by \( y = b/(2x) \).

**Example 5.2.1** Solve the equation \( z^2 = 1 + i \).

**Solution.** Put \( z = x + iy \). Then the equation becomes

\[
 (x + iy)^2 = x^2 - y^2 + 2xyi = 1 + i,
\]

so equating real and imaginary parts gives

\[
 x^2 - y^2 = 1 \) and \( 2xy = 1 \).
\]

Hence \( x \neq 0 \) and \( y = b/(2x) \). Consequently

\[
 x^2 - \left( \frac{1}{2x} \right)^2 = 1,
\]

so \( 4x^4 - 4x^2 - 1 = 0 \). Hence

\[
 x^2 = \frac{4 \pm \sqrt{16 + 16}}{8} = \frac{1 \pm \sqrt{2}}{2}.
\]

Hence

\[
 x^2 = \frac{1 + \sqrt{2}}{2} \) and \( x = \pm \sqrt{\frac{1 + \sqrt{2}}{2}}.
\]
Then
\[ y = \frac{1}{2x} = \pm \frac{1}{\sqrt{2} \sqrt{1 + \sqrt{2}}}. \]

Hence the solutions are
\[ z = \pm \left( \sqrt{\frac{1 + \sqrt{2}}{2}} \pm \frac{i}{\sqrt{2} \sqrt{1 + \sqrt{2}}} \right). \]

**EXAMPLE 5.2.2** Solve the equation \( z^2 + (\sqrt{3} + i)z + 1 = 0. \)

**Solution.** Because every complex number has a square root, the familiar formula
\[ z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \]
for the solution of the general quadratic equation \( az^2 + bz + c = 0 \) can be used, where now \( a(\neq 0), b, c \in \mathbb{C} \). Hence
\[
\begin{align*}
z &= \frac{-\sqrt{3} + i \pm \sqrt{(\sqrt{3} + i)^2 - 4}}{2} \\
&= \frac{-\sqrt{3} + i \pm \sqrt{3 + 2\sqrt{3}i - 1} - 4}{2} \\
&= \frac{-\sqrt{3} + i \pm \sqrt{-2 + 2\sqrt{3}i}}{2}.
\end{align*}
\]

Now we have to solve \( w^2 = -2 + 2\sqrt{3}i \). Put \( w = x + iy \). Then \( w^2 = x^2 - y^2 + 2xyi = -2 + 2\sqrt{3}i \) and equating real and imaginary parts gives \( x^2 - y^2 = -2 \) and \( 2xy = 2\sqrt{3} \). Hence \( y = \sqrt{3}/x \) and so \( x^2 - 3/x^2 = -2 \). So \( x^4 + 2x^2 - 3 = 0 \) and \( (x^2 + 3)(x^2 - 1) = 0 \). Hence \( x^2 - 1 = 0 \) and \( x = \pm 1 \). Then \( y = \pm\sqrt{3} \). Hence \( (1 + \sqrt{3}i)^2 = -2 + 2\sqrt{3}i \) and the formula for \( z \) now becomes
\[
\begin{align*}
z &= \frac{-\sqrt{3} - i \pm (1 + \sqrt{3}i)}{2} \\
&= \frac{1 - \sqrt{3} + (1 + \sqrt{3}i)i}{2} \text{ or } \frac{-1 - \sqrt{3} - (1 + \sqrt{3}i)}{2}.
\end{align*}
\]

**EXAMPLE 5.2.3** Find the cube roots of 1.
5.3. **GEOMETRIC REPRESENTATION OF \( \mathbb{C} \)**

**Solution.** We have to solve the equation \( z^3 = 1 \), or \( z^3 - 1 = 0 \). Now \( z^3 - 1 = (z - 1)(z^2 + z + 1) \). So \( z^3 - 1 = 0 \Rightarrow z - 1 = 0 \) or \( z^2 + z + 1 = 0 \). But

\[
z^2 + z + 1 = 0 \Rightarrow z = \frac{-1 \pm \sqrt{1^2 - 4}}{2} = \frac{-1 \pm \sqrt{3}i}{2}.
\]

So there are 3 cube roots of 1, namely \( i \) and \((−1 ± \sqrt{3}i)/2\).

We state the next theorem without proof. It states that every non–constant polynomial with complex number coefficients has a root in the field of complex numbers.

**THEOREM 5.2.2 (Gauss)** If \( f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \), where \( a_n \neq 0 \) and \( n \geq 1 \), then \( f(z) = 0 \) for some \( z \in \mathbb{C} \).

It follows that in view of the factor theorem, which states that if \( a \in F \) is a root of a polynomial \( f(z) \) with coefficients from a field \( F \), then \( z - a \) is a factor of \( f(z) \), that is \( f(z) = (z - a)g(z) \), where the coefficients of \( g(z) \) also belong to \( F \). By repeated application of this result, we can factorize any polynomial with complex coefficients into a product of linear factors with complex coefficients:

\[
f(z) = a_n(z - z_1)(z - z_2) \cdots (z - z_n).
\]

There are available a number of computational algorithms for finding good approximations to the roots of a polynomial with complex coefficients.

### 5.3 Geometric representation of \( \mathbb{C} \)

Complex numbers can be represented as points in the plane, using the correspondence \( x + iy \leftrightarrow (x, y) \). The representation is known as the Argand diagram or complex plane. The real complex numbers lie on the \( x \)-axis, which is then called the real axis, while the imaginary numbers lie on the \( y \)-axis, which is known as the imaginary axis. The complex numbers with positive imaginary part lie in the upper half plane, while those with negative imaginary part lie in the lower half plane.

Because of the equation

\[
(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2),
\]

complex numbers add vectorially, using the parallelogram law. Similarly, the complex number \( z_1 - z_2 \) can be represented by the vector from \((x_2, y_2)\) to \((x_1, y_1)\), where \( z_1 = x_1 + iy_1 \) and \( z_2 = x_2 + iy_2 \). (See Figure 5.1.)
Figure 5.1: Complex addition and subtraction.

The geometrical representation of complex numbers can be very useful when complex number methods are used to investigate properties of triangles and circles. It is very important in the branch of calculus known as Complex Function theory, where geometric methods play an important role.

We mention that the line through two distinct points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ has the form $z = (1 - t)z_1 + tz_2$, $t \in \mathbb{R}$, where $z = x + iy$ is any point on the line and $z_i = x_i + iy_i$, $i = 1, 2$. For the line has parametric equations

$$x = (1 - t)x_1 + tx_2, \quad y = (1 - t)y_1 + ty_2$$

and these can be combined into a single equation $z = (1 - t)z_1 + tz_2$.

Circles have various equation representations in terms of complex numbers, as will be seen later.

### 5.4 Complex conjugate

**DEFINITION 5.4.1 (Complex conjugate)** If $z = x + iy$, the complex conjugate of $z$ is the complex number defined by $\overline{z} = x - iy$. Geometrically, the complex conjugate of $z$ is obtained by reflecting $z$ in the real axis (see Figure 5.2).

The following properties of the complex conjugate are easy to verify:
5.4. COMPLEX CONJUGATE

Figure 5.2: The complex conjugate of $z$: $\overline{z}$.

1. $z_1 + z_2 = \overline{z_1} + \overline{z_2}$;
2. $-\overline{z} = -\overline{z}$.
3. $z_1 - z_2 = \overline{z_1} - \overline{z_2}$;
4. $z_1 \overline{z_2} = \overline{z_1 z_2}$;
5. $(1/z) = 1/\overline{z}$;
6. $(z_1/z_2) = \overline{z_1}/\overline{z_2}$;
7. $z$ is real if and only if $\overline{z} = z$;
8. With the standard convention that the real and imaginary parts are denoted by $\text{Re} z$ and $\text{Im} z$, we have
   \[ \text{Re} z = \frac{z + \overline{z}}{2}, \quad \text{Im} z = \frac{z - \overline{z}}{2i}; \]
9. If $z = x + iy$, then $z \overline{z} = x^2 + y^2$.

**THEOREM 5.4.1** If $f(z)$ is a polynomial with real coefficients, then its non–real roots occur in complex–conjugate pairs, i.e. if $f(z) = 0$, then $f(\overline{z}) = 0$.

**Proof.** Suppose $f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$, where $a_n, \ldots, a_0$ are real. Then
\[
0 = \overline{0} = \overline{f(z)} = \overline{a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0}
\]
\[
= a_n \overline{z^n} + a_{n-1} \overline{z^{n-1}} + \cdots + a_1 \overline{z} + a_0
\]
\[
= a_n \overline{z}^n + a_{n-1} \overline{z}^{n-1} + \cdots + a_1 \overline{z} + a_0
\]
\[
= f(\overline{z}).
\]
EXAMPLE 5.4.1 Discuss the position of the roots of the equation
\[ z^4 = -1 \]
in the complex plane.

**Solution.** The equation \( z^4 = -1 \) has real coefficients and so its roots come in complex conjugate pairs. Also if \( z \) is a root, so is \(-z\). Also there are clearly no real roots and no imaginary roots. So there must be one root \( w \) in the first quadrant, with all remaining roots being given by \( w \), \(-w\) and \(-\bar{w}\). In fact, as we shall soon see, the roots lie evenly spaced on the unit circle.

The following theorem is useful in deciding if a polynomial \( f(z) \) has a multiple root \( a \); that is if \((z - a)^m\) divides \( f(z) \) for some \( m \geq 2 \). (The proof is left as an exercise.)

**THEOREM 5.4.2** If \( f(z) = (z - a)^m g(z) \), where \( m \geq 2 \) and \( g(z) \) is a polynomial, then \( f'(a) = 0 \) and the polynomial and its derivative have a common root.

From theorem 5.4.1 we obtain a result which is very useful in the explicit integration of rational functions (i.e. ratios of polynomials) with real coefficients.

**THEOREM 5.4.3** If \( f(z) \) is a non–constant polynomial with real coefficients, then \( f(z) \) can be factorized as a product of real linear factors and real quadratic factors.

**Proof.** In general \( f(z) \) will have \( r \) real roots \( z_1, \ldots, z_r \) and \( 2s \) non–real roots \( z_{r+1}, \bar{z}_{r+1}, \ldots, z_{r+s}, \bar{z}_{r+s} \), occurring in complex–conjugate pairs by theorem 5.4.1. Then if \( a_n \) is the coefficient of highest degree in \( f(z) \), we have the factorization
\[
\begin{align*}
\ f(z) &= a_n(z - z_1) \cdots (z - z_r) \times \\
&\quad \times (z - z_{r+1})(z - \bar{z}_{r+1}) \cdots (z - z_{r+s})(z - \bar{z}_{r+s}).
\end{align*}
\]
We then use the following identity for \( j = r + 1, \ldots, r + s \) which in turn shows that paired terms give rise to real quadratic factors:
\[
\begin{align*}
(z - z_j)(z - \bar{z}_j) &= z^2 - (z_j + \bar{z}_j)z + z_j\bar{z}_j \\
&= z^2 - 2\text{Re}\ z_j + (x_j^2 + y_j^2),
\end{align*}
\]
where \( z_j = x_j + iy_j \).

A well–known example of such a factorization is the following:
EXAMPLE 5.4.2 Find a factorization of \( z^4 + 1 \) into real linear and quadratic factors.

Solution. Clearly there are no real roots. Also we have the preliminary factorization \( z^4 + 1 = (z^2 - i)(z^2 + i) \). Now the roots of \( z^2 - i \) are easily verified to be \( \pm(1 + i)/\sqrt{2} \), so the roots of \( z^2 + i \) must be \( \pm(1 - i)/\sqrt{2} \). In other words the roots are \( w = (1 + i)/\sqrt{2} \) and \( \bar{w}, -w, -\bar{w} \). Grouping conjugate–complex terms gives the factorization

\[
    z^4 + 1 = (z - w)(z - \bar{w})(z + w)(z + \bar{w}) \\
    = (z^2 - 2z\text{Re } w + w\bar{w})(z^2 + 2z\text{Re } w + w\bar{w}) \\
    = (z^2 - \sqrt{2}z + 1)(z^2 + \sqrt{2}z + 1).
\]

5.5 Modulus of a complex number

DEFINITION 5.5.1 (Modulus) If \( z = x + iy \), the modulus of \( z \) is the non–negative real number \(|z|\) defined by \(|z| = \sqrt{x^2 + y^2}\). Geometrically, the modulus of \( z \) is the distance from \( z \) to 0 (see Figure 5.3).

More generally, \(|z_1 - z_2|\) is the distance between \( z_1 \) and \( z_2 \) in the complex plane. For

\[
    |z_1 - z_2| = |(x_1 + iy_1) - (x_2 + iy_2)| = |(x_1 - x_2) + i(y_1 - y_2)| \\
    = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.
\]

The following properties of the modulus are easy to verify, using the identity \(|z|^2 = z\bar{z}\):

(i) \(|z_1z_2| = |z_1||z_2|\);
(ii) \( |z^{-1}| = |z|^{-1}; \)

(iii) \( \left| \frac{z_1}{z_2} \right| = \left| \frac{z_1}{|z_2|} \right| . \)

For example, to prove (i):

\[
|z_1 z_2|^2 = (z_1 z_2)\overline{z_1 z_2} = (z_1 \overline{z_2})(\overline{z_1} z_2) = (z_1 \overline{z_1})(z_2 \overline{z_2}) = |z_1|^2 |z_2|^2 = (|z_1||z_2|)^2 .
\]

Hence \( |z_1 z_2| = |z_1||z_2| . \)

EXAMPLE 5.5.1 Find \( |z| \) when \( z = \frac{(1 + i)^4}{(1 + 6i)(2 - 7i)} . \)

Solution.

\[
|z| = \frac{|1 + i|^4}{|1 + 6i||2 - 7i|} = \frac{(\sqrt{1^2 + 1^2})^4}{\sqrt{1^2 + 6^2}\sqrt{2^2 + (-7)^2}} = \frac{4}{\sqrt{37}\sqrt{53}} .
\]

THEOREM 5.5.1 (Ratio formulae) If \( z \) lies on the line through \( z_1 \) and \( z_2 \):

\[
z = (1 - t)z_1 + tz_2, \quad t \in \mathbb{R},
\]

we have the useful ratio formulae:

(i) \( \left| \frac{z - z_1}{z - z_2} \right| = \left| \frac{t}{1 - t} \right| \) if \( z \neq z_2 , \)

(ii) \( \left| \frac{z - z_1}{z_1 - z_2} \right| = |t| . \)

Circle equations. The equation \( |z - z_0| = r \), where \( z_0 \in \mathbb{C} \) and \( r > 0 \), represents the circle centre \( z_0 \) and radius \( r \). For example the equation \( |z - (1 + 2i)| = 3 \) represents the circle \( (x - 1)^2 + (y - 2)^2 = 9 . \)

Another useful circle equation is the circle of Apollonius:

\[
\left| \frac{z - a}{z - b} \right| = \lambda ,
\]
where $a$ and $b$ are distinct complex numbers and $\lambda$ is a positive real number, $\lambda \neq 1$. (If $\lambda = 1$, the above equation represents the perpendicular bisector of the segment joining $a$ and $b$.)

An algebraic proof that the above equation represents a circle, runs as follows. We use the following identities:

1. $|z - a|^2 = |z|^2 - 2\text{Re}(\overline{z}a) + |a|^2$
2. $\text{Re}(z_1 \pm z_2) = \text{Re}z_1 \pm \text{Re}z_2$
3. $\text{Re}(tz) = t\text{Re}z$ if $t \in \mathbb{R}$.

We have

\[
\left| \frac{z - a}{z - b} \right| = \lambda \iff |z - a|^2 = \lambda^2|z - b|^2
\]

\[
|z|^2 - 2\text{Re}\{\overline{z}a\} + |a|^2 = \lambda^2(|z|^2 - 2\text{Re}\{\overline{z}b\} + |b|^2)
\]

\[
(1 - \lambda^2)|z|^2 - 2\text{Re}\{\overline{z}(a - \lambda^2b)\} = \lambda^2|b|^2 - |a|^2
\]

\[
|z|^2 - 2\text{Re}\left\{\frac{a - \lambda^2b}{1 - \lambda^2}\right\} = \frac{\lambda^2|b|^2 - |a|^2}{1 - \lambda^2}
\]

\[
|z|^2 - 2\text{Re}\left\{\frac{a - \lambda^2b}{1 - \lambda^2}\right\} + \left|\frac{a - \lambda^2b}{1 - \lambda^2}\right|^2 = \frac{\lambda^2|b|^2 - |a|^2}{1 - \lambda^2} + \left|\frac{a - \lambda^2b}{1 - \lambda^2}\right|^2.
\]
Now it is easily verified that
\[ |a - \lambda^2 b|^2 + (1 - \lambda^2)(\lambda^2 |b|^2 - |a|^2) = \lambda^2 |a - b|^2. \]
So we obtain
\[ \frac{|z - a|}{|z - b|} = \lambda \iff \left| z - \left( \frac{a - \lambda^2 b}{1 - \lambda^2} \right) \right|^2 = \frac{\lambda^2 |a - b|^2}{|1 - \lambda^2|^2} \]
\[ \iff \left| z - \left( \frac{a - \lambda^2 b}{1 - \lambda^2} \right) \right| = \frac{\lambda |a - b|}{|1 - \lambda^2|}. \]
The last equation represents a circle centre \( z_0 \), radius \( r \), where
\[ z_0 = \frac{a - \lambda^2 b}{1 - \lambda^2} \quad \text{and} \quad r = \frac{\lambda |a - b|}{|1 - \lambda^2|}. \]
There are two special points on the circle of Apollonius, the points \( z_1 \) and \( z_2 \) defined by
\[ \frac{z_1 - a}{z_1 - b} = \lambda \quad \text{and} \quad \frac{z_2 - a}{z_2 - b} = -\lambda, \]
or
\[ z_1 = \frac{a - \lambda b}{1 - \lambda} \quad \text{and} \quad z_2 = \frac{a + \lambda b}{1 + \lambda}. \quad (5.3) \]
It is easy to verify that \( z_1 \) and \( z_2 \) are distinct points on the line through \( a \) and \( b \) and that \( z_0 = \frac{z_1 + z_2}{2} \). Hence the circle of Apollonius is the circle based on the segment \( z_1, z_2 \) as diameter.

**EXAMPLE 5.5.2** Find the centre and radius of the circle
\[ |z - 1 - i| = 2|z - 5 - 2i|. \]

**Solution.** Method 1. Proceed algebraically and simplify the equation
\[ |x + iy - 1 - i| = 2|x + iy - 5 - 2i| \]
or
\[ |x - 1 + i(y - 1)| = 2|x - 5 + i(y - 2)|. \]
Squaring both sides gives
\[ (x - 1)^2 + (y - 1)^2 = 4((x - 5)^2 + (y - 2)^2), \]
which reduces to the circle equation
\[ x^2 + y^2 - \frac{38}{3}x - \frac{14}{3}y + 38 = 0. \]
Completing the square gives
\[(x - \frac{19}{3})^2 + (y - \frac{7}{3})^2 = \left(\frac{19}{3}\right)^2 + \left(\frac{7}{3}\right)^2 - 38 = \frac{68}{9},\]
so the centre is \((\frac{19}{3}, \frac{7}{3})\) and the radius is \(\sqrt{\frac{68}{9}}\).

Method 2. Calculate the diametrical points \(z_1\) and \(z_2\) defined above by equations 5.3:

\[z_1 - 1 - i = 2(z_1 - 5 - 2i),\]
\[z_2 - 1 - i = -2(z_2 - 5 - 2i).\]

We find \(z_1 = 9 + 3i\) and \(z_2 = (11 + 5i)/3\). Hence the centre \(z_0\) is given by
\[z_0 = \frac{z_1 + z_2}{2} = \frac{19}{3} + \frac{7}{3}i\]
and the radius \(r\) is given by
\[r = |z_1 - z_0| = \left|\left(\frac{19}{3} + \frac{7}{3}i\right) - (9 + 3i)\right| = \left|\frac{8}{3} - \frac{2}{3}i\right| = \frac{\sqrt{68}}{3}.

5.6 Argument of a complex number

Let \(z = x + iy\) be a non–zero complex number, \(r = |z| = \sqrt{x^2 + y^2}\). Then we have \(x = r \cos \theta, y = r \sin \theta\), where \(\theta\) is the angle made by \(z\) with the positive \(x\)–axis. So \(\theta\) is unique up to addition of a multiple of \(2\pi\) radians.

DEFINITION 5.6.1 (Argument) Any number \(\theta\) satisfying the above pair of equations is called an argument of \(z\) and is denoted by \(\arg z\). The particular argument of \(z\) lying in the range \(-\pi < \theta \leq \pi\) is called the principal argument of \(z\) and is denoted by \(\text{Arg } z\) (see Figure 5.5).

We have \(z = r \cos \theta + ir \sin \theta = r(\cos \theta + i \sin \theta)\) and this representation of \(z\) is called the polar representation or modulus–argument form of \(z\).

EXAMPLE 5.6.1 Arg \(1 = 0\), Arg \((-1) = \pi\), Arg \(i = \frac{\pi}{2}\), Arg \((-i) = -\frac{\pi}{2}\).

We note that \(y/x = \tan \theta\) if \(x \neq 0\), so \(\theta\) is determined by this equation up to a multiple of \(\pi\). In fact
\[\text{Arg } z = \tan^{-1} \frac{y}{x} + k \pi,\]
where \( k = 0 \) if \( x > 0 \); \( k = 1 \) if \( x < 0, y > 0 \); \( k = -1 \) if \( x < 0, y < 0 \).

To determine \( \text{Arg} \ z \) graphically, it is simplest to draw the triangle formed by the points \( 0, x, z \) on the complex plane, mark in the positive acute angle \( \alpha \) between the rays \( 0, x \) and \( 0, z \) and determine \( \text{Arg} z \) geometrically, using the fact that \( \alpha = \tan^{-1}(\vert y/\vert x\vert) \), as in the following examples:

**EXAMPLE 5.6.2** Determine the principal argument of \( z \) for the following complex numbers:

\[ z = 4 + 3i, \quad -4 + 3i, \quad -4 - 3i, \quad 4 - 3i. \]

**Solution.** Referring to Figure 5.6, we see that \( \text{Arg} \ z \) has the values

\[ \alpha, \pi - \alpha, \quad -\pi + \alpha, \quad -\alpha, \]

where \( \alpha = \tan^{-1} \frac{3}{4} \).

An important property of the argument of a complex number states that the sum of the arguments of two non–zero complex numbers is an argument of their product:

**THEOREM 5.6.1** If \( \theta_1 \) and \( \theta_2 \) are arguments of \( z_1 \) and \( z_2 \), then \( \theta_1 + \theta_2 \) is an argument of \( z_1z_2 \).

**Proof.** Let \( z_1 \) and \( z_2 \) have polar representations \( z_1 = r_1(\cos \theta_1 + i \sin \theta_1) \) and \( z_2 = r_2(\cos \theta_2 + i \sin \theta_2) \). Then

\[
z_1z_2 = r_1(\cos \theta_1 + i \sin \theta_1)r_2(\cos \theta_2 + i \sin \theta_2) = r_1r_2(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 + i(\cos \theta_1 \sin \theta_2 + \sin \theta_1 \cos \theta_2)) = r_1r_2(\cos (\theta_1 + \theta_2) + i \sin (\theta_1 + \theta_2)),
\]

where \( \alpha = \tan^{-1} \frac{3}{4} \).
which is the polar representation of \( z_1 z_2 \), as \( r_1 r_2 = |z_1||z_2| = |z_1 z_2| \). Hence \( \theta_1 + \theta_2 \) is an argument of \( z_1 z_2 \).

An easy induction gives the following generalization to a product of \( n \) complex numbers:

**Corollary 5.6.1** If \( \theta_1, \ldots, \theta_n \) are arguments for \( z_1, \ldots, z_n \) respectively, then \( \theta_1 + \cdots + \theta_n \) is an argument for \( z_1 \cdots z_n \).

Taking \( \theta_1 = \cdots = \theta_n = \theta \) in the previous corollary gives

**Corollary 5.6.2** If \( \theta \) is an argument of \( z \), then \( n \theta \) is an argument for \( z^n \).

**Theorem 5.6.2** If \( \theta \) is an argument of the non–zero complex number \( z \), then \( -\theta \) is an argument of \( z^{-1} \).

**Proof.** Let \( \theta \) be an argument of \( z \). Then \( z = r(\cos \theta + i \sin \theta) \), where \( r = |z| \). Hence

\[
    z^{-1} = r^{-1}(\cos \theta + i \sin \theta)^{-1} \\
    = r^{-1}(\cos \theta - i \sin \theta) \\
    = r^{-1}(\cos(-\theta) + i \sin(-\theta)).
\]
Now \( r^{-1} = |z|^{-1} = |z^{-1}|, \) so \(-\theta\) is an argument of \(z^{-1}\).

**COROLLARY 5.6.3** If \(\theta_1\) and \(\theta_2\) are arguments of \(z_1\) and \(z_2\), then \(\theta_1 - \theta_2\) is an argument of \(z_1/z_2\).

In terms of principal arguments, we have the following equations:

(i) \(\text{Arg}(z_1 z_2) = \text{Arg}z_1 + \text{Arg}z_2 + 2k_1\pi\),
(ii) \(\text{Arg}(z^{-1}) = -\text{Arg}z + 2k_2\pi\),
(iii) \(\text{Arg}(z_1/z_2) = \text{Arg}z_1 - \text{Arg}z_2 + 2k_3\pi\),
(iv) \(\text{Arg}(z_1 \cdots z_n) = \text{Arg}z_1 + \cdots + \text{Arg}z_n + 2k_4\pi\),
(v) \(\text{Arg}(z^n) = n \text{Arg}z + 2k_5\pi\),

where \(k_1, k_2, k_3, k_4, k_5\) are integers.

In numerical examples, we can write (i), for example, as

\[\text{Arg}(z_1 z_2) \equiv \text{Arg}z_1 + \text{Arg}z_2.\]

**EXAMPLE 5.6.3** Find the modulus and principal argument of
\[z = \left(\frac{\sqrt{3} + i}{1 + i}\right)^{17}\]

and hence express \(z\) in modulus–argument form.

**Solution.** \(|z| = \left|\frac{\sqrt{3} + i}{1 + i}\right|^{17} = \frac{2^{17}}{\sqrt{2}} = 2^{17/2}.\)

\[\text{Arg}z \equiv 17\text{Arg}\left(\frac{\sqrt{3} + i}{1 + i}\right)\]
\[= 17(\text{Arg}(\sqrt{3} + i) - \text{Arg}(1 + i))\]
\[= 17\left(\frac{\pi}{6} - \frac{\pi}{4}\right) = -\frac{17\pi}{12}.\]

Hence \(\text{Arg}z = \left(-\frac{17\pi}{12}\right) + 2k\pi\), where \(k\) is an integer. We see that \(k = 1\) and hence \(\text{Arg}z = \frac{7\pi}{12}\). Consequently \(z = 2^{17/2} \left(\cos \frac{7\pi}{12} + i \sin \frac{7\pi}{12}\right)\).

**DEFINITION 5.6.2** If \(\theta\) is a real number, then we define \(e^{i\theta}\) by
\[e^{i\theta} = \cos \theta + i \sin \theta.\]

More generally, if \(z = x + iy\), then we define \(e^z\) by
\[e^z = e^x e^{iy}.\]
For example,

\[ e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{-i\pi/2} = -i. \]

The following properties of the complex exponential function are left as exercises:

**THEOREM 5.6.3**

(i) \[ e^{z_1}e^{z_2} = e^{z_1+z_2}, \]

(ii) \[ e^{z_1}\ldots e^{z_n} = e^{z_1+\ldots+z_n}, \]

(iii) \[ e^z \neq 0, \]

(iv) \[ (e^z)^{-1} = e^{-z}, \]

(v) \[ e^{z_1}/e^{z_2} = e^{z_1-z_2}, \]

(vi) \[ e^{\bar{z}} = \overline{e^z}. \]

**THEOREM 5.6.4** The equation

\[ e^z = 1 \]

has the complete solution \( z = 2k\pi i, k \in \mathbb{Z}. \)

**Proof.** First we observe that

\[ e^{2k\pi i} = \cos (2k\pi) + i \sin (2k\pi) = 1. \]

Conversely, suppose \( e^z = 1, z = x + iy. \) Then \( e^x(\cos y + i \sin y) = 1. \) Hence \( e^x \cos y = 1 \) and \( e^x \sin y = 0. \) Hence \( \sin y = 0 \) and so \( y = n\pi, n \in \mathbb{Z}. \) Then \( e^x \cos (n\pi) = 1, \) so \( e^x(-1)^n = 1, \) from which follows \( (-1)^n = 1 \) as \( e^x > 0. \) Hence \( n = 2k, k \in \mathbb{Z} \) and \( e^x = 1. \) Hence \( x = 0 \) and \( z = 2k\pi i. \)

### 5.7 De Moivre’s theorem

The next theorem has many uses and is a special case of theorem 5.6.3(ii). Alternatively it can be proved directly by induction on \( n. \)

**THEOREM 5.7.1 (De Moivre)** If \( n \) is a positive integer, then

\[ (\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta. \]

As a first application, we consider the equation \( z^n = 1. \)

**THEOREM 5.7.2** The equation \( z^n = 1 \) has \( n \) distinct solutions, namely the complex numbers \( \zeta_k = e^{2k\pi i/n}, k = 0, 1, \ldots, n-1. \) These lie equally spaced on the unit circle \( |z| = 1 \) and are obtained by starting at 1, moving round the circle anti-clockwise, incrementing the argument in steps of \( \frac{2\pi}{n}. \)

(See Figure 5.7)

We notice that the roots are the powers of the special root \( \zeta = e^{2\pi i/n}. \)
CHAPTER 5. COMPLEX NUMBERS

Figure 5.7: The $n$th roots of unity.

**Proof.** With $\zeta_k$ defined as above,

$$
\zeta_k^n = \left( e^{\frac{2k\pi i}{n}} \right)^n = e^{\frac{2k\pi i}{n}n} = 1,
$$

by De Moivre’s theorem. However $|\zeta_k| = 1$ and $\arg \zeta_k = \frac{2k\pi}{n}$, so the complex numbers $\zeta_k$, $k = 0, 1, \ldots, n - 1$, lie equally spaced on the unit circle. Consequently these numbers must be precisely all the roots of $z^n - 1$. For the polynomial $z^n - 1$, being of degree $n$ over a field, can have at most $n$ distinct roots in that field.

The more general equation $z^n = a$, where $a \in \mathbb{C}, a \neq 0$, can be reduced to the previous case:

Let $\alpha$ be argument of $z$, so that $a = |a|e^{i\alpha}$. Then if $w = |a|^{1/n}e^{i\alpha}$, we have

$$
w^n = \left( |a|^{1/n}e^{i\alpha} \right)^n = (|a|^{1/n})^n \left( e^{i\alpha} \right)^n = |a|e^{i\alpha} = a.
$$

So $w$ is a particular solution. Substituting for $a$ in the original equation, we get $z^n = w^n$, or $(z/w)^n = 1$. Hence the complete solution is $z/w =$
5.7. De Moivre’s Theorem

Figure 5.8: The roots of $z^n = a$.

e^{2k\pi i/n}$, $k = 0, 1, \ldots, n-1$, or

$$z_k = |a|^{1/n} e^{\frac{\alpha}{n} i} e^{\frac{2k\pi i}{n}} = |a|^{1/n} e^{\frac{i(\alpha+2k\pi)}{n}}, \quad (5.4)$$

$k = 0, 1, \ldots, n-1$. So the roots are equally spaced on the circle $|z| = |a|^{1/n}$ and are generated from the special solution having argument equal to $(\arg a)/n$, by incrementing the argument in steps of $2\pi/n$. (See Figure 5.8.)

**Example 5.7.1** Factorize the polynomial $z^5 - 1$ as a product of real linear and quadratic factors.

**Solution.** The roots are $1, e^{\frac{2\pi i}{5}}, e^{\frac{-2\pi i}{5}}, e^{\frac{4\pi i}{5}}, e^{\frac{-4\pi i}{5}}$, using the fact that non–real roots come in conjugate–complex pairs. Hence

$$z^5 - 1 = (z - 1)(z - e^{\frac{2\pi i}{5}})(z - e^{\frac{-2\pi i}{5}})(z - e^{\frac{4\pi i}{5}})(z - e^{\frac{-4\pi i}{5}}).$$

Now

$$(z - e^{\frac{2\pi i}{5}})(z - e^{\frac{-2\pi i}{5}}) = z^2 - z(e^{\frac{2\pi i}{5}} + e^{\frac{-2\pi i}{5}}) + 1 = z^2 - 2z \cos \frac{2\pi}{5} + 1.$$
Similarly
\[(z - e^{\frac{4\pi i}{5}})(z - e^{-\frac{4\pi i}{5}}) = z^2 - 2z \cos \frac{4\pi}{5} + 1.\]

This gives the desired factorization.

EXAMPLE 5.7.2 Solve \(z^3 = i\).

Solution. \(|i| = 1\) and \(\text{Arg } i = \frac{\pi}{2} = \alpha\). So by equation 5.4, the solutions are

\[z_k = |i|^{1/3} e^{i(\alpha + 2k\pi)/3}, k = 0, 1, 2.\]

First, \(k = 0\) gives

\[z_0 = e^{i\frac{\pi}{6}} = \cos \frac{\pi}{6} + i \sin \frac{\pi}{6} = \frac{\sqrt{3}}{2} + \frac{i}{2}.\]

Next, \(k = 1\) gives

\[z_1 = e^{i\frac{5\pi}{6}} = \cos \frac{5\pi}{6} + i \sin \frac{5\pi}{6} = -\frac{\sqrt{3}}{2} + \frac{i}{2}.\]

Finally, \(k = 2\) gives

\[z_1 = e^{i\frac{9\pi}{6}} = \cos \frac{9\pi}{6} + i \sin \frac{9\pi}{6} = -i.\]

We finish this chapter with two more examples of De Moivre’s theorem.

EXAMPLE 5.7.3 If

\[C = 1 + \cos \theta + \cdots + \cos (n-1)\theta,\]
\[S = \sin \theta + \cdots + \sin (n-1)\theta,\]

prove that

\[C = \frac{\sin \frac{n\theta}{2}}{\sin \frac{\theta}{2}} \cos \frac{(n-1)\theta}{2} \quad \text{and} \quad S = \frac{\sin \frac{n\theta}{2}}{\sin \frac{\theta}{2}} \sin \frac{(n-1)\theta}{2},\]

if \(\theta \neq 2k\pi, k \in \mathbb{Z} \).
Solution.

\[ C + iS = 1 + (\cos \theta + i \sin \theta) + \cdots + (\cos (n-1)\theta + i \sin (n-1)\theta) \]
\[ = 1 + e^{i\theta} + \cdots + e^{i(n-1)\theta} \]
\[ = 1 + z + \cdots + z^{n-1}, \text{ where } z = e^{i\theta} \]
\[ = \frac{1 - z^n}{1 - z}, \text{ if } z \neq 1, \text{ i.e. } \theta \neq 2k\pi, \]
\[ = \frac{1 - e^{in\theta}}{1 - e^{i\theta}} = \frac{\sin \frac{n\theta}{2}}{\tan \frac{n\theta}{2}} \]
\[ = e^{i(n-1)\frac{\theta}{2}} \sin \frac{n\theta}{2} \]

The result follows by equating real and imaginary parts.

**EXAMPLE 5.7.4** Express \( \cos n\theta \) and \( \sin n\theta \) in terms of \( \cos \theta \) and \( \sin \theta \), using the equation \( \cos \theta + \sin \theta = (\cos \theta + i \sin \theta)^n \).

**Solution.** The binomial theorem gives

\[ (\cos \theta + i \sin \theta)^n = \cos^n \theta + \binom{n}{1} \cos^{n-1} \theta (i \sin \theta) + \binom{n}{2} \cos^{n-2} \theta (i \sin \theta)^2 + \cdots \]
\[ + (i \sin \theta)^n. \]

Equating real and imaginary parts gives

\[ \cos n\theta = \cos^n \theta - \binom{n}{2} \cos^{n-2} \theta \sin^2 \theta + \cdots \]
\[ \sin n\theta = \binom{n}{1} \cos^{n-1} \theta \sin \theta - \binom{n}{3} \cos^{n-3} \theta \sin^3 \theta + \cdots . \]

### 5.8 PROBLEMS

1. Express the following complex numbers in the form \( x + iy, x, y \) real:

   (i) \((-3 + i)(14 - 2i)\); (ii) \(\frac{2 + 3i}{1 - 4i}\); (iii) \(\frac{(1 + 2i)^2}{1 - i}\).

   [Answers: (i) \(-40 + 20i\); (ii) \(-\frac{10}{17} + \frac{11}{17}i\); (iii) \(\frac{-7}{2} + \frac{i}{2}\).]

2. Solve the following equations:
112  

CHAPTER 5. COMPLEX NUMBERS

(i) \( iz + (2 - 10i)z = 3z + 2i, \)

(ii) \( (1 + i)z + (2 - i)w = -3i \)
\( (1 + 2i)z + (3 + i)w = 2 + 2i. \)

[Answers: (i) \( z = \frac{-9}{41} - \frac{41}{41}i; \) (ii) \( z = -1 + 5i, w = \frac{19}{5} - \frac{8}{5}i. \)]

3. Express \( 1 + (1 + i) + (1 + i)^2 + \ldots + (1 + i)^{99} \) in the form \( x + iy, x, y \) real. [Answer: \( 1 + 2^{50}i. \)]

4. Solve the equations: (i) \( z^2 = -8 - 6i; \) (ii) \( z^2 - (3 + i)z + 4 + 3i = 0. \)

[Answers: (i) \( z = \pm (1 - 3i); \) (ii) \( z = 2 - i, 1 + 2i. \)]

5. Find the modulus and principal argument of each of the following complex numbers:

(i) \( 4 + i; \) (ii) \( -\frac{3}{2} - \frac{1}{2}i; \) (iii) \( -1 + 2i; \) (iv) \( \frac{1}{2}(-1 + i\sqrt{3}). \)

[Answers: (i) \( \sqrt{17}, \tan^{-1}\frac{1}{4} \); (ii) \( \sqrt{\frac{10}{2}}, -\pi + \tan^{-1}\frac{1}{3} \); (iii) \( \sqrt{5}, \pi - \tan^{-1}\frac{1}{2}. \)]

6. Express the following complex numbers in modulus-argument form:

(i) \( z = (1 + i)(1 + i\sqrt{3})(\sqrt{3} - i). \)

(ii) \( z = \frac{(1 + i)^5(1 - i\sqrt{3})^5}{(\sqrt{3} + i)^4}. \)

[Answers: (i) \( 4\sqrt{2}(\cos \frac{5\pi}{12} + i\sin \frac{5\pi}{12}); \) (ii) \( 2^{7/2}(\cos \frac{11\pi}{12} + i\sin \frac{11\pi}{12}). \)]

7. (i) If \( z = 2(\cos \frac{\pi}{4} + i\sin \frac{\pi}{4}) \) and \( w = 3(\cos \frac{\pi}{6} + i\sin \frac{\pi}{6}) \), find the polar form of

(a) \( zw; \) (b) \( \frac{z}{w}; \) (c) \( w; \) (d) \( \frac{z^5}{w^2}. \)

(ii) Express the following complex numbers in the form \( x + iy: \)

(a) \( (1 + i)^{12}; \) (b) \( \left(\frac{1 + i}{\sqrt{2}}\right)^{-6}. \)

[Answers: (i): (a) \( 6(\cos \frac{5\pi}{12} + i\sin \frac{5\pi}{12}); \) (b) \( \frac{2}{3}(\cos \frac{\pi}{12} + i\sin \frac{\pi}{12}); \)
(c) \( \frac{3}{2}(\cos -\frac{\pi}{12} + i\sin -\frac{\pi}{12}); \) (d) \( \frac{32}{9}(\cos \frac{11\pi}{12} + i\sin \frac{11\pi}{12}); \)
(ii): (a) \(-64; \) (b) \(-i. \)]
8. Solve the equations:
   (i) $z^2 = 1 + i\sqrt{3}$; (ii) $z^4 = i$; (iii) $z^3 = -8i$; (iv) $z^4 = 2 - 2i$.

   [Answers: (i) $z = \pm (\sqrt{3} + i \sqrt{2})$; (ii) $i^k (\cos \frac{\pi}{8} + i \sin \frac{\pi}{8})$, $k = 0, 1, 2, 3$; (iii) $z = 2i, -\sqrt{3} - i, \sqrt{3} - i$; (iv) $z = i^k 2^\frac{3}{8} (\cos \frac{\pi}{16} - i \sin \frac{\pi}{16})$, $k = 0, 1, 2, 3$.]

9. Find the reduced row–echelon form of the complex matrix

   \[
   \begin{bmatrix}
   2 + i & -1 + 2i & 2 \\
   1 + i & -1 + i & 1 \\
   1 + 2i & -2 + i & 1 + i 
   \end{bmatrix}
   \]

   [Answer: \[
   \begin{bmatrix}
   1 & i & 0 \\
   0 & 0 & 1 \\
   0 & 0 & 0 
   \end{bmatrix}
   \].]

10. (i) Prove that the line equation $lx + my = n$ is equivalent to

   \[
   \bar{p}z + p\bar{z} = 2n,
   \]

   where $p = l + im$.

   (ii) Use (ii) to deduce that reflection in the straight line

   \[
   \bar{p}z + p\bar{z} = n
   \]

   is described by the equation

   \[
   \bar{p}w + p\bar{z} = n.
   \]

   [Hint: The complex number $l + im$ is perpendicular to the given line.]

   (iii) Prove that the line $|z - a| = |z - b|$ may be written as $\bar{p}z + p\bar{z} = n$, where $p = b - a$ and $n = |b|^2 - |a|^2$. Deduce that if $z$ lies on the Apollonius circle $|z-a| = \lambda$, then $w$, the reflection of $z$ in the line $|z - a| = |z - b|$, lies on the Apollonius circle $|z-a| = \frac{1}{\lambda}$.

11. Let $a$ and $b$ be distinct complex numbers and $0 < \alpha < \pi$.

   (i) Prove that each of the following sets in the complex plane represents a circular arc and sketch the circular arcs on the same diagram:
\[ \text{Arg} \frac{z - a}{z - b} = \alpha, -\alpha, \pi - \alpha, \alpha - \pi. \]

Also show that \( \text{Arg} \frac{z - a}{z - b} = \pi \) represents the line segment joining \( a \) and \( b \), while \( \text{Arg} \frac{z - a}{z - b} = 0 \) represents the remaining portion of the line through \( a \) and \( b \).

(ii) Use (i) to prove that four distinct points \( z_1, z_2, z_3, z_4 \) are concyclic or collinear, if and only if the cross–ratio
\[
\frac{z_4 - z_1}{z_4 - z_2} / \frac{z_3 - z_1}{z_3 - z_2}
\]
is real.

(iii) Use (ii) to derive Ptolemy’s Theorem: Four distinct points \( A, B, C, D \) are concyclic or collinear, if and only if one of the following holds:
\[
\begin{align*}
AB \cdot CD + BC \cdot AD &= AC \cdot BD \\
BD \cdot AC + AD \cdot BC &= AB \cdot CD \\
BD \cdot AC + AB \cdot CD &= AD \cdot BC.
\end{align*}
\]
Chapter 6

EIGENVALUES AND EIGENVECTORS

6.1 Motivation

We motivate the chapter on eigenvalues by discussing the equation

\[ ax^2 + 2hxy + by^2 = c, \]

where not all of \( a, h, b \) are zero. The expression \( ax^2 + 2hxy + by^2 \) is called a quadratic form in \( x \) and \( y \) and we have the identity

\[ ax^2 + 2hxy + by^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & h \\ h & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = X^tAX, \]

where \( X = \begin{bmatrix} x \\ y \end{bmatrix} \) and \( A = \begin{bmatrix} a & h \\ h & b \end{bmatrix} \). \( A \) is called the matrix of the quadratic form.

We now rotate the \( x, y \) axes anticlockwise through \( \theta \) radians to new \( x_1, y_1 \) axes. The equations describing the rotation of axes are derived as follows:

Let \( P \) have coordinates \((x, y)\) relative to the \( x, y \) axes and coordinates \((x_1, y_1)\) relative to the \( x_1, y_1 \) axes. Then referring to Figure 6.1:
\[ x = OQ = OP \cos (\theta + \alpha) \]
\[ = OP(\cos \theta \cos \alpha - \sin \theta \sin \alpha) \]
\[ = (OP \cos \alpha) \cos \theta - (OP \sin \alpha) \sin \theta \]
\[ = OR \cos \theta - PR \sin \theta \]
\[ = x_1 \cos \theta - y_1 \sin \theta. \]

Similarly \( y = x_1 \sin \theta + y_1 \cos \theta. \)

We can combine these transformation equations into the single matrix equation:
\[ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \]

or \( X = PY, \) where \( X = \begin{bmatrix} x \\ y \end{bmatrix}, \ Y = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \) and \( P = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}. \)

We note that the columns of \( P \) give the directions of the positive \( x_1 \) and \( y_1 \) axes. Also \( P \) is an orthogonal matrix – we have \( PP^t = I_2 \) and so \( P^{-1} = P^t. \)

The matrix \( P \) has the special property that \( \det P = 1. \)

A matrix of the type \( P = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \) is called a rotation matrix.

We shall show soon that any \( 2 \times 2 \) real orthogonal matrix with determinant
6.1. MOTIVATION

equal to 1 is a rotation matrix.

We can also solve for the new coordinates in terms of the old ones:

\[
\begin{bmatrix}
  x_1 \\
y_1
\end{bmatrix}
= \begin{bmatrix}
  \cos \theta & \sin \theta \\
  -\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
  x \\
y
\end{bmatrix},
\]

so \( x_1 = x \cos \theta + y \sin \theta \) and \( y_1 = -x \sin \theta + y \cos \theta \). Then

\[
X^tAX = (PY)^tA(PY) = Y^t(P^tAP)Y.
\]

Now suppose, as we later show, that it is possible to choose an angle \( \theta \) so that \( P^tAP \) is a diagonal matrix, say \( \text{diag}(\lambda_1, \lambda_2) \). Then

\[
X^tAX = \begin{bmatrix}
  x_1 & y_1
\end{bmatrix}
\begin{bmatrix}
  \lambda_1 & 0 \\
  0 & \lambda_2
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
y_1
\end{bmatrix}
= \lambda_1 x_1^2 + \lambda_2 y_1^2
\]

and relative to the new axes, the equation \( ax^2 + 2hxy + by^2 = c \) becomes \( \lambda_1 x_1^2 + \lambda_2 y_1^2 = c \), which is quite easy to sketch. This curve is symmetrical about the \( x_1 \) and \( y_1 \) axes, with \( P_1 \) and \( P_2 \), the respective columns of \( P \), giving the directions of the axes of symmetry.

Also it can be verified that \( P_1 \) and \( P_2 \) satisfy the equations

\[
AP_1 = \lambda_1 P_1 \quad \text{and} \quad AP_2 = \lambda_2 P_2.
\]

These equations force a restriction on \( \lambda_1 \) and \( \lambda_2 \). For if \( P_1 = \begin{bmatrix} u_1 \\ v_1 \end{bmatrix} \), the first equation becomes

\[
\begin{bmatrix}
  a & h \\
h & b
\end{bmatrix}
\begin{bmatrix}
  u_1 \\
v_1
\end{bmatrix}
= \lambda_1
\begin{bmatrix}
  u_1 \\
v_1
\end{bmatrix}
\text{ or }
\begin{bmatrix}
  a - \lambda_1 & h \\
h & b - \lambda_1
\end{bmatrix}
\begin{bmatrix}
  u_1 \\
v_1
\end{bmatrix}
= \begin{bmatrix}
  0 \\
0
\end{bmatrix}.
\]

Hence we are dealing with a homogeneous system of two linear equations in two unknowns, having a non–trivial solution \( (u_1, v_1) \). Hence

\[
\begin{vmatrix}
  a - \lambda_1 & h \\
h & b - \lambda_1
\end{vmatrix}
= 0.
\]

Similarly, \( \lambda_2 \) satisfies the same equation. In expanded form, \( \lambda_1 \) and \( \lambda_2 \) satisfy

\[
\lambda^2 - (a + b)\lambda + ab - h^2 = 0.
\]

This equation has real roots

\[
\lambda = \frac{a + b \pm \sqrt{(a + b)^2 - 4(ab - h^2)}}{2} = \frac{a + b \pm \sqrt{(a - b)^2 + 4h^2}}{2}
\]

(6.2)

(The roots are distinct if \( a \neq b \) or \( h \neq 0 \). The case \( a = b \) and \( h = 0 \) needs no investigation, as it gives an equation of a circle.)

The equation \( \lambda^2 - (a + b)\lambda + ab - h^2 = 0 \) is called the eigenvalue equation of the matrix \( A \).
CHAPTER 6. EIGENVALUES AND EIGENVECTORS

6.2 Definitions and examples

**DEFINITION 6.2.1 (Eigenvalue, eigenvector)**
Let \( A \) be a complex square matrix. Then if \( \lambda \) is a complex number and \( X \) a non–zero complex column vector satisfying \( AX = \lambda X \), we call \( X \) an eigenvector of \( A \), while \( \lambda \) is called an eigenvalue of \( A \). We also say that \( X \) is an eigenvector corresponding to the eigenvalue \( \lambda \).

So in the above example \( P_1 \) and \( P_2 \) are eigenvectors corresponding to \( \lambda_1 \) and \( \lambda_2 \), respectively. We shall give an algorithm which starts from the eigenvalues of \( A = \begin{bmatrix} a & b \\ h & b \end{bmatrix} \) and constructs a rotation matrix \( P \) such that \( P^t AP \) is diagonal.

As noted above, if \( \lambda \) is an eigenvalue of an \( n \times n \) matrix \( A \), with corresponding eigenvector \( X \), then \((A - \lambda I_n)X = 0\), with \( X \neq 0 \), so \( \det(A - \lambda I_n) = 0 \) and there are at most \( n \) distinct eigenvalues of \( A \).

Conversely if \( \det(A - \lambda I_n) = 0 \), then \((A - \lambda I_n)X = 0\) has a non–trivial solution \( X \) and so \( \lambda \) is an eigenvalue of \( A \) with \( X \) a corresponding eigenvector.

**DEFINITION 6.2.2 (Characteristic equation, polynomial)**
The equation \( \det(A - \lambda I_n) = 0 \) is called the characteristic equation of \( A \), while the polynomial \( \det(A - \lambda I_n) \) is called the characteristic polynomial of \( A \). The characteristic polynomial of \( A \) is often denoted by \( ch_A(\lambda) \).

Hence the eigenvalues of \( A \) are the roots of the characteristic polynomial of \( A \).

For a \( 2 \times 2 \) matrix \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \), it is easily verified that the characteristic polynomial is \( \lambda^2 - (\text{trace} A)\lambda + \det A \), where \( \text{trace} A = a + d \) is the sum of the diagonal elements of \( A \).

**EXAMPLE 6.2.1** Find the eigenvalues of \( A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \) and find all eigenvectors.

**Solution.** The characteristic equation of \( A \) is \( \lambda^2 - 4\lambda + 3 = 0 \), or

\[
(\lambda - 1)(\lambda - 3) = 0.
\]

Hence \( \lambda = 1 \) or 3. The eigenvector equation \((A - \lambda I_n)X = 0\) reduces to

\[
\begin{bmatrix}
2 - \lambda & 1 \\
1 & 2 - \lambda
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0
\end{bmatrix},
\]
or

\[(2 - \lambda)x + y = 0\]
\[x + (2 - \lambda)y = 0.\]

Taking \(\lambda = 1\) gives

\[x + y = 0\]
\[x + y = 0,\]

which has solution \(x = -y, y\) arbitrary. Consequently the eigenvectors corresponding to \(\lambda = 1\) are the vectors \([-y \\ y]\), with \(y \neq 0\).

Taking \(\lambda = 3\) gives

\[-x + y = 0\]
\[x - y = 0,\]

which has solution \(x = y, y\) arbitrary. Consequently the eigenvectors corresponding to \(\lambda = 3\) are the vectors \([y \\ y]\), with \(y \neq 0\).

Our next result has wide applicability:

**THEOREM 6.2.1** Let \(A\) be a \(2 \times 2\) matrix having distinct eigenvalues \(\lambda_1\) and \(\lambda_2\) and corresponding eigenvectors \(X_1\) and \(X_2\). Let \(P\) be the matrix whose columns are \(X_1\) and \(X_2\), respectively. Then \(P\) is non–singular and

\[P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.\]

**Proof.** Suppose \(AX_1 = \lambda_1 X_1\) and \(AX_2 = \lambda_2 X_2\). We show that the system of homogeneous equations

\[xX_1 + yX_2 = 0\]

has only the trivial solution. Then by theorem 2.5.10 the matrix \(P = [X_1|X_2]\) is non–singular. So assume

\[xX_1 + yX_2 = 0. \quad (6.3)\]

Then \(A(xX_1 + yX_2) = A0 = 0\), so \(x(AX_1) + y(AX_2) = 0\). Hence

\[x\lambda_1 X_1 + y\lambda_2 X_2 = 0. \quad (6.4)\]
Multiplying equation 6.3 by \( \lambda_1 \) and subtracting from equation 6.4 gives
\[
(\lambda_2 - \lambda_1)yX_2 = 0.
\]
Hence \( y = 0 \), as \( (\lambda_2 - \lambda_1) \neq 0 \) and \( X_2 \neq 0 \). Then from equation 6.3, \( xX_1 = 0 \) and hence \( x = 0 \).

Then the equations \( AX_1 = \lambda_1 X_1 \) and \( AX_2 = \lambda_2 X_2 \) give
\[
AP = A[X_1|X_2] = [AX_1|AX_2] = [\lambda_1 X_1|\lambda_2 X_2] = [X_1|X_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},
\]
so
\[
P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.
\]

**EXAMPLE 6.2.2** Let \( A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \) be the matrix of example 6.2.1. Then \( X_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \) and \( X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \) are eigenvectors corresponding to eigenvalues 1 and 3, respectively. Hence if \( P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \), we have
\[
P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.
\]

There are two immediate applications of theorem 6.2.1. The first is to the calculation of \( A^n \): If \( P^{-1}AP = \text{diag} (\lambda_1, \lambda_2) \), then \( A = P \text{diag} (\lambda_1, \lambda_2)P^{-1} \) and
\[
A^n = \left( P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1} \right)^n = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n P^{-1} = P \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} P^{-1}.
\]

The second application is to solving a system of linear differential equations
\[
\begin{align*}
\frac{dx}{dt} &= ax + by \\
\frac{dy}{dt} &= cx + dy,
\end{align*}
\]
where \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) is a matrix of real or complex numbers and \( x \) and \( y \) are functions of \( t \). The system can be written in matrix form as \( \dot{X} = AX \), where
\[
X = \begin{bmatrix} x \\ y \end{bmatrix} \text{ and } \dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix}.
\]
We make the substitution \( X = PY \), where \( Y = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \). Then \( x_1 \) and \( y_1 \) are also functions of \( t \) and

\[
\dot{X} = P\dot{Y} = AX = A(PY), \quad \text{so} \quad \dot{Y} = (P^{-1}AP)Y = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}Y.
\]

Hence \( \dot{x}_1 = \lambda_1 x_1 \) and \( \dot{y}_1 = \lambda_2 y_1 \).

These differential equations are well-known to have the solutions \( x_1 = x_1(0)e^{\lambda_1 t} \) and \( x_2 = x_2(0)e^{\lambda_2 t} \), where \( x_1(0) \) is the value of \( x_1 \) when \( t = 0 \).

\[\frac{dx}{dt} = kx, \quad \text{where} \ k \ \text{is a constant, then} \]

\[
\frac{d}{dt}(e^{-kt}x) = -ke^{-kt}x + e^{-kt}\frac{dx}{dt} = -ke^{-kt}x + e^{-kt}kx = 0.
\]

Hence \( e^{-kt}x \) is constant, so \( e^{-kt}x = e^{-k_0}x(0) = x(0). \) Hence \( x = x(0)e^{kt} \).

However \( \begin{bmatrix} x_1(0) \\ y_1(0) \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} \), so this determines \( x_1(0) \) and \( y_1(0) \) in terms of \( x(0) \) and \( y(0) \). Hence ultimately \( x \) and \( y \) are determined as explicit functions of \( t \), using the equation \( X = PY \).

**EXAMPLE 6.2.3** Let \( A = \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix} \). Use the eigenvalue method to derive an explicit formula for \( A^n \) and also solve the system of differential equations

\[
\frac{dx}{dt} = 2x - 3y \\
\frac{dy}{dt} = 4x - 5y,
\]

given \( x = 7 \) and \( y = 13 \) when \( t = 0 \).

**Solution.** The characteristic polynomial of \( A \) is \( \lambda^2 + 3\lambda + 2 \) which has distinct roots \( \lambda_1 = -1 \) and \( \lambda_2 = -2 \). We find corresponding eigenvectors \( X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \)

and \( X_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \). Hence if \( P = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \), we have \( P^{-1}AP = \text{diag}(-1, -2) \).

Hence

\[
A^n = (P\text{diag}(-1, -2)P^{-1})^n = P\text{diag}((-1)^n, (-2)^n)P^{-1} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & (-2)^n \end{bmatrix} \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}.
\]
To solve the differential equation system, make the substitution $X = PY$. Then $x = x_1 + 3y_1$, $y = x_1 + 4y_1$. The system then becomes
\[
\begin{align*}
\dot{x}_1 &= -x_1 \\
\dot{y}_1 &= -2y_1,
\end{align*}
\]
so $x_1 = x_1(0)e^{-t}$, $y_1 = y_1(0)e^{-2t}$. Now
\[
\begin{bmatrix} x_1(0) \\ y_1(0) \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 13 \end{bmatrix} = \begin{bmatrix} -11 \\ 6 \end{bmatrix},
\]
so $x_1 = -11e^{-t}$ and $y_1 = 6e^{-2t}$. Hence $x = -11e^{-t} + 3(6e^{-2t}) = -11e^{-t} + 18e^{-2t}$, $y = -11e^{-t} + 4(6e^{-2t}) = -11e^{-t} + 24e^{-2t}$.

For a more complicated example we solve a system of inhomogeneous recurrence relations.

**EXAMPLE 6.2.4** Solve the system of recurrence relations
\[
\begin{align*}
x_{n+1} &= 2x_n - y_n - 1 \\
y_{n+1} &= -x_n + 2y_n + 2,
\end{align*}
\]
given that $x_0 = 0$ and $y_0 = -1$.

**Solution.** The system can be written in matrix form as
\[
X_{n+1} = AX_n + B,
\]
where
\[
A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.
\]
It is then an easy induction to prove that
\[
X_n = A^n X_0 + (A^{n-1} + \cdots + A + I_2)B. \quad (6.5)
\]
Also it is easy to verify by the eigenvalue method that
\[ A^n = \frac{1}{2} \begin{bmatrix} 1 + 3^n & 1 - 3^n \\ 1 - 3^n & 1 + 3^n \end{bmatrix} = \frac{1}{2} U + \frac{3^n}{2} V, \]
where \( U = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \) and \( V = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}. \) Hence
\[ A^{n-1} + \cdots + A + I_2 = \frac{n}{2} U + \frac{(3^n - 1 + \cdots + 3 + 1)}{2} V = \frac{n}{2} U + \frac{(3^n - 1)}{4} V. \]
Then equation 6.5 gives
\[ X_n = \left( \frac{1}{2} U + \frac{3^n}{2} V \right) \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \left( \frac{n}{2} U + \frac{(3^n - 1)}{4} V \right) \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \]
which simplifies to
\[ \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} \frac{(2n + 1 - 3^n)}{4} \\ \frac{(2n - 5 + 3^n)}{4} \end{bmatrix}. \]
Hence \( x_n = (2n - 1 + 3^n)/4 \) and \( y_n = (2n - 5 + 3^n)/4. \)

**REMARK 6.2.1** If \((A - I_2)^{-1}\) existed (that is, if \( \text{det}(A - I_2) \neq 0 \), or equivalently, if 1 is not an eigenvalue of \( A \)), then we could have used the formula
\[ A^{n-1} + \cdots + A + I_2 = (A^n - I_2)(A - I_2)^{-1}. \]
However the eigenvalues of \( A \) are 1 and 3 in the above problem, so formula 6.6 cannot be used there.

Our discussion of eigenvalues and eigenvectors has been limited to \( 2 \times 2 \) matrices. The discussion is a more complicated for matrices of size greater than two and is best left to a second course in linear algebra. Nevertheless the following result is a useful generalization of theorem 6.2.1. The reader is referred to [28, page 350] for a proof.

**THEOREM 6.2.2** Let \( A \) be an \( n \times n \) matrix having distinct eigenvalues \( \lambda_1, \ldots, \lambda_n \) and corresponding eigenvectors \( X_1, \ldots, X_n \). Let \( P \) be the matrix whose columns are respectively \( X_1, \ldots, X_n \). Then \( P \) is non–singular and
\[
P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.
\]
Another useful result which covers the case where there are multiple eigenvalues is the following (The reader is referred to [28, pages 351–352] for a proof):

**THEOREM 6.2.3** Suppose the characteristic polynomial of $A$ has the factorization
\[
\det(\lambda I_n - A) = (\lambda - c_1)^{n_1} \cdots (\lambda - c_t)^{n_t},
\]
where $c_1, \ldots, c_t$ are the distinct eigenvalues of $A$. Suppose that for $i = 1, \ldots, t$, we have nullity $(c_i I_n - A) = n_i$. For each $i$, choose a basis $X_{i1}, \ldots, X_{in_i}$ for the eigenspace $N(c_i I_n - A)$. Then the matrix
\[
P = [X_{11} | \cdots | X_{1n_1} | \cdots | X_{t1} | \cdots | X_{tn_t}]
\]
is non–singular and $P^{-1}AP$ is the following diagonal matrix
\[
P^{-1}AP = \begin{bmatrix}
c_1 I_{n_1} & 0 & \cdots & 0 \\
0 & c_2 I_{n_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & c_t I_{n_t}
\end{bmatrix}.
\]
(The notation means that on the diagonal there are $n_1$ elements $c_1$, followed by $n_2$ elements $c_2$, \ldots, $n_t$ elements $c_t$.)

### 6.3 PROBLEMS

1. Let $A = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$. Find a non–singular matrix $P$ such that $P^{-1}AP = \text{diag}(1, 3)$ and hence prove that
\[
A^n = \frac{3^n - 1}{2} A + \frac{3 - 3^n}{2} I_2.
\]

2. If $A = \begin{bmatrix} 0.6 & 0.8 \\ 0.4 & 0.2 \end{bmatrix}$, prove that $A^n$ tends to a limiting matrix
\[
\begin{bmatrix}
2/3 & 2/3 \\
1/3 & 1/3
\end{bmatrix}
\]
as $n \to \infty$. 
3. Solve the system of differential equations
\[ \begin{align*}
\frac{dx}{dt} &= 3x - 2y \\
\frac{dy}{dt} &= 5x - 4y,
\end{align*} \]
given \( x = 13 \) and \( y = 22 \) when \( t = 0 \).
[Answer: \( x = 7e^t + 6e^{-2t} \), \( y = 7e^t + 15e^{-2t} \).]

4. Solve the system of recurrence relations
\[ \begin{align*}
x_{n+1} &= 3x_n - y_n \\
y_{n+1} &= -x_n + 3y_n,
\end{align*} \]
given that \( x_0 = 1 \) and \( y_0 = 2 \).
[Answer: \( x_n = 2^{n-1}(3 - 2^n), \ y_n = 2^{n-1}(3 + 2^n) \).]

5. Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) be a real or complex matrix with distinct eigenvalues \( \lambda_1, \lambda_2 \) and corresponding eigenvectors \( X_1, X_2 \). Also let \( P = [X_1|X_2] \).

(a) Prove that the system of recurrence relations
\[ \begin{align*}
x_{n+1} &= ax_n + by_n \\
y_{n+1} &= cx_n + dy_n
\end{align*} \]
has the solution
\[ \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \alpha \lambda_1^n X_1 + \beta \lambda_2^n X_2, \]
where \( \alpha \) and \( \beta \) are determined by the equation
\[ \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = P^{-1} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}. \]

(b) Prove that the system of differential equations
\[ \begin{align*}
\frac{dx}{dt} &= ax + by \\
\frac{dy}{dt} &= cx + dy
\end{align*} \]
has the solution
\[ \begin{bmatrix} x \\ y \end{bmatrix} = \alpha e^{\lambda_1 t} X_1 + \beta e^{\lambda_2 t} X_2, \]
where $\alpha$ and $\beta$ are determined by the equation

$$
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = P^{-1} \begin{bmatrix}
x(0) \\
y(0)
\end{bmatrix}.
$$

6. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real matrix with non-real eigenvalues $\lambda = a + ib$ and $\bar{\lambda} = a - ib$, with corresponding eigenvectors $X = U + iV$ and $\bar{X} = U - iV$, where $U$ and $V$ are real vectors. Also let $P$ be the real matrix defined by $P = [U|V]$. Finally let $a + ib = r e^{i\theta}$, where $r > 0$ and $\theta$ is real.

(a) Prove that

$$
AU = aU - bV,
$$
$$
AV = bU + aV.
$$

(b) Deduce that

$$
P^{-1}AP = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.
$$

(c) Prove that the system of recurrence relations

$$
x_{n+1} = ax_n + by_n
$$
$$
y_{n+1} = cx_n + dy_n
$$

has the solution

$$
\begin{bmatrix}
x_n \\
y_n
\end{bmatrix} = r^n \{(\alpha U + \beta V) \cos n\theta + (\beta U - \alpha V) \sin n\theta\},
$$

where $\alpha$ and $\beta$ are determined by the equation

$$
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = P^{-1} \begin{bmatrix}
x_0 \\
y_0
\end{bmatrix}.
$$

(d) Prove that the system of differential equations

$$
\frac{dx}{dt} = ax + by
$$
$$
\frac{dy}{dt} = cx + dy
$$
6.3. PROBLEMS

has the solution
\[
\begin{bmatrix}
x \\
y
\end{bmatrix} = e^{at}\{ (\alpha U + \beta V) \cos bt + (\beta U - \alpha V) \sin bt \},
\]
where \( \alpha \) and \( \beta \) are determined by the equation
\[
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = P^{-1}\begin{bmatrix}
x(0) \\
y(0)
\end{bmatrix}.
\]

[Hint: Let \( \begin{bmatrix}
x \\
y
\end{bmatrix} = P\begin{bmatrix}
x_1 \\
y_1
\end{bmatrix} \). Also let \( z = x_1 + iy_1 \). Prove that
\[
\dot{z} = (a - ib)z
\]
and deduce that
\[
x_1 + iy_1 = e^{at}(\alpha + i\beta)(\cos bt + i \sin bt).
\]
Then equate real and imaginary parts to solve for \( x_1, y_1 \) and hence \( x, y \).]

7. (The case of repeated eigenvalues.) Let \( A = \begin{bmatrix}
a & b \\
c & d
\end{bmatrix} \) and suppose
that the characteristic polynomial of \( A \), \( \lambda^2 - (a + d)\lambda + (ad - bc) \), has
a repeated root \( \alpha \). Also assume that \( A \neq \alpha I_2 \). Let \( B = A - \alpha I_2 \).

(i) Prove that \( (a - d)^2 + 4bc = 0 \).
(ii) Prove that \( B^2 = 0 \).
(iii) Prove that \( BX_2 \neq 0 \) for some vector \( X_2 \); indeed, show that \( X_2 \)
can be taken to be \( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \) or \( \begin{bmatrix} 0 \\ 1 \end{bmatrix} \).
(iv) Let \( X_1 = BX_2 \). Prove that \( P = [X_1|X_2] \) is non–singular,
\[
AX_1 = \alpha X_1 \text{ and } AX_2 = \alpha X_2 + X_1
\]
and deduce that
\[
P^{-1}AP = \begin{bmatrix}
\alpha & 1 \\
0 & \alpha
\end{bmatrix}.
\]

8. Use the previous result to solve system of the differential equations
\[
\frac{dx}{dt} = 4x - y
\]
\[
\frac{dy}{dt} = 4x + 8y,
\]
given that \( x = 1 = y \) when \( t = 0 \).

To solve the differential equation

\[
\frac{dx}{dt} - kx = f(t), \quad k \text{ a constant,}
\]

multiply throughout by \( e^{-kt} \), thereby converting the left–hand side to \( \frac{dx}{dt}(e^{-kt}x) \).

[Answer: \( x = (1 - 3t)e^{6t}, \ y = (1 + 6t)e^{6t}. \)]

9. Let

\[
A = \begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{2}
\end{bmatrix}.
\]

(a) Verify that \( \det(\lambda I_3 - A) \), the characteristic polynomial of \( A \), is given by

\[
(\lambda - 1)\lambda(\lambda - \frac{1}{4}).
\]

(b) Find a non–singular matrix \( P \) such that \( P^{-1}AP = \text{diag}(1, 0, \frac{1}{4}). \)

(c) Prove that

\[
A^n = \frac{1}{3} \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix} + \frac{1}{3 \cdot 4^n} \begin{bmatrix}
2 & 2 & -4 \\
-1 & -1 & 2 \\
-1 & -1 & 2
\end{bmatrix}
\]

if \( n \geq 1. \)

10. Let

\[
A = \begin{bmatrix}
5 & 2 & -2 \\
2 & 5 & -2 \\
-2 & -2 & 5
\end{bmatrix}.
\]

(a) Verify that \( \det(\lambda I_3 - A) \), the characteristic polynomial of \( A \), is given by

\[
(\lambda - 3)^2(\lambda - 9).
\]

(b) Find a non–singular matrix \( P \) such that \( P^{-1}AP = \text{diag}(3, 3, 9). \)


Chapter 7

Identifying second degree equations

7.1 The eigenvalue method

In this section we apply eigenvalue methods to determine the geometrical nature of the second degree equation

\[ ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0, \]  

(7.1)

where not all of \( a, h, b \) are zero.

Let \( A = \begin{bmatrix} a & h \\ h & b \end{bmatrix} \) be the matrix of the quadratic form \( ax^2 + 2hxy + by^2 \).

We saw in section 6.1, equation 6.2 that \( A \) has real eigenvalues \( \lambda_1 \) and \( \lambda_2 \), given by

\[ \lambda_1 = \frac{a + b - \sqrt{(a - b)^2 + 4h^2}}{2}, \quad \lambda_2 = \frac{a + b + \sqrt{(a - b)^2 + 4h^2}}{2}. \]

We show that it is always possible to rotate the \( x, y \) axes to \( x_1, x_2 \) axes whose positive directions are determined by eigenvectors \( X_1 \) and \( X_2 \) corresponding to \( \lambda_1 \) and \( \lambda_2 \) in such a way that relative to the \( x_1, y_1 \) axes, equation 7.1 takes the form

\[ a'x^2 + b'y^2 + 2g'x + 2f'y + c = 0. \]  

(7.2)

Then by completing the square and suitably translating the \( x_1, y_1 \) axes, to new \( x_2, y_2 \) axes, equation 7.2 can be reduced to one of several standard forms, each of which is easy to sketch. We need some preliminary definitions.
DEFINITION 7.1.1 (Orthogonal matrix) An $n \times n$ real matrix $P$ is called orthogonal if

$$P^t P = I_n.$$ 

It follows that if $P$ is orthogonal, then $\det P = \pm 1$. For

$$\det (P^t P) = \det P^t \det P = (\det P)^2,$$

so $(\det P)^2 = \det I_n = 1$. Hence $\det P = \pm 1$.

If $P$ is an orthogonal matrix with $\det P = 1$, then $P$ is called a proper orthogonal matrix.

**Theorem 7.1.1** If $P$ is a $2 \times 2$ orthogonal matrix with $\det P = 1$, then

$$P = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

for some $\theta$.

**Remark 7.1.1** Hence, by the discussion at the beginning of Chapter 6, if $P$ is a proper orthogonal matrix, the coordinate transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

represents a rotation of the axes, with new $x_1$ and $y_1$ axes given by the respective columns of $P$.

**Proof.** Suppose that $P^t P = I_2$, where $\Delta = \det P = 1$. Let

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$ 

Then the equation

$$P^t = P^{-1} = \frac{1}{\Delta} \text{adj} P$$

gives

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$ 

Hence $a = d$, $b = -c$ and so

$$P = \begin{bmatrix} a & -c \\ c & a \end{bmatrix},$$

where $a^2 + c^2 = 1$. But then the point $(a, c)$ lies on the unit circle, so $a = \cos \theta$ and $c = \sin \theta$, where $\theta$ is uniquely determined up to multiples of $2\pi$. 
7.1. THE EIGENVALUE METHOD

**DEFINITION 7.1.2** (Dot product). If \( X = \begin{bmatrix} a \\ b \end{bmatrix} \) and \( Y = \begin{bmatrix} c \\ d \end{bmatrix} \), then

\( X \cdot Y \), the *dot product* of \( X \) and \( Y \), is defined by

\[
X \cdot Y = ac + bd.
\]

The dot product has the following properties:

(i) \( X \cdot (Y + Z) = X \cdot Y + X \cdot Z \);

(ii) \( X \cdot Y = Y \cdot X \);

(iii) \( tX \cdot Y = t(X \cdot Y) \);

(iv) \( X \cdot X = a^2 + b^2 \) if \( X = \begin{bmatrix} a \\ b \end{bmatrix} \);

(v) \( X \cdot Y = X^tY \).

The *length* of \( X \) is defined by

\[
||X|| = \sqrt{a^2 + b^2} = (X \cdot X)^{1/2}.
\]

We see that \( ||X|| \) is the distance between the origin \( O = (0, 0) \) and the point \( (a, b) \).

**THEOREM 7.1.2** (Geometrical interpretation of the dot product)

Let \( A = (x_1, y_1) \) and \( B = (x_2, y_2) \) be points, each distinct from the origin \( O = (0, 0) \). Then if \( X = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \) and \( Y = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \), we have

\[
X \cdot Y = OA \cdot OB \cos \theta,
\]

where \( \theta \) is the angle between the rays \( OA \) and \( OB \).

**Proof.** By the cosine law applied to triangle \( OAB \), we have

\[
AB^2 = OA^2 + OB^2 - 2OA \cdot OB \cos \theta. \tag{7.3}
\]

Now \( AB^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 \), \( OA^2 = x_1^2 + y_1^2 \), \( OB^2 = x_2^2 + y_2^2 \).

Substituting in equation 7.3 then gives

\[
(x_2 - x_1)^2 + (y_2 - y_1)^2 = (x_1^2 + y_1^2) + (x_2^2 + y_2^2) - 2OA \cdot OB \cos \theta,
\]
which simplifies to give
\[ OA \cdot OB \cos \theta = x_1 x_2 + y_1 y_2 = X \cdot Y. \]

It follows from theorem 7.1.2 that if \( A = (x_1, y_1) \) and \( B = (x_2, y_2) \) are points distinct from \( O = (0, 0) \) and \( X = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \) and \( Y = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \), then \( X \cdot Y = 0 \) means that the rays \( OA \) and \( OB \) are perpendicular. This is the reason for the following definition:

**DEFINITION 7.1.3 (Orthogonal vectors)** Vectors \( X \) and \( Y \) are called orthogonal if
\[ X \cdot Y = 0. \]

There is also a connection with orthogonal matrices:

**THEOREM 7.1.3** Let \( P \) be a \( 2 \times 2 \) real matrix. Then \( P \) is an orthogonal matrix if and only if the columns of \( P \) are orthogonal and have unit length.

**Proof.** \( P \) is orthogonal if and only if \( P^t P = I_2 \). Now if \( P = [X_1|X_2] \), the matrix \( P^t P \) is an important matrix called the *Gram* matrix of the column vectors \( X_1 \) and \( X_2 \). It is easy to prove that
\[
P^t P = [X_1 \cdot X_2] = \begin{bmatrix} X_1 \cdot X_1 & X_1 \cdot X_2 \\ X_2 \cdot X_1 & X_2 \cdot X_2 \end{bmatrix}.
\]

Hence the equation \( P^t P = I_2 \) is equivalent to
\[
\begin{bmatrix} X_1 \cdot X_1 & X_1 \cdot X_2 \\ X_2 \cdot X_1 & X_2 \cdot X_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\]
or, equating corresponding elements of both sides:
\[
X_1 \cdot X_1 = 1, \ X_1 \cdot X_2 = 0, \ X_2 \cdot X_2 = 1,
\]
which says that the columns of \( P \) are orthogonal and of unit length.

The next theorem describes a fundamental property of real symmetric matrices and the proof generalizes to symmetric matrices of any size.

**THEOREM 7.1.4** If \( X_1 \) and \( X_2 \) are eigenvectors corresponding to distinct eigenvalues \( \lambda_1 \) and \( \lambda_2 \) of a real symmetric matrix \( A \), then \( X_1 \) and \( X_2 \) are orthogonal vectors.
Proof. Suppose
\[ AX_1 = \lambda_1 X_1, \ AX_2 = \lambda_2 X_2, \] (7.4)
where \( X_1 \) and \( X_2 \) are non–zero column vectors, \( A^t = A \) and \( \lambda_1 \neq \lambda_2 \).
We have to prove that \( X_1^t X_2 = 0 \). From equation 7.4,
\[ X_2^t AX_1 = \lambda_1 X_2^t X_1 \] (7.5)
and
\[ X_1^t AX_2 = \lambda_2 X_1^t X_2. \] (7.6)
From equation 7.5, taking transposes,
\[ (X_2^t AX_1)^t = (\lambda_1 X_2^t X_1)^t \]
so
\[ X_1^t A^t X_2 = \lambda_1 X_1^t X_2. \]
Hence
\[ X_1^t AX_2 = \lambda_1 X_2^t X_2. \] (7.7)
Finally, subtracting equation 7.6 from equation 7.7, we have
\[ (\lambda_1 - \lambda_2) X_1^t X_2 = 0 \]
and hence, since \( \lambda_1 \neq \lambda_2 \),
\[ X_1^t X_2 = 0. \]

**Theorem 7.1.5** Let \( A \) be a real \( 2 \times 2 \) symmetric matrix with distinct eigenvalues \( \lambda_1 \) and \( \lambda_2 \). Then a proper orthogonal \( 2 \times 2 \) matrix \( P \) exists such that
\[ P^t AP = \text{diag}(\lambda_1, \lambda_2). \]
Also the rotation of axes
\[ \begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \]
“diagonalizes” the quadratic form corresponding to \( A \):
\[ X^t AX = \lambda_1 x_1^2 + \lambda_2 y_1^2. \]
Proof. Let $X_1$ and $X_2$ be eigenvectors corresponding to $\lambda_1$ and $\lambda_2$. Then by theorem 7.1.4, $X_1$ and $X_2$ are orthogonal. By dividing $X_1$ and $X_2$ by their lengths (i.e. normalizing $X_1$ and $X_2$) if necessary, we can assume that $X_1$ and $X_2$ have unit length. Then by theorem 7.1.1, $P = [X_1 \vert X_2]$ is an orthogonal matrix. By replacing $X_1$ by $-X_1$, if necessary, we can assume that $\det P = 1$. Then by theorem 6.2.1, we have

$$P^t A P = P^{-1} A P = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$ 

Also under the rotation $X = PY$,

$$X^t A X = (PY)^t A (PY) = Y^t (P^t A P) Y = Y^t \text{diag}(\lambda_1, \lambda_2) Y = \lambda_1 x_1^2 + \lambda_2 y_1^2.$$

**EXAMPLE 7.1.1** Let $A$ be the symmetric matrix

$$A = \begin{bmatrix} 12 & -6 \\ -6 & 7 \end{bmatrix}.$$ 

Find a proper orthogonal matrix $P$ such that $P^t A P$ is diagonal.

**Solution.** The characteristic equation of $A$ is $\lambda^2 - 19\lambda + 48 = 0$, or

$$(\lambda - 16)(\lambda - 3) = 0.$$ 

Hence $A$ has distinct eigenvalues $\lambda_1 = 16$ and $\lambda_2 = 3$. We find corresponding eigenvectors

$$X_1 = \begin{bmatrix} -3 \\ 2 \end{bmatrix} \text{ and } X_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$ 

Now $||X_1|| = ||X_2|| = \sqrt{13}$. So we take

$$X_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} -3 \\ 2 \end{bmatrix} \text{ and } X_2 = \frac{1}{\sqrt{13}} \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$ 

Then if $P = [X_1 \vert X_2]$, the proof of theorem 7.1.5 shows that

$$P^t A P = \begin{bmatrix} 16 & 0 \\ 0 & 3 \end{bmatrix}.$$ 

However $\det P = -1$, so replacing $X_1$ by $-X_1$ will give $\det P = 1.$
7.1. THE EIGENVALUE METHOD

Figure 7.1: $12x^2 - 12xy + 7y^2 + 60x - 38y + 31 = 0$.

**REMARK 7.1.2 (A shortcut)** Once we have determined one eigenvector $X_1 = \begin{bmatrix} a \\ b \end{bmatrix}$, the other can be taken to be $\begin{bmatrix} -b \\ a \end{bmatrix}$, as these vectors are always orthogonal. Also $P = [X_1|X_2]$ will have $\det P = a^2 + b^2 > 0$.

We now apply the above ideas to determine the geometric nature of second degree equations in $x$ and $y$.

**EXAMPLE 7.1.2** Sketch the curve determined by the equation

$$12x^2 - 12xy + 7y^2 + 60x - 38y + 31 = 0.$$ 

**Solution.** With $P$ taken to be the proper orthogonal matrix defined in the previous example by

$$P = \begin{bmatrix} 3/\sqrt{13} & 2/\sqrt{13} \\ -2/\sqrt{13} & 3/\sqrt{13} \end{bmatrix},$$

then as theorem 7.1.1 predicts, $P$ is a rotation matrix and the transformation

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = PY = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$
or more explicitly

\[
x = \frac{3x_1 + 2y_1}{\sqrt{13}}, \quad y = \frac{-2x_1 + 3y_1}{\sqrt{13}},
\]

(7.8)

will rotate the \(x, y\) axes to positions given by the respective columns of \(P\).

(More generally, we can always arrange for the \(x_1\) axis to point either into the first or fourth quadrant.)

Now \(A = \begin{bmatrix} 12 & -6 \\ -6 & 7 \end{bmatrix}\) is the matrix of the quadratic form

\[12x^2 - 12xy + 7y^2,\]

so we have, by Theorem 7.1.5

\[12x^2 - 12xy + 7y^2 = 16x_1^2 + 3y_1^2.\]

Then under the rotation \(X = PY\), our original quadratic equation becomes

\[16x_1^2 + 3y_1^2 + \frac{60}{\sqrt{13}}(3x_1 + 2y_1) - \frac{38}{\sqrt{13}}(-2x_1 + 3y_1) + 31 = 0,
\]

or

\[16x_1^2 + 3y_1^2 + \frac{256}{\sqrt{13}}x_1 + \frac{6}{\sqrt{13}}y_1 + 31 = 0.
\]

Now complete the square in \(x_1\) and \(y_1\):

\[16 \left( x_1^2 + \frac{16}{\sqrt{13}}x_1 \right) + 3 \left( y_1^2 + \frac{2}{\sqrt{13}}y_1 \right) + 31 = 0,
\]

\[16 \left( x_1 + \frac{8}{\sqrt{13}} \right)^2 + 3 \left( y_1 + \frac{1}{\sqrt{13}} \right)^2 = 16 \left( \frac{8}{\sqrt{13}} \right)^2 + 3 \left( \frac{1}{\sqrt{13}} \right)^2 - 31
\]

\[= 48.\]

(7.9)

Then if we perform a translation of axes to the new origin \((x_1, y_1) = (-\frac{8}{\sqrt{13}}, -\frac{1}{\sqrt{13}})\):

\[x_2 = x_1 + \frac{8}{\sqrt{13}}, \quad y_2 = y_1 + \frac{1}{\sqrt{13}},\]

equation 7.9 reduces to

\[16x_2^2 + 3y_2^2 = 48,
\]

or

\[\frac{x_2^2}{3} + \frac{y_2^2}{16} = 1.
\]
This equation is now in one of the standard forms listed below as Figure 7.2 and is that of an ellipse whose centre is at \((x_2, y_2) = (0, 0)\) and whose axes of symmetry lie along the \(x_2, y_2\) axes. In terms of the original \(x, y\) coordinates, we find that the centre is \((x, y) = (-2, 1)\). Also \(Y = P^t X\), so equations 7.8 can be solved to give

\[
x_1 = \frac{3x_1 - 2y_1}{\sqrt{13}}, \quad y_1 = \frac{2x_1 + 3y_1}{\sqrt{13}}.
\]

Hence the \(y_2\)-axis is given by

\[
0 = x_2 = x_1 + \frac{8}{\sqrt{13}} = \frac{3x - 2y}{\sqrt{13}} + \frac{8}{\sqrt{13}},
\]

or \(3x - 2y + 8 = 0\). Similarly the \(x_2\) axis is given by \(2x + 3y + 1 = 0\).

This ellipse is sketched in Figure 7.1.

Figures 7.2, 7.3, 7.4 and 7.5 are a collection of standard second degree equations: Figure 7.2 is an ellipse; Figures 7.3 are hyperbolas (in both these examples, the asymptotes are the lines \(y = \pm \frac{b}{a}x\)); Figures 7.4 and 7.5 represent parabolas.

**EXAMPLE 7.1.3** Sketch \(y^2 - 4x - 10y - 7 = 0\).
CHAPTER 7. IDENTIFYING SECOND DEGREE EQUATIONS

Figure 7.3: (i) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$; (ii) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$, $0 < b$, $0 < a$.

Figure 7.4: (i) $y^2 = 4ax$, $a > 0$; (ii) $y^2 = 4ax$, $a < 0$. 
7.1. THE EIGENVALUE METHOD

Figure 7.5: (iii) $x^2 = 4ay$, $a > 0$; (iv) $x^2 = 4ay$, $a < 0$.

**Solution.** Complete the square:

$$y^2 - 10y + 25 - 4x - 32 = 0$$

$$(y - 5)^2 = 4x + 32 = 4(x + 8),$$

or $y_1^2 = 4x_1$, under the translation of axes $x_1 = x + 8$, $y_1 = y - 5$. Hence we get a parabola with vertex at the new origin $(x_1, y_1) = (0, 0)$, i.e. $(x, y) = (-8, 5)$.

The parabola is sketched in Figure 7.6.

**EXAMPLE 7.1.4** Sketch the curve $x^2 - 4xy + 4y^2 + 5y - 9 = 0$.

**Solution.** We have $x^2 - 4xy + 4y^2 = X^t AX$, where

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}.$$

The characteristic equation of $A$ is $\lambda^2 - 5\lambda = 0$, so $A$ has distinct eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 0$. We find corresponding unit length eigenvectors

$$X_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad X_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Then $P = [X_1 | X_2]$ is a proper orthogonal matrix and under the rotation of axes $X = PY$, or

$$x = \frac{x_1 + 2y_1}{\sqrt{5}}, \quad y = \frac{-2x_1 + y_1}{\sqrt{5}},$$
Figure 7.6: \( y^2 - 4x - 10y - 7 = 0 \).

we have

\[
x^2 - 4xy + 4y^2 = \lambda_1 x_1^2 + \lambda_2 y_1^2 = 5x_1^2.
\]

The original quadratic equation becomes

\[
5x_1 + \frac{\sqrt{5}}{\sqrt{3}}(-2x_1 + y_1) - 9 = 0
\]
\[
5(x_1^2 - \frac{2}{\sqrt{5}}x_1) + \sqrt{3}y_1 - 9 = 0
\]
\[
5(x_1 - \frac{1}{\sqrt{5}})^2 = 10 - \sqrt{5}y_1 = \sqrt{5}(y_1 - 2\sqrt{5}),
\]

or \( 5x_2^2 = -\frac{1}{\sqrt{5}}y_2 \), where the \( x_1, y_1 \) axes have been translated to \( x_2, y_2 \) axes using the transformation

\[
x_2 = x_1 - \frac{1}{\sqrt{5}}, \quad y_2 = y_1 - 2\sqrt{5}.
\]

Hence the vertex of the parabola is at \((x_2, y_2) = (0, 0)\), i.e. \((x_1, y_1) = (\frac{1}{\sqrt{5}}, 2\sqrt{5})\), or \((x, y) = (\frac{21}{5}, \frac{8}{5})\). The axis of symmetry of the parabola is the line \( x_2 = 0 \), i.e. \( x_1 = 1/\sqrt{5} \). Using the rotation equations in the form

\[
x_1 = \frac{x - 2y}{\sqrt{5}}
\]
7.2. A CLASSIFICATION ALGORITHM

There are several possible degenerate cases that can arise from the general second degree equation. For example $x^2 + y^2 = 0$ represents the point $(0, 0)$; $x^2 + y^2 = -1$ defines the empty set, as does $x^2 = -1$ or $y^2 = -1$; $x^2 = 0$ defines the line $x = 0$; $(x + y)^2 = 0$ defines the line $x + y = 0$; $x^2 - y^2 = 0$ defines the lines $x - y = 0$, $x + y = 0$; $x^2 = 1$ defines the parallel lines $x = \pm 1$; $(x + y)^2 = 1$ likewise defines two parallel lines $x + y = \pm 1$.

We state without proof a complete classification \(^1\) of the various cases

---

\(^1\) This classification forms the basis of a computer program which was used to produce the diagrams in this chapter. I am grateful to Peter Adams for his programming assistance.
that can possibly arise for the general second degree equation
\[ ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0. \] \hspace{1cm} (7.10)

It turns out to be more convenient to first perform a suitable translation of
axes, before rotating the axes. Let
\[
\Delta = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}, \quad C = ab - h^2, \quad A = bc - f^2, \quad B = ca - g^2.
\]

If \( C \neq 0 \), let
\[
\alpha = -\frac{\begin{vmatrix} g & h \\ f & b \end{vmatrix}}{C}, \quad \beta = -\frac{\begin{vmatrix} a & g \\ h & f \end{vmatrix}}{C}.
\] \hspace{1cm} (7.11)

**CASE 1.** \( \Delta = 0 \).

1. \( C \neq 0 \). Translate axes to the new origin \((\alpha, \beta)\), where \( \alpha \) and \( \beta \) are
given by equations 7.11:
\[ x = x_1 + \alpha, \quad y = y_1 + \beta. \]

Then equation 7.10 reduces to
\[ ax_1^2 + 2hxy + by_1^2 = 0. \]

(a) \( C > 0 \): **Single point** \((x, y) = (\alpha, \beta)\).

(b) \( C < 0 \): **Two non–parallel lines** intersecting in \((x, y) = (\alpha, \beta)\).
The lines are
\[
\frac{y - \beta}{x - \alpha} = \frac{-h \pm \sqrt{-C}}{b} \quad \text{if } b \neq 0,
\]
\[ x = \alpha \quad \text{and} \quad \frac{y - \beta}{x - \alpha} = \frac{a}{2bh} \quad \text{if } b = 0. \]

2. \( C = 0 \).

(a) \( h = 0 \).

(i) \( a = g = 0 \).

(A) \( A > 0 \): **Empty set**.

(B) \( A = 0 \): **Single line** \( y = -f/b \).
7.2. A CLASSIFICATION ALGORITHM

(C) \( A < 0 \): Two parallel lines

\[
y = \frac{-f \pm \sqrt{-A}}{b}
\]

(ii) \( b = f = 0 \).

(A) \( B > 0 \): Empty set.

(B) \( B = 0 \): Single line \( x = -g/a \).

(C) \( B < 0 \): Two parallel lines

\[
x = \frac{-g \pm \sqrt{-B}}{a}
\]

(b) \( h \neq 0 \).

(i) \( B > 0 \): Empty set.

(ii) \( B = 0 \): Single line \( ax + hy = -g \).

(iii) \( B < 0 \): Two parallel lines

\[
ax + hy = -g \pm \sqrt{-B}
\]

CASE 2. \( \Delta \neq 0 \).

(2.1) \( C \neq 0 \). Translate axes to the new origin \((\alpha, \beta)\), where \( \alpha \) and \( \beta \) are given by equations 7.11:

\[
x = x_1 + \alpha, \quad y = y_1 + \beta.
\]

Equation 7.10 becomes

\[
ax_1^2 + 2hx_1y_1 + by_1^2 = -\frac{\Delta}{C}.
\]

CASE 2.1(i) \( h = 0 \). Equation 7.12 becomes \( ax_1^2 + by_1^2 = \frac{-\Delta}{C} \).

(a) \( C < 0 \): Hyperbola.

(b) \( C > 0 \) and \( a\Delta > 0 \): Empty set.

(c) \( C > 0 \) and \( a\Delta < 0 \).

(i) \( a = b \): Circle, centre \((\alpha, \beta)\), radius \( \sqrt{\frac{g^2 + f^2 - ac}{a}} \).

(ii) \( a \neq b \): Ellipse.
CASE 2.1(ii) $h \neq 0$.

Rotate the $(x_1, y_1)$ axes with the new positive $x_2$-axis in the direction of

$$[(b - a + R)/2, -h],$$

where $R = \sqrt{(a - b)^2 + 4h^2}$.

Then equation 7.12 becomes

$$\lambda_1 x_2^2 + \lambda_2 y_2^2 = -\frac{\Delta}{C}. \quad (7.13)$$

where

$$\lambda_1 = (a + b - R)/2, \quad \lambda_2 = (a + b + R)/2,$$

Here $\lambda_1 \lambda_2 = C$.

(a) $C < 0$: Hyperbola.

Here $\lambda_2 > 0 > \lambda_1$ and equation 7.13 becomes

$$\frac{x_2^2}{u^2} - \frac{y_2^2}{v^2} = -\frac{\Delta}{|\Delta|},$$

where

$$u = \sqrt{\frac{|\Delta|}{C\lambda_1}}, \quad v = \sqrt{\frac{|\Delta|}{-C\lambda_2}}.$$

(b) $C > 0$ and $a\Delta > 0$: Empty set.

(c) $C > 0$ and $a\Delta < 0$: Ellipse.

Here $\lambda_1, \lambda_2, a, b$ have the same sign and $\lambda_1 \neq \lambda_2$ and equation 7.13 becomes

$$\frac{x_2^2}{u^2} + \frac{y_2^2}{v^2} = 1,$$

where

$$u = \sqrt{\frac{\Delta}{-C\lambda_1}}, \quad v = \sqrt{\frac{\Delta}{-C\lambda_2}}.$$

(2.1) $C = 0$.

(a) $h = 0$.

(i) $a = 0$: Then $b \neq 0$ and $g \neq 0$. Parabola with vertex

$$\left(\frac{-A}{2gb}, \frac{-f}{b}\right).$$
Translate axes to \((x_1, y_1)\) axes:

\[
y_1^2 = -\frac{2g}{b}x_1.
\]

(ii) \(b = 0\): Then \(a \neq 0\) and \(f \neq 0\). **Parabola** with vertex

\[
\left( -\frac{g}{a}, -\frac{B}{2fa} \right).
\]

Translate axes to \((x_1, y_1)\) axes:

\[
x_1^2 = -\frac{2f}{a}y_1.
\]

(b) \(h \neq 0\): **Parabola.** Let

\[
k = \frac{ga + bf}{a + b}.
\]

The vertex of the parabola is

\[
\left( \frac{(2akf - h^2 - hac)}{d}, \frac{a(k^2 + ac - 2kg)}{d} \right).
\]

Now translate to the vertex as the new origin, then rotate to \((x_2, y_2)\) axes with the positive \(x_2\)-axis along \([sa, -sh]\), where \(s = \text{sign}(a)\).

(The positive \(x_2\)-axis points into the first or fourth quadrant.) Then the parabola has equation

\[
x_2^2 = -\frac{2st}{\sqrt{a^2 + h^2}}y_2,
\]

where \(t = (af - gh)/(a + b)\).

**REMARK 7.2.1** If \(\Delta = 0\), it is not necessary to rotate the axes. Instead it is always possible to translate the axes suitably so that the coefficients of the terms of the first degree vanish.

**EXAMPLE 7.2.1** Identify the curve

\[
2x^2 + xy - y^2 + 6y - 8 = 0.
\]
CHAPTER 7. IDENTIFYING SECOND DEGREE EQUATIONS

Solution. Here
\[ \Delta = \begin{vmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & -1 & 3 \\ 0 & 3 & -8 \end{vmatrix} = 0. \]

Let \( x = x_1 + \alpha, \ y = y_1 + \beta \) and substitute in equation \( 7.14 \) to get
\[ 2(x_1 + \alpha)^2 + (x_1 + \alpha)(y_1 + \beta) - (y_1 + \beta)^2 + 4(y_1 + \beta) - 8 = 0. \quad (7.15) \]

Then equating the coefficients of \( x_1 \) and \( y_1 \) to 0 gives
\[
\begin{align*}
4\alpha + \beta &= 0 \\
\alpha + 2\beta + 4 &= 0,
\end{align*}
\]

which has the unique solution \( \alpha = -\frac{2}{3}, \ \beta = \frac{8}{3} \). Then equation \( 7.15 \) simplifies to
\[ 2x_1^2 + 2x_1y_1 - y_1^2 = 0 = (2x_1 - y_1)(x_1 + y_1), \]

so relative to the \( x_1, y_1 \) coordinates, equation \( 7.14 \) describes two lines: \( 2x_1 - y_1 = 0 \) or \( x_1 + y_1 = 0 \). In terms of the original \( x, y \) coordinates, these lines become \( 2(x + \frac{2}{3}) - (y - \frac{8}{3}) = 0 \) and \( (x + \frac{2}{3}) + (y - \frac{8}{3}) = 0 \), i.e. \( 2x - y + 4 = 0 \) and \( x + y - 2 = 0 \), which intersect in the point
\[ (x, y) = (\alpha, \beta) = \left(-\frac{2}{3}, \frac{8}{3}\right). \]

EXAMPLE 7.2.2 Identify the curve
\[ x^2 + 2xy + y^2 + 2x + 2y + 1 = 0. \quad (7.16) \]

Solution. Here
\[ \Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0. \]

Let \( x = x_1 + \alpha, \ y = y_1 + \beta \) and substitute in equation \( 7.16 \) to get
\[ (x_1 + \alpha)^2 + 2(x_1 + \alpha)(y_1 + \beta) + (y_1 + \beta)^2 + 2(x_1 + \alpha) + 2(y_1 + \beta) + 1 = 0. \quad (7.17) \]

Then equating the coefficients of \( x_1 \) and \( y_1 \) to 0 gives the same equation
\[ 2\alpha + 2\beta + 2 = 0. \]

Take \( \alpha = 0, \ \beta = -1 \). Then equation \( 7.17 \) simplifies to
\[ x_1^2 + 2x_1y_1 + y_1^2 = 0 = (x_1 + y_1)^2, \]

and in terms of \( x, y \) coordinates, equation \( 7.16 \) becomes
\[ (x + y + 1)^2 = 0, \ \text{or} \ x + y + 1 = 0. \]
7.3 PROBLEMS

1. Sketch the curves
   (i) \( x^2 - 8x + 8y + 8 = 0; \)
   (ii) \( y^2 - 12x + 2y + 25 = 0. \)

2. Sketch the hyperbola
   \[ 4xy - 3y^2 = 8 \]
   and find the equations of the asymptotes.
   [Answer: \( y = 0 \) and \( y = \frac{4}{3}x. \)]

3. Sketch the ellipse
   \[ 8x^2 - 4xy + 5y^2 = 36 \]
   and find the equations of the axes of symmetry.
   [Answer: \( y = 2x \) and \( x = -2y. \)]

4. Sketch the conics defined by the following equations. Find the centre when the conic is an ellipse or hyperbola, asymptotes if an hyperbola, the vertex and axis of symmetry if a parabola:
   (i) \( 4x^2 - 9y^2 - 24x - 36y - 36 = 0; \)
   (ii) \( 5x^2 - 4xy + 8y^2 + 4\sqrt{5}x - 16\sqrt{5}y + 4 = 0; \)
   (iii) \( 4x^2 + y^2 - 4xy - 10y - 19 = 0; \)
   (iv) \( 77x^2 + 78xy - 27y^2 + 70x - 30y + 29 = 0. \)
   [Answers: (i) hyperbola, centre \((3, -2)\), asymptotes \(2x - 3y - 12 = 0, 2x + 3y = 0; \)
   (ii) ellipse, centre \((0, \sqrt{5})\);
   (iii) parabola, vertex \((-\frac{7}{5}, -\frac{9}{5})\), axis of symmetry \(2x - y + 1 = 0; \)
   (iv) hyperbola, centre \((-\frac{1}{10}, \frac{7}{10})\), asymptotes \(7x + 9y + 7 = 0 \) and \(11x - 3y - 1 = 0. \)]

5. Identify the lines determined by the equations:
   (i) \( 2x^2 + y^2 + 3xy - 5x - 4y + 3 = 0; \)
(ii) \(9x^2 + y^2 - 6xy + 6x - 2y + 1 = 0;\)

(iii) \(x^2 + 4xy + 4y^2 - x - 2y - 2 = 0.\)

[Answers: (i) \(2x + y - 3 = 0\) and \(x + y - 1 = 0;\) (ii) \(3x - y + 1 = 0;\)
(iii) \(x + 2y + 1 = 0\) and \(x + 2y - 2 = 0.\)]
Chapter 8

THREE–DIMENSIONAL GEOMETRY

8.1 Introduction

In this chapter we present a vector–algebra approach to three–dimensional geometry. The aim is to present standard properties of lines and planes, with minimum use of complicated three–dimensional diagrams such as those involving similar triangles. We summarize the chapter:

Points are defined as ordered triples of real numbers and the distance between points $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ is defined by the formula

$$P_1P_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$ 

Directed line segments $\vec{AB}$ are introduced as three–dimensional column vectors: If $A = (x_1, y_1, z_1)$ and $B = (x_2, y_2, z_2)$, then

$$\vec{AB} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}.$$ 

If $P$ is a point, we let $\vec{OP}$ and call $P$ the position vector of $P$.

With suitable definitions of lines, parallel lines, there are important geometrical interpretations of equality, addition and scalar multiplication of vectors.

(i) Equality of vectors: Suppose $A, B, C, D$ are distinct points such that no three are collinear. Then $\vec{AB} = \vec{CD}$ if and only if $\vec{AB} \parallel \vec{CD}$ and $\vec{AC} \parallel \vec{BD}$ (See Figure 8.1.)
(ii) Addition of vectors obeys the parallelogram law. Let $A$, $B$, $C$ be non-collinear. Then

\[ \overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AD}, \]

where $D$ is the point such that $\overrightarrow{AB} \parallel \overrightarrow{CD}$ and $\overrightarrow{AC} \parallel \overrightarrow{BD}$. (See Figure 8.1.)

(iii) Scalar multiplication of vectors: Let $\overrightarrow{AP} = t \overrightarrow{AB}$, where $A$ and $B$ are distinct points. Then $P$ is on the line $AB$,

\[ \frac{AP}{AB} = |t| \]

and

(a) $P = A$ if $t = 0$, $P = B$ if $t = 1$;

(b) $P$ is between $A$ and $B$ if $0 < t < 1$;

(c) $B$ is between $A$ and $P$ if $1 < t$;

(d) $A$ is between $P$ and $B$ if $t < 0$.

(See Figure 8.2.)
The **dot product** $X \cdot Y$ of vectors $X = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ and $Y = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$, is defined by

$$X \cdot Y = a_1a_2 + b_1b_2 + c_1c_2.$$ 

The **length** $||X||$ of a vector $X$ is defined by

$$||X|| = (X \cdot X)^{1/2}$$

and the **Cauchy–Schwarz inequality** holds:

$$|X \cdot Y| \leq ||X|| \cdot ||Y||.$$ 

The **triangle inequality** for vector length now follows as a simple deduction:

$$||X + Y|| \leq ||X|| + ||Y||.$$ 

Using the equation

$$AB = ||\overrightarrow{AB}||,$$

we deduce the corresponding familiar **triangle inequality** for distance:

$$AB \leq AC + CB.$$
The angle \( \theta \) between two non-zero vectors \( X \) and \( Y \) is then defined by
\[
\cos \theta = \frac{X \cdot Y}{||X|| \cdot ||Y||}, \quad 0 \leq \theta \leq \pi.
\]
This definition makes sense. For by the Cauchy–Schwarz inequality,
\[
-1 \leq \frac{X \cdot Y}{||X|| \cdot ||Y||} \leq 1.
\]
Vectors \( X \) and \( Y \) are said to be perpendicular or orthogonal if \( X \cdot Y = 0 \). Vectors of unit length are called unit vectors. The vectors
\[
i = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad j = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\]
are unit vectors and every vector is a linear combination of \( i \), \( j \) and \( k \):
\[
\begin{bmatrix} a \\ b \\ c \end{bmatrix} = ai + bj + ck.
\]
Non-zero vectors \( X \) and \( Y \) are parallel or proportional if the angle between \( X \) and \( Y \) equals 0 or \( \pi \); equivalently if \( X = tY \) for some real number \( t \). Vectors \( X \) and \( Y \) are then said to have the same or opposite direction, according as \( t > 0 \) or \( t < 0 \).

We are then led to study straight lines. If \( A \) and \( B \) are distinct points, it is easy to show that \( AP + PB = AB \) holds if and only if
\[
\overrightarrow{AP} = t \overrightarrow{AB}, \quad \text{where } 0 \leq t \leq 1.
\]
A line is defined as a set consisting of all points \( P \) satisfying
\[
P = P_0 + tX, \quad t \in \mathbb{R} \quad \text{or equivalently } \overrightarrow{P_0P} = tX,
\]
for some fixed point \( P_0 \) and fixed non-zero vector \( X \) called a direction vector for the line.

Equivalently, in terms of coordinates,
\[
x = x_0 + ta, \quad y = y_0 + tb, \quad z = z_0 + tc,
\]
where \( P_0 = (x_0, y_0, z_0) \) and not all of \( a, b, c \) are zero.
There is then one and only one line passing through two distinct points \( A \) and \( B \). It consists of the points \( P \) satisfying

\[
\overrightarrow{AP} = t \overrightarrow{AB},
\]

where \( t \) is a real number.

The cross-product \( X \times Y \) provides us with a vector which is perpendicular to both \( X \) and \( Y \). It is defined in terms of the components of \( X \) and \( Y \):

Let \( X = a_1i + b_1j + c_1k \) and \( Y = a_2i + b_2j + c_2k \). Then

\[
X \times Y = a \hat{i} + b \hat{j} + c \hat{k},
\]

where

\[
a = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad b = -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, \quad c = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.
\]

The cross-product enables us to derive elegant formulae for the distance from a point to a line, the area of a triangle and the distance between two skew lines.

Finally, we turn to the geometrical concept of a plane in three-dimensional space.

A plane is a set of points \( P \) satisfying an equation of the form

\[
\mathbf{P} = \mathbf{P}_0 + s\mathbf{X} + t\mathbf{Y}, \quad s, t \in \mathbb{R},
\]

(8.1)

where \( \mathbf{X} \) and \( \mathbf{Y} \) are non-zero, non-parallel vectors.

In terms of coordinates, equation 8.1 takes the form

\[
x = x_0 + sa_1 + ta_2 \\
y = y_0 + sb_1 + tb_2 \\
z = z_0 + sc_1 + tc_2,
\]

where \( \mathbf{P}_0 = (x_0, y_0, z_0) \).

There is then one and only one plane passing through three non-collinear points \( A, B, C \). It consists of the points \( P \) satisfying

\[
\overrightarrow{AP} = s \overrightarrow{AB} + t \overrightarrow{AC},
\]

where \( s \) and \( t \) are real numbers.

The cross-product enables us to derive a concise equation for the plane through three non-collinear points \( A, B, C \), namely

\[
\overrightarrow{AP} \cdot (\overrightarrow{AB} \times \overrightarrow{AC}) = 0.
\]
When expanded, this equation has the form
\[ ax + by + cz = d, \]
where \( ai + bj + ck \) is a non-zero vector which is perpendicular to \( \overrightarrow{P_1P_2} \) for all points \( P_1, P_2 \) lying in the plane. Any vector with this property is said to be a normal to the plane.

It is then easy to prove that two planes with non-parallel normal vectors must intersect in a line.

We conclude the chapter by deriving a formula for the distance from a point to a plane.

### 8.2 Three-dimensional space

**Definition 8.2.1** Three-dimensional space is the set \( E^3 \) of ordered triples \( (x, y, z) \), where \( x, y, z \) are real numbers. The triple \( (x, y, z) \) is called a point in \( E^3 \) and we write \( P = (x, y, z) \). The numbers \( x, y, z \) are called, respectively, the \( x, y, z \) coordinates of \( P \).

The coordinate axes are the sets of points:
\[ \{(x, 0, 0)\} \quad (x\text{-axis}), \quad \{(0, y, 0)\} \quad (y\text{-axis}), \quad \{(0, 0, z)\} \quad (z\text{-axis}). \]

The only point common to all three axes is the origin \( O = (0, 0, 0) \).

The coordinate planes are the sets of points:
\[ \{(x, y, 0)\} \quad (xy\text{-plane}), \quad \{(0, y, z)\} \quad (yz\text{-plane}), \quad \{(x, 0, z)\} \quad (xz\text{-plane}). \]

The positive octant consists of the points \( (x, y, z) \), where \( x > 0, y > 0, z > 0 \).

We think of the points \( (x, y, z) \) with \( z > 0 \) as lying above the \( xy \)-plane, and those with \( z < 0 \) as lying beneath the \( xy \)-plane. A point \( P = (x, y, z) \) will be represented as in Figure 8.3. The point illustrated lies in the positive octant.

**Definition 8.2.2** The distance \( P_1P_2 \) between points \( P_1 = (x_1, y_1, z_1) \) and \( P_2 = (x_2, y_2, z_2) \) is defined by the formula
\[ P_1P_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \]

For example, if \( P = (x, y, z) \),
\[ OP = \sqrt{x^2 + y^2 + z^2}. \]
8.2. THREE–DIMENSIONAL SPACE

Figure 8.3: Representation of three-dimensional space.

Figure 8.4: The vector $\overrightarrow{AB}$. 
DEFINITION 8.2.3 If \( A = (x_1, y_1, z_1) \) and \( B = (x_2, y_2, z_2) \) we define the symbol \( \overrightarrow{AB} \) to be the column vector

\[
\overrightarrow{AB} = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix}.
\]

We let \( \overrightarrow{OP} \) and call \( \overrightarrow{P} \) the position vector of \( P \).

The components of \( \overrightarrow{AB} \) are the coordinates of \( B \) when the axes are translated to \( A \) as origin of coordinates.

We think of \( \overrightarrow{AB} \) as being represented by the directed line segment from \( A \) to \( B \) and think of it as an arrow whose tail is at \( A \) and whose head is at \( B \). (See Figure 8.4.)

Some mathematicians think of \( \overrightarrow{AB} \) as representing the translation of space which takes \( A \) into \( B \).

The following simple properties of \( \overrightarrow{AB} \) are easily verified and correspond to how we intuitively think of directed line segments:

(i) \( \overrightarrow{AB} = 0 \iff A = B \);

(ii) \( \overrightarrow{BA} = -\overrightarrow{AB} \);

(iii) \( \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC} \) (the triangle law);

(iv) \( \overrightarrow{BC} = \overrightarrow{AC} - \overrightarrow{AB} = \overrightarrow{C} - \overrightarrow{B} \);

(v) if \( X \) is a vector and \( A \) a point, there is exactly one point \( B \) such that \( \overrightarrow{AB} = X \), namely that defined by \( \overrightarrow{B} = \overrightarrow{A} + X \).

To derive properties of the distance function and the vector function \( \overrightarrow{P_1P_2} \), we need to introduce the dot product of two vectors in \( \mathbb{R}^3 \).

### 8.3 Dot product

DEFINITION 8.3.1 If \( X = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \) and \( Y = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix} \), then \( X \cdot Y \), the dot product of \( X \) and \( Y \), is defined by

\[
X \cdot Y = a_1a_2 + b_1b_2 + c_1c_2.
\]
The dot product has the following properties:

(i) \( X \cdot (Y + Z) = X \cdot Y + X \cdot Z; \)

(ii) \( X \cdot Y = Y \cdot X; \)

(iii) \( (tX) \cdot Y = t(X \cdot Y); \)

(iv) \( X \cdot X = a^2 + b^2 + c^2 \) if \( X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}; \)

(v) \( X \cdot Y = X^t Y; \)

(vi) \( X \cdot X = 0 \) if and only if \( X = 0. \)

The length of \( X \) is defined by

\[ ||X|| = \sqrt{a^2 + b^2 + c^2} = (X \cdot X)^{1/2}. \]

We see that \( ||P|| = OP \) and more generally \( || \text{vec}_1 \text{vec}_2 || = \text{vec}_1 \text{vec}_2 \), the distance between \( \text{vec}_1 \) and \( \text{vec}_2 \).
CHAPTER 8. THREE–DIMENSIONAL GEOMETRY

Vectors having unit length are called unit vectors.

The vectors

$$
\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
$$

are unit vectors. Every vector is a linear combination of \( \mathbf{i} \), \( \mathbf{j} \) and \( \mathbf{k} \):

$$
\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}.
$$

(See Figure 8.7.)

It is easy to prove that

$$
||tX|| = |t| \cdot ||X||,
$$

if \( t \) is a real number. Hence if \( X \) is a non–zero vector, the vectors

$$
\pm \frac{1}{||X||} X
$$

are unit vectors.

A useful property of the length of a vector is

$$
||X \pm Y||^2 = ||X||^2 \pm 2X \cdot Y + ||Y||^2. \quad (8.2)
$$
8.3. DOT PRODUCT

The following important property of the dot product is widely used in mathematics:

**THEOREM 8.3.1 (The Cauchy–Schwarz inequality)**

If $X$ and $Y$ are vectors in $\mathbb{R}^3$, then

$$|X \cdot Y| \leq ||X|| \cdot ||Y||. \quad (8.3)$$

Moreover if $X \neq 0$ and $Y \neq 0$, then

$$X \cdot Y = ||X|| \cdot ||Y|| \Leftrightarrow Y = tX, \; t > 0,$$
$$X \cdot Y = -||X|| \cdot ||Y|| \Leftrightarrow Y = tX, \; t < 0.$$

**Proof.** If $X = 0$, then inequality 8.3 is trivially true. So assume $X \neq 0$. Now if $t$ is any real number, by equation 8.2,

$$0 \leq ||tX - Y||^2 = ||tX||^2 - 2(tX) \cdot Y + ||Y||^2$$
$$= t^2||X||^2 - 2(X \cdot Y)t + ||Y||^2$$
$$= at^2 - 2bt + c,$$

where $a = ||X||^2 > 0$, $b = X \cdot Y$, $c = ||Y||^2$.

Hence

$$a(t^2 - \frac{2bt}{a} + \frac{c}{a}) \geq 0$$
$$\left( t - \frac{b}{a} \right)^2 + \frac{ca - b^2}{a^2} \geq 0.$$

Substituting $t = b/a$ in the last inequality then gives

$$\frac{ac - b^2}{a^2} \geq 0,$$

so

$$|b| \leq \sqrt{ac} = \sqrt{a\sqrt{c}}$$

and hence inequality 8.3 follows.

To discuss equality in the Cauchy–Schwarz inequality, assume $X \neq 0$ and $Y \neq 0$.

Then if $X \cdot Y = ||X|| \cdot ||Y||$, we have for all $t$

$$||tX - Y||^2 = t^2||X||^2 - 2tX \cdot Y + ||Y||^2$$
$$= t^2||X||^2 - 2t||X|| \cdot ||Y|| + ||Y||^2$$
$$= ||tX - Y||^2.$$

Taking $t = ||X||/||Y||$ then gives $||tX - Y||^2 = 0$ and hence $tX - Y = 0$. Hence $Y = tX$, where $t > 0$. The case $X \cdot Y = -||X|| \cdot ||Y||$ is proved similarly.
COROLLARY 8.3.1 (The triangle inequality for vectors)
If $X$ and $Y$ are vectors, then
\[ ||X + Y|| \leq ||X|| + ||Y||. \] (8.4)
Moreover if $X \neq 0$ and $Y \neq 0$, then equality occurs in inequality 8.4 if and only if $Y = tX$, where $t > 0$.

**Proof.**
\[
||X + Y||^2 = ||X||^2 + 2X \cdot Y + ||Y||^2 \\
\leq ||X||^2 + 2||X|| \cdot ||Y|| + ||Y||^2 \\
= (||X|| + ||Y||)^2
\]
and inequality 8.4 follows.
If $||X + Y|| = ||X|| + ||Y||$, then the above proof shows that
\[ X \cdot Y = ||X|| \cdot ||Y||. \]
Hence if $X \neq 0$ and $Y \neq 0$, the first case of equality in the Cauchy–Schwarz inequality shows that $Y = tX$ with $t > 0$.

The triangle inequality for vectors gives rise to a corresponding inequality for the distance function:

THEOREM 8.3.2 (The triangle inequality for distance)
If $A$, $B$, $C$ are points, then
\[ AC \leq AB + BC. \] (8.5)
Moreover if $B \neq A$ and $B \neq C$, then equality occurs in inequality 8.5 if and only if $\overrightarrow{AB} = r \overrightarrow{AC}$, where $0 < r < 1$.

**Proof.**
\[
AC = ||\overrightarrow{AC}|| = ||\overrightarrow{AB} + \overrightarrow{BC}|| \\
\leq ||\overrightarrow{AB}|| + ||\overrightarrow{BC}|| \\
= AB + BC.
\]
Moreover if equality occurs in inequality 8.5 and $B \neq A$ and $B \neq C$, then $X = \overrightarrow{AB} \neq 0$ and $Y = \overrightarrow{BC} \neq 0$ and the equation $AC = AB + BC$ becomes
8.4. LINES

\[ ||X + Y|| = ||X|| + ||Y||. \] Hence the case of equality in the vector triangle inequality gives

\[ Y = \overrightarrow{BC} = t \overrightarrow{X}, \quad t > 0. \]

Then

\[
\begin{align*}
\overrightarrow{BC} &= \overrightarrow{AC} - \overrightarrow{AB} = t \overrightarrow{AB} \\
\overrightarrow{AC} &= (1 + t) \overrightarrow{AB} \\
\overrightarrow{AB} &= r \overrightarrow{AC},
\end{align*}
\]

where \( r = 1/(t + 1) \) satisfies \( 0 < r < 1 \).

8.4 Lines

**DEFINITION 8.4.1** A line in \( E^3 \) is the set \( \mathcal{L}(P_0, X) \) consisting of all points \( P \) satisfying

\[
P = P_0 + tX, \quad t \in \mathbb{R} \quad \text{or equivalently} \quad \overrightarrow{P_0}P = tX, \quad (8.6)
\]

for some fixed point \( P_0 \) and fixed non-zero vector \( X \). (See Figure 8.8.)

Equivalently, in terms of coordinates, equation 8.6 becomes

\[
x = x_0 + ta, \ y = y_0 + tb, \ z = z_0 + tc,
\]

where not all of \( a, b, c \) are zero.

The following familiar property of straight lines is easily verified.

**THEOREM 8.4.1** If \( A \) and \( B \) are distinct points, there is one and only one line containing \( A \) and \( B \), namely \( \mathcal{L}(A, \overrightarrow{AB}) \) or more explicitly the line defined by \( \overrightarrow{AP} = t \overrightarrow{AB} \), or equivalently, in terms of position vectors:

\[
P = (1 - t)A + tB \quad \text{or} \quad P = A + t \overrightarrow{AB}, \quad (8.7)
\]

Equations 8.7 may be expressed in terms of coordinates: if \( A = (x_1, y_1, z_1) \) and \( B = (x_2, y_2, z_2) \), then

\[
x = (1 - t)x_1 + tx_2, \ y = (1 - t)y_1 + ty_2, \ z = (1 - t)z_1 + tz_2.
\]
Figure 8.8: Representation of a line.

Figure 8.9: The line segment $AB$. 

$$\overrightarrow{P_0P} = t \overrightarrow{CD}$$

$$\mathbf{P} = \mathbf{A} + t \overrightarrow{AB}, \ 0 < t < 1$$
There is an important geometric significance in the number \( t \) of the above equation of the line through \( A \) and \( B \). The proof is left as an exercise:

**THEOREM 8.4.2 (Joachimsthal’s ratio formulae)**

If \( t \) is the parameter occurring in theorem 8.4.1, then

(i) \(|t| = \frac{AP}{AB};
(ii) \left| \frac{t}{1-t} \right| = \frac{AP}{PB} \quad \text{if} \ P \neq B.

Also

(iii) \( P \) is between \( A \) and \( B \) if \( 0 < t < 1 \);
(iv) \( B \) is between \( A \) and \( P \) if \( 1 < t \);
(v) \( A \) is between \( P \) and \( B \) if \( t < 0 \).

(See Figure 8.9.)

For example, \( t = \frac{1}{2} \) gives the mid–point \( P \) of the segment \( AB \):

\[ P = \frac{1}{2}(A + B). \]

**EXAMPLE 8.4.1** \( L \) is the line \( AB \), where \( A = (-4, 3, 1) \), \( B = (1, 1, 0) \);
\( M \) is the line \( CD \), where \( C = (2, 0, 2) \), \( D = (-1, 3, -2) \);
\( N \) is the line \( EF \), where \( E = (1, 4, 7) \), \( F = (-4, -3, -13) \). Find which pairs of lines intersect and also the points of intersection.

**Solution.** In fact only \( L \) and \( N \) intersect, in the point \((-\frac{2}{3}, \frac{5}{3}, \frac{1}{3})\). For example, to determine if \( L \) and \( N \) meet, we start with vector equations for \( L \) and \( N \):

\[ P = A + t \overrightarrow{AB}, \quad Q = E + s \overrightarrow{EF}, \]

equate \( P \) and \( Q \) and solve for \( s \) and \( t \):

\[ (-4i + 3j + k) + t(5i - 2j - k) = (i + 4j + 7k) + s(-5i - 7j - 20k), \]

which on simplifying, gives

\[
\begin{align*}
5t + 5s &= 5 \\
-2t + 7s &= 1 \\
-t + 20s &= 6
\end{align*}
\]

This system has the unique solution \( t = \frac{2}{7}, s = \frac{1}{7} \) and this determines a corresponding point \( P \) where the lines meet, namely \( P = (-\frac{2}{3}, \frac{5}{3}, \frac{1}{3}) \).

The same method yields inconsistent systems when applied to the other pairs of lines.
EXAMPLE 8.4.2 If $A = (5, 0, 7)$ and $B = (2, -3, 6)$, find the points $P$ on the line $AB$ which satisfy $AP/PB = 3$.

Solution. Use the formulae

$$P = A + t \overrightarrow{AB} \text{ and } \left| \frac{t}{1-t} \right| = \frac{AP}{PB} = 3.$$ 

Then

$$\frac{t}{1-t} = 3 \text{ or } -3,$$

so $t = \frac{3}{4}$ or $t = \frac{3}{2}$. The corresponding points are $(\frac{11}{4}, \frac{9}{4}, \frac{25}{4})$ and $(\frac{1}{2}, \frac{9}{2}, \frac{11}{2})$.

**DEFINITION 8.4.2** Let $X$ and $Y$ be non–zero vectors. Then $X$ is parallel or proportional to $Y$ if $X = tY$ for some $t \in \mathbb{R}$. We write $X \parallel Y$ if $X$ is parallel to $Y$. If $X = tY$, we say that $X$ and $Y$ have the same or opposite direction, according as $t > 0$ or $t < 0$.

**DEFINITION 8.4.3** If $A$ and $B$ are distinct points on a line $\mathcal{L}$, the non-zero vector $\overrightarrow{AB}$ is called a direction vector for $\mathcal{L}$.

It is easy to prove that any two direction vectors for a line are parallel.

**DEFINITION 8.4.4** Let $\mathcal{L}$ and $\mathcal{M}$ be lines having direction vectors $X$ and $Y$, respectively. Then $\mathcal{L}$ is parallel to $\mathcal{M}$ if $X$ is parallel to $Y$. Clearly any line is parallel to itself.

It is easy to prove that the line through a given point $A$ and parallel to a given line $CD$ has an equation $P = A + t \overrightarrow{CD}$.

**THEOREM 8.4.3** Let $X = a_1i + b_1j + c_1k$ and $Y = a_2i + b_2j + c_2k$ be non-zero vectors. Then $X$ is parallel to $Y$ if and only if

$$\left| \begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \end{array} \right| = \left| \begin{array}{cc} b_1 & c_1 \\ b_2 & c_2 \end{array} \right| = \left| \begin{array}{cc} a_1 & c_1 \\ a_2 & c_2 \end{array} \right| = 0. \tag{8.8}$$

Proof. The case of equality in the Cauchy–Schwarz inequality (theorem 8.3.1) shows that $X$ and $Y$ are parallel if and only if

$$|X \cdot Y| = ||X|| \cdot ||Y||.$$ 

Squaring gives the equivalent equality

$$(a_1a_2 + b_1b_2 + c_1c_2)^2 = (a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2),$$
which simplifies to
\[(a_1b_2 - a_2b_1)^2 + (b_1c_2 - b_2c_1)^2 + (a_1c_2 - a_2c_1)^2 = 0,\]
which is equivalent to
\[a_1b_2 - a_2b_1 = 0, b_1c_2 - b_2c_1 = 0, a_1c_2 - a_2c_1 = 0,\]
which is equation 8.8.

Equality of geometrical vectors has a fundamental geometrical interpretation:

**THEOREM 8.4.4** Suppose \(A, B, C, D\) are distinct points such that no three are collinear. Then \(\overrightarrow{AB} = \overrightarrow{CD}\) if and only if \(\overrightarrow{AB} \parallel \overrightarrow{CD}\) and \(\overrightarrow{AC} \parallel \overrightarrow{BD}\) (See Figure 8.1.)

**Proof.** If \(\overrightarrow{AB} = \overrightarrow{CD}\) then
\[
\begin{align*}
\overrightarrow{B - A} &= \overrightarrow{D - C}, \\
\overrightarrow{C - A} &= \overrightarrow{D - B}
\end{align*}
\]
and so \(\overrightarrow{AC} = \overrightarrow{BD}\). Hence \(\overrightarrow{AB} \parallel \overrightarrow{CD}\) and \(\overrightarrow{AC} \parallel \overrightarrow{BD}\).

Conversely, suppose that \(\overrightarrow{AB} \parallel \overrightarrow{CD}\) and \(\overrightarrow{AC} \parallel \overrightarrow{BD}\). Then
\[
\overrightarrow{AB} = s \overrightarrow{CD} \quad \text{and} \quad \overrightarrow{AC} = t \overrightarrow{BD},
\]
or
\[
\overrightarrow{B - A} = s(\overrightarrow{D - C}) \quad \text{and} \quad \overrightarrow{C - A} = t(\overrightarrow{D - B}).
\]
We have to prove \(s = 1\) or equivalently, \(t = 1\).

Now subtracting the second equation above from the first, gives
\[
\overrightarrow{B - C} = s(\overrightarrow{D - C}) - t(\overrightarrow{D - B}),
\]
so
\[
(1 - t)\overrightarrow{B} = (1 - s)\overrightarrow{C} + (s - t)\overrightarrow{D}.
\]
If \(t \neq 1\), then
\[
\overrightarrow{B} = \frac{1 - s}{1 - t} \overrightarrow{C} + \frac{s - t}{1 - t} \overrightarrow{D}
\]
and \(B\) would lie on the line \(CD\). Hence \(t = 1\).
8.5 The angle between two vectors

**DEFINITION 8.5.1** Let \( X \) and \( Y \) be non-zero vectors. Then the angle between \( X \) and \( Y \) is the unique value of \( \theta \) defined by

\[
\cos \theta = \frac{X \cdot Y}{\|X\| \cdot \|Y\|}, \quad 0 \leq \theta \leq \pi.
\]

**REMARK 8.5.1** By Cauchy’s inequality, we have

\[
-1 \leq \frac{X \cdot Y}{\|X\| \cdot \|Y\|} \leq 1,
\]

so the above equation does define an angle \( \theta \).

In terms of components, if \( X = [a_1, b_1, c_1]^t \) and \( Y = [a_2, b_2, c_2]^t \), then

\[
\cos \theta = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2} \sqrt{a_2^2 + b_2^2 + c_2^2}}. \tag{8.9}
\]

The next result is the well-known cosine rule for a triangle.

**THEOREM 8.5.1 (Cosine rule)** If \( A, B, C \) are points with \( A \neq B \) and \( A \neq C \), then the angle \( \theta \) between vectors \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \) satisfies

\[
\cos \theta = \frac{AB^2 + AC^2 - BC^2}{2AB \cdot AC}, \tag{8.10}
\]

or equivalently

\[
BC^2 = AB^2 + AC^2 - 2AB \cdot AC \cos \theta.
\]

(See Figure 8.10.)

**Proof.** Let \( A = (x_1, y_1, z_1) \), \( B = (x_2, y_2, z_2) \), \( C = (x_3, y_3, z_3) \). Then

\[
\overrightarrow{AB} = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}, \\
\overrightarrow{AC} = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}, \\
\overrightarrow{BC} = (a_2 - a_1)\mathbf{i} + (b_2 - b_1)\mathbf{j} + (c_2 - c_1)\mathbf{k},
\]

where

\[
a_i = x_{i+1} - x_1, \quad b_i = y_{i+1} - y_1, \quad c_i = z_{i+1} - z_1, \quad i = 1, 2.
\]
8.5. THE ANGLE BETWEEN TWO VECTORS

Now by equation 8.9,

\[
\cos \theta = \frac{a_1 a_2 + b_1 b_2 + c_1 c_2}{AB \cdot AC}.
\]

Also

\[
AB^2 + AC^2 - BC^2 = (a_1^2 + b_1^2 + c_1^2) + (a_2^2 + b_2^2 + c_2^2)
- ((a_2 - a_1)^2 + (b_2 - b_1)^2 + (c_2 - c_1)^2)
= 2a_1 a_2 + 2b_1 b_2 + c_1 c_2.
\]

Equation 8.10 now follows, since

\[
\overrightarrow{AB} \cdot \overrightarrow{AC} = a_1 a_2 + b_1 b_2 + c_1 c_2.
\]

**EXAMPLE 8.5.1** Let \( A = (2, 1, 0) \), \( B = (3, 2, 0) \), \( C = (5, 0, 1) \). Find the angle \( \theta \) between vectors \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \).

**Solution.**

\[
\cos \theta = \frac{\overrightarrow{AB} \cdot \overrightarrow{AC}}{AB \cdot AC}.
\]

Now

\[
\overrightarrow{AB} = \mathbf{i} + \mathbf{j} \quad \text{and} \quad \overrightarrow{AC} = 3\mathbf{i} - \mathbf{j} + \mathbf{k}.
\]
Figure 8.11: Pythagoras’ theorem for a right–angled triangle.

Hence

\[ \cos \theta = \frac{1 \times 3 + 1 \times (-1) + 0 \times 1}{\sqrt{1^2 + 1^2 + 0^2}\sqrt{3^2 + (-1)^2 + 1^2}} = \frac{2}{\sqrt{2}\sqrt{11}} = \frac{\sqrt{2}}{\sqrt{11}}. \]

Hence \( \theta = \cos^{-1} \frac{\sqrt{2}}{\sqrt{11}}. \)

**DEFINITION 8.5.2** If \( X \) and \( Y \) are vectors satisfying \( X \cdot Y = 0 \), we say \( X \) is orthogonal or perpendicular to \( Y \).

**REMARK 8.5.2** If \( A, B, C \) are points forming a triangle and \( \overrightarrow{AB} \) is orthogonal to \( \overrightarrow{AC} \), then the angle \( \theta \) between \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \) satisfies \( \cos \theta = 0 \) and hence \( \theta = \frac{\pi}{2} \) and the triangle is right–angled at \( A \).

Then we have Pythagoras’ theorem:

\[ BC^2 = AB^2 + AC^2. \quad (8.11) \]

We also note that \( BC \geq AB \) and \( BC \geq AC \) follow from equation 8.11. (See Figure 8.11.)

**EXAMPLE 8.5.2** Let \( A = (2, 9, 8), B = (6, 4, -2), C = (7, 15, 7) \). Show that \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \) are perpendicular and find the point \( D \) such that \( ABDC \) forms a rectangle.
8.5. THE ANGLE BETWEEN TWO VECTORS

Solution.

\[ \overrightarrow{AB} \cdot \overrightarrow{AC} = (4\mathbf{i} - 5\mathbf{j} - 10\mathbf{k}) \cdot (5\mathbf{i} + 6\mathbf{j} - \mathbf{k}) = 20 - 30 + 10 = 0. \]

Hence \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \) are perpendicular. Also, the required fourth point \( D \) clearly has to satisfy the equation

\[ \overrightarrow{BD} = -\overrightarrow{AC}, \] or equivalently \( \overrightarrow{D} - \overrightarrow{B} = -\overrightarrow{AC}. \)

Hence

\[ \overrightarrow{D} = \overrightarrow{B} + \overrightarrow{AC} = (6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}) + (5\mathbf{i} + 6\mathbf{j} - \mathbf{k}) = 11\mathbf{i} + 10\mathbf{j} - 3\mathbf{k}, \]

so \( D = (11, 10, -3). \)

**THEOREM 8.5.2 (Distance from a point to a line)** If \( C \) is a point and \( L \) is the line through \( A \) and \( B \), then there is exactly one point \( P \) on \( L \) such that \( \overrightarrow{CP} \) is perpendicular to \( \overrightarrow{AB} \), namely

\[ P = \overrightarrow{A} + t \overrightarrow{AB}, \quad t = \frac{\overrightarrow{AC} \cdot \overrightarrow{AB}}{\overrightarrow{AB}^2}. \quad (8.12) \]

Moreover if \( Q \) is any point on \( L \), then \( CQ \geq CP \) and hence \( P \) is the point on \( L \) closest to \( C \).
The shortest distance \( CP \) is given by
\[
CP = \sqrt{AC^2 AB^2 - (\vec{AC} \cdot \vec{AB})^2} / AB.
\] (8.13)

(See Figure 8.12.)

**Proof.** Let \( P = A + t \vec{AB} \) and assume that \( \vec{CP} \) is perpendicular to \( \vec{AB} \).

Then
\[
\begin{align*}
\vec{CP} \cdot \vec{AB} & = 0 \\
(P - C) \cdot \vec{AB} & = 0 \\
(A + t \vec{AB} - C) \cdot \vec{AB} & = 0 \\
(\vec{CA} + t \vec{AB}) \cdot \vec{AB} & = 0 \\
\vec{CA} \cdot \vec{AB} + t(\vec{AB} \cdot \vec{AB}) & = 0 \\
- \vec{AC} \cdot \vec{AB} + t(\vec{AB} \cdot \vec{AB}) & = 0,
\end{align*}
\]

so equation 8.12 follows.

The inequality \( CQ \geq CP \), where \( Q \) is any point on \( L \), is a consequence of Pythagoras’ theorem.

Finally, as \( \vec{CP} \) and \( \vec{PA} \) are perpendicular, Pythagoras’ theorem gives
\[
CP^2 = AC^2 - PA^2 = AC^2 - ||t \vec{AB}||^2 = AC^2 - t^2 AB^2 = AC^2 - \left( \frac{\vec{AC} \cdot \vec{AB}}{AB^2} \right)^2 AB^2 = AC^2 AB^2 - \left( \frac{\vec{AC} \cdot \vec{AB}}{AB^2} \right)^2 AB^2,
\]
as required.

**EXAMPLE 8.5.3** The closest point on the line through \( A = (1, 2, 1) \) and \( B = (2, -1, 3) \) to the origin is \( P = (\frac{17}{17}, \frac{19}{17}, \frac{20}{17}) \) and the corresponding shortest distance equals \( \frac{5}{17} \sqrt{42} \).

Another application of theorem 8.5.2 is to the projection of a line segment on another line:
THEOREM 8.5.3 (The projection of a line segment onto a line)
Let $C_1, C_2$ be points and $P_1, P_2$ be the feet of the perpendiculars from $C_1$ and $C_2$ to the line $AB$. Then

$$P_1P_2 = |\overrightarrow{C_1C_2} \cdot \hat{n}|,$$

where

$$\hat{n} = \frac{1}{AB} \overrightarrow{AB}.$$  

Also

$$C_1C_2 \geq P_1P_2. \quad (8.14)$$

(See Figure 8.13.)

**Proof.** Using equations 8.12, we have

$$\mathbf{P}_1 = \mathbf{A} + t_1 \overrightarrow{AB}, \quad \mathbf{P}_2 = \mathbf{A} + t_2 \overrightarrow{AB},$$

where

$$t_1 = \frac{\overrightarrow{AC}_1 \cdot \overrightarrow{AB}}{AB^2}, \quad t_2 = \frac{\overrightarrow{AC}_2 \cdot \overrightarrow{AB}}{AB^2}.$$  

Hence

$$\overrightarrow{P_1P_2} = (\mathbf{A} + t_2 \overrightarrow{AB}) - (\mathbf{A} + t_1 \overrightarrow{AB})$$

$$= (t_2 - t_1) \overrightarrow{AB},$$
CHAPTER 8. THREE–DIMENSIONAL GEOMETRY

so

\[ P_1 P_2 = ||P_1 P_2|| = |t_2 - t_1|AB \]
\[ = \left| \frac{\overrightarrow{AC_2} \cdot \overrightarrow{AB}}{AB^2} - \frac{\overrightarrow{AC_1} \cdot \overrightarrow{AB}}{AB^2} \right| AB \]
\[ = \left| \frac{\overrightarrow{C_1 C_2} \cdot \overrightarrow{AB}}{AB^2} \right| AB \]
\[ = \left| \overrightarrow{C_1 C_2} \cdot \hat{n} \right| , \]

where \( \hat{n} \) is the unit vector

\[ \hat{n} = \frac{1}{AB} \overrightarrow{AB} . \]

Inequality 8.14 then follows from the Cauchy–Schwarz inequality 8.3.

**DEFINITION 8.5.3** Two non–intersecting lines are called *skew* if they have non–parallel direction vectors.

Theorem 8.5.3 has an application to the problem of showing that two skew lines have a shortest distance between them. (The reader is referred to problem 16 at the end of the chapter.)

Before we turn to the study of planes, it is convenient to introduce the cross–product of two vectors.

### 8.6 The cross–product of two vectors

**DEFINITION 8.6.1** Let \( X = a_1 \mathbf{i} + b_1 \mathbf{j} + c_1 \mathbf{k} \) and \( Y = a_2 \mathbf{i} + b_2 \mathbf{j} + c_2 \mathbf{k} \). Then \( X \times Y \), the *cross–product* of \( X \) and \( Y \), is defined by

\[ X \times Y = ai + bj + ck , \]

where

\[ a = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} , \quad b = - \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} , \quad c = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} . \]

The vector cross–product has the following properties which follow from properties of \( 2 \times 2 \) and \( 3 \times 3 \) determinants:

(i) \( \mathbf{i} \times \mathbf{j} = \mathbf{k} , \quad \mathbf{j} \times \mathbf{k} = \mathbf{i} , \quad \mathbf{k} \times \mathbf{i} = \mathbf{j} ; \)
(ii) \( X \times X = 0 \);

(iii) \( Y \times X = -X \times Y \);

(iv) \( X \times (Y + Z) = X \times Y + X \times Z \);

(v) \( (tX) \times Y = t(X \times Y) \);

(vi) (Scalar triple product formula) if \( Z = a_3 i + b_3 j + c_3 k \), then

\[
X \cdot (Y \times Z) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = (X \times Y) \cdot Z;
\]

(vii) \( X \cdot (X \times Y) = 0 = Y \cdot (X \times Y) \);

(viii) \( ||X \times Y|| = \sqrt{||X||^2 ||Y||^2 - (X \cdot Y)^2} \);

(ix) if \( X \) and \( Y \) are non–zero vectors and \( \theta \) is the angle between \( X \) and \( Y \), then

\[
||X \times Y|| = ||X|| \cdot ||Y|| \sin \theta.
\]

(See Figure 8.14.)

From theorem 8.4.3 and the definition of cross–product, it follows that non–zero vectors \( X \) and \( Y \) are parallel if and only if \( X \times Y = 0 \); hence by (vii), the cross–product of two non–parallel, non–zero vectors \( X \) and \( Y \), is a non–zero vector perpendicular to both \( X \) and \( Y \).

**Lemma 8.6.1** Let \( X \) and \( Y \) be non–zero, non–parallel vectors.

(i) \( Z \) is a linear combination of \( X \) and \( Y \), if and only if \( Z \) is perpendicular to \( X \times Y \);

(ii) \( Z \) is perpendicular to \( X \) and \( Y \), if and only if \( Z \) is parallel to \( X \times Y \).

**Proof.** Let \( X \) and \( Y \) be non–zero, non–parallel vectors. Then

\[
X \times Y \neq 0.
\]

Then if \( X \times Y = a i + bj + ck \), we have

\[
\det [X \times Y | X | Y]^t = \begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = (X \times Y) \cdot (X \times Y) > 0.
\]
Hence the matrix \([X \times Y\,|\,X\,|\,Y]\) is non–singular. Consequently the linear system

\[
r(X \times Y) + sX + tY = Z
\]

has a unique solution \(r, s, t\).

(i) Suppose \(Z = sX + tY\). Then

\[
Z \cdot (X \times Y) = sX \cdot (X \times Y) + tY \cdot (X \times Y) = s0 + t0 = 0.
\]

Conversely, suppose that

\[
Z \cdot (X \times Y) = 0.
\]

Now from equation 8.15, \(r, s, t\) exist satisfying

\[
Z = r(X \times Y) + sX + tY.
\]

Then equation 8.16 gives

\[
0 = (r(X \times Y) + sX + tY) \cdot (X \times Y)
= r||X \times Y||^2 + sX \cdot (X \times Y) + tY \cdot (Y \times X)
= r||X \times Y||^2.
\]

Hence \(r = 0\) and \(Z = sX + tY\), as required.

(ii) Suppose \(Z = \lambda(X \times Y)\). Then clearly \(Z\) is perpendicular to \(X\) and \(Y\).
Conversely suppose that $Z$ is perpendicular to $X$ and $Y$.
Now from equation 8.15, $r, s, t$ exist satisfying
\[ Z = r(X \times Y) + sX + tY. \]
Then
\[ sX \cdot X + tX \cdot Y = X \cdot Z = 0 \]
\[ sY \cdot X + tY \cdot Y = Y \cdot Z = 0, \]
from which it follows that
\[ (sX + tY) \cdot (sX + tY) = 0. \]
Hence $sX + tY = 0$ and so $s = 0, t = 0$. Consequently $Z = r(X \times Y)$, as required.

The cross–product gives a compact formula for the distance from a point to a line, as well as the area of a triangle.

**THEOREM 8.6.1 (Area of a triangle)**
If $A, B, C$ are distinct non–collinear points, then

(i) the distance $d$ from $C$ to the line $AB$ is given by
\[ d = \frac{|| \overrightarrow{AB} \times \overrightarrow{AC} ||}{\overrightarrow{AB}}, \]  
(8.17)

(ii) the area of the triangle $ABC$ equals
\[ \frac{|| \overrightarrow{AB} \times \overrightarrow{AC} ||}{2} = \frac{|| \overrightarrow{A} \times \overrightarrow{B} + \overrightarrow{B} \times \overrightarrow{C} + \overrightarrow{C} \times \overrightarrow{A} ||}{2}. \]  
(8.18)

**Proof.** The area $\Delta$ of triangle $ABC$ is given by
\[ \Delta = \frac{AB \cdot CP}{2}, \]
where $P$ is the foot of the perpendicular from $C$ to the line $AB$. Now by formula 8.13, we have
\[ CP = \frac{\sqrt{AC^2 \cdot AB^2 - (\overrightarrow{AC} \cdot \overrightarrow{AB})^2}}{AB} \]
\[ = \frac{|| \overrightarrow{AB} \times \overrightarrow{AC} ||}{\overrightarrow{AB}}, \]
which, by property (viii) of the cross-product, gives formula 8.17. The second formula of equation 8.18 follows from the equations

\[
\overrightarrow{AB} \times \overrightarrow{AC} = (\overrightarrow{B} - \overrightarrow{A}) \times (\overrightarrow{C} - \overrightarrow{A})
\]

\[
= \{(\overrightarrow{B} - \overrightarrow{A}) \times \overrightarrow{C}\} - \{(\overrightarrow{C} - \overrightarrow{A}) \times \overrightarrow{A}\}
\]

\[
= (\overrightarrow{B} \times \overrightarrow{C} - \overrightarrow{A} \times \overrightarrow{C}) - (\overrightarrow{C} \times \overrightarrow{A} - \overrightarrow{A} \times \overrightarrow{A})
\]

\[
= B \times C - A \times C - B \times A
\]

\[
= B \times C + C \times A + A \times B,
\]

as required.

8.7 Planes

**DEFINITION 8.7.1** A plane is a set of points \( P \) satisfying an equation of the form

\[
P = P_0 + sX + tY, \quad s, t \in \mathbb{R},
\]

(8.19)

where \( X \) and \( Y \) are non-zero, non-parallel vectors.

For example, the \( xy \)-plane consists of the points \( P = (x, y, 0) \) and corresponds to the plane equation

\[
P = xi + yj = O + xi + yj.
\]

In terms of coordinates, equation 8.19 takes the form

\[
x = x_0 + sa_1 + ta_2
\]

\[
y = y_0 + sb_1 + tb_2
\]

\[
z = z_0 + sc_1 + tc_2,
\]

where \( P_0 = (x_0, y_0, z_0) \) and \( (a_1, b_1, c_1) \) and \( (a_2, b_2, c_2) \) are non-zero and non-proportional.

**THEOREM 8.7.1** Let \( A, B, C \) be three non-collinear points. Then there is one and only one plane through these points, namely the plane given by the equation

\[
P = A + s \overrightarrow{AB} + t \overrightarrow{AC},
\]

(8.20)

or equivalently

\[
\overrightarrow{AP} = s \overrightarrow{AB} + t \overrightarrow{AC}.
\]

(8.21)

(See Figure 8.15.)
Proof. First note that equation 8.20 is indeed the equation of a plane through $A$, $B$ and $C$, as $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are non-zero and non-parallel and $(s, t) = (0, 0), (1, 0)$ and $(0, 1)$ give $P = A$, $B$ and $C$, respectively. Call this plane $P$.

Conversely, suppose $P = P_0 + sX + tY$ is the equation of a plane $Q$ passing through $A$, $B$, $C$. Then $A = P_0 + s_0X + t_0Y$, so the equation for $Q$ may be written

$$P = A + (s - s_0)X + (t - t_0)Y = A + s'X + t'Y;$$

so in effect we can take $P_0 = A$ in the equation of $Q$. Then the fact that $B$ and $C$ lie on $Q$ gives equations

$$B = A + s_1X + t_1Y, \quad C = A + s_2X + t_2Y,$$

or

$$\overrightarrow{AB} = s_1X + t_1Y, \quad \overrightarrow{AC} = s_2X + t_2Y. \quad (8.22)$$

Then equations 8.22 and equation 8.20 show that $P \subseteq Q$.

Conversely, it is straightforward to show that because $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are not parallel, we have

$$\begin{vmatrix} s_1 & t_1 \\ s_2 & t_2 \end{vmatrix} \neq 0.$$
CHAPTER 8. THREE–DIMENSIONAL GEOMETRY

Figure 8.16: Normal equation of the plane $ABC$.

Hence equations 8.22 can be solved for $X$ and $Y$ as linear combinations of $\overrightarrow{AB}$ and $\overrightarrow{AC}$, allowing us to deduce that

$$Q \subseteq P.$$ 

Hence

$$Q = P.$$ 

THEOREM 8.7.2 (Normal equation for a plane) Let

$$A = (x_1, y_1, z_1), B = (x_2, y_2, z_2), C = (x_3, y_3, z_3)$$

be three non–collinear points. Then the plane through $A, B, C$ is given by

$$\overrightarrow{AP} \cdot (\overrightarrow{AB} \times \overrightarrow{AC}) = 0,$$ 

(8.23)

or equivalently,

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0,$$ 

(8.24)

where $P = (x, y, z)$. (See Figure 8.16.)
REMARK 8.7.1 Equation 8.24 can be written in more symmetrical form as

\[
\begin{vmatrix}
  x & y & z & 1 \\
  x_1 & y_1 & z_1 & 1 \\
  x_2 & y_2 & z_2 & 1 \\
  x_3 & y_3 & z_3 & 1 \\
\end{vmatrix} = 0.
\]

(8.25)

Proof. Let \( P \) be the plane through \( A, B, C \). Then by equation 8.21, we have \( P \in P \) if and only if \( \overrightarrow{AP} \) is a linear combination of \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \) and so by lemma 8.6.1(i), using the fact that \( \overrightarrow{AB} \times \overrightarrow{AC} \neq 0 \) here, if and only if \( \overrightarrow{AP} \) is perpendicular to \( \overrightarrow{AB} \times \overrightarrow{AC} \). This gives equation 8.23.

Equation 8.24 is the scalar triple product version of equation 8.23, taking into account the equations

\[
\begin{align*}
\overrightarrow{AP} &= (x - x_1)i + (y - y_1)j + (z - z_1)k, \\
\overrightarrow{AB} &= (x_2 - x_1)i + (y_2 - y_1)j + (z_2 - z_1)k, \\
\overrightarrow{AC} &= (x_3 - x_1)i + (y_3 - y_1)j + (z_3 - z_1)k.
\end{align*}
\]

REMARK 8.7.2 Equation 8.24 gives rise to a linear equation in \( x, y \) and \( z \):

\[
a x + b y + c z = d,
\]
where $a \mathbf{i} + b \mathbf{j} + c \mathbf{k} \neq 0$. For

\[
\begin{vmatrix}
  x - x_1 & y - y_1 & z - z_1 \\
  x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\
  x_3 - x_1 & y_3 - y_1 & z_3 - z_1
\end{vmatrix}
\]

and expanding the first determinant on the right-hand side of equation 8.26 along row 1 gives an expression

\[
ax + by + cz
\]

where

\[
a = \begin{vmatrix}
  y_2 - y_1 & z_2 - z_1 \\
  y_3 - y_1 & z_3 - z_1
\end{vmatrix},
\]

\[
b = -\begin{vmatrix}
  x_2 - x_1 & z_2 - z_1 \\
  x_3 - x_1 & z_3 - z_1
\end{vmatrix},
\]

\[
c = \begin{vmatrix}
  x_2 - x_1 & y_2 - y_1 \\
  x_3 - x_1 & y_3 - y_1
\end{vmatrix}.
\]

But $a$, $b$, $c$ are the components of $\mathbf{AB} \times \mathbf{AC}$, which in turn is non-zero, as $A$, $B$, $C$ are non-collinear here.

Conversely if $a \mathbf{i} + b \mathbf{j} + c \mathbf{k} \neq 0$, the equation

\[
a x + by + cz = d
\]

does indeed represent a plane. For if say $a \neq 0$, the equation can be solved for $x$ in terms of $y$ and $z$:

\[
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} = \begin{bmatrix}
  -\frac{d}{a} \\
  0 \\
  0
\end{bmatrix} + y \begin{bmatrix}
  -\frac{b}{a} \\
  1 \\
  0
\end{bmatrix} + z \begin{bmatrix}
  -\frac{c}{a} \\
  0 \\
  1
\end{bmatrix},
\]

which gives the plane

\[
\mathbf{P} = \mathbf{P}_0 + y \mathbf{X} + z \mathbf{Y},
\]

where $\mathbf{P}_0 = (-\frac{d}{a}, 0, 0)$ and $\mathbf{X} = -\frac{b}{a} \mathbf{i} + \mathbf{j}$ and $\mathbf{Y} = -\frac{c}{a} \mathbf{i} + \mathbf{k}$ are evidently non-parallel vectors.

**Remark 8.7.3** The plane equation $a x + by + cz = d$ is called the normal form, as it is easy to prove that if $P_1$ and $P_2$ are two points in the plane, then $a \mathbf{i} + b \mathbf{j} + c \mathbf{k}$ is perpendicular to $\overrightarrow{P_1P_2}$. Any non-zero vector with this property is called a normal to the plane. (See Figure 8.17.)
By lemma 8.6.1(ii), it follows that every vector \( X \) normal to a plane through three non–collinear points \( A, B, C \) is parallel to \( \overrightarrow{AB} \times \overrightarrow{AC} \), since \( X \) is perpendicular to \( \overrightarrow{AB} \) and \( \overrightarrow{AC} \).

**EXAMPLE 8.7.1** Show that the planes

\[
x + y - 2z = 1 \quad \text{and} \quad x + 3y - z = 4
\]

intersect in a line and find the distance from the point \( C = (1, 0, 1) \) to this line.

**Solution.** Solving the two equations simultaneously gives

\[
x = -\frac{1}{2} + \frac{5}{2}z, \quad y = \frac{3}{2} - \frac{1}{2}z, \quad (8.27)
\]

where \( z \) is arbitrary. Hence

\[
xi + yj + zk = -\frac{1}{2}i - \frac{3}{2}j + z(\frac{5}{2}i - \frac{1}{2}j + k),
\]

which is the equation of a line \( L \) through \( A = (-\frac{1}{2}, -\frac{3}{2}, 0) \) and having direction vector \( \frac{5}{2}i - \frac{1}{2}j + k \).

We can now proceed in one of three ways to find the closest point on \( L \) to \( A \).

One way is to use equation 8.17 with \( B \) defined by

\[
\overrightarrow{AB} = \frac{5}{2}i - \frac{1}{2}j + k.
\]

Another method minimizes the distance \( CP \), where \( P \) ranges over \( L \).

A third way is to find an equation for the plane through \( C \), having \( \frac{5}{2}i - \frac{1}{2}j + k \) as a normal. Such a plane has equation

\[
5x - y + 2z = d,
\]

where \( d \) is found by substituting the coordinates of \( C \) in the last equation.

\[
d = 5 \times 1 - 0 + 2 \times 1 = 7.
\]

We now find the point \( P \) where the plane intersects the line \( L \). Then \( \overrightarrow{CP} \) will be perpendicular to \( L \) and \( CP \) will be the required shortest distance from \( C \) to \( L \). We find using equations 8.27 that

\[
5(-\frac{1}{2} + \frac{5}{2}z) - (\frac{3}{2} - \frac{1}{2}z) + 2z = 7,
\]
so \( z = \frac{11}{15} \). Hence \( P = \left( \frac{4}{3}, \frac{17}{15}, \frac{11}{15} \right) \).

It is clear that through a given line and a point not on that line, there passes exactly one plane. If the line is given as the intersection of two planes, each in normal form, there is a simple way of finding an equation for this plane. More explicitly we have the following result:

**THEOREM 8.7.3** Suppose the planes

\[
\begin{align*}
  a_1x + b_1y + c_1z &= d_1 \\  a_2x + b_2y + c_2z &= d_2
\end{align*}
\]

have non–parallel normals. Then the planes intersect in a line \( \mathcal{L} \).

Moreover the equation

\[
\lambda(a_1x + b_1y + c_1z - d_1) + \mu(a_2x + b_2y + c_2z - d_2) = 0,
\]

where \( \lambda \) and \( \mu \) are not both zero, gives all planes through \( \mathcal{L} \).

(See Figure 8.18.)

**Proof.** Assume that the normals \( a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k} \) and \( a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k} \) are non–parallel. Then by theorem 8.4.3, not all of

\[
\Delta_1 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}
\]

(8.31)
8.7. PLANES

are zero. If say $\Delta_1 \neq 0$, we can solve equations 8.28 and 8.29 for $x$ and $y$ in terms of $z$, as we did in the previous example, to show that the intersection forms a line $L$.

We next have to check that if $\lambda$ and $\mu$ are not both zero, then equation 8.30 represents a plane. (Whatever set of points equation 8.30 represents, this set certainly contains $L$.)

$$(\lambda a_1 + \mu a_2)x + (\lambda b_1 + \mu b_2)y + (\lambda c_1 + \mu c_2)z - (\lambda d_1 + \mu d_2) = 0.$$ Then we clearly cannot have all the coefficients $\lambda a_1 + \mu a_2$, $\lambda b_1 + \mu b_2$, $\lambda c_1 + \mu c_2$ zero, as otherwise the vectors $a_1 \mathbf{i} + b_1 \mathbf{j} + c_1 \mathbf{k}$ and $a_2 \mathbf{i} + b_2 \mathbf{j} + c_2 \mathbf{k}$ would be parallel.

Finally, if $P$ is a plane containing $L$, let $P_0 = (x_0, y_0, z_0)$ be a point not on $L$. Then if we define $\lambda$ and $\mu$ by

$$\lambda = -(a_2 x_0 + b_2 y_0 + c_2 z_0 - d_2), \quad \mu = a_1 x_0 + b_1 y_0 + c_1 z_0 - d_1,$$

then at least one of $\lambda$ and $\mu$ is non–zero. Then the coordinates of $P_0$ satisfy equation 8.30, which therefore represents a plane passing through $L$ and $P_0$ and hence identical with $P$.

**EXAMPLE 8.7.2** Find an equation for the plane through $P_0 = (1, 0, 1)$ and passing through the line of intersection of the planes

$$x + y - 2z = 1 \quad \text{and} \quad x + 3y - z = 4.$$ Solution. The required plane has the form

$$\lambda(x + y - 2z - 1) + \mu(x + 3y - z - 4) = 0,$$

where not both of $\lambda$ and $\mu$ are zero. Substituting the coordinates of $P_0$ into this equation gives

$$-2\lambda + \mu(-4) = 0, \quad \lambda = -2\mu.$$ So the required equation is

$$-2\mu(x + y - 2z - 1) + \mu(x + 3y - z - 4) = 0,$$

or

$$-x + y + 3z - 2 = 0.$$ Our final result is a formula for the distance from a point to a plane.
THEOREM 8.7.4 (Distance from a point to a plane)

Let \( P_0 = (x_0, y_0, z_0) \) and \( \mathcal{P} \) be the plane

\[
ax + by + cz = d. \tag{8.32}
\]

Then there is a unique point \( P \) on \( \mathcal{P} \) such that \( \overrightarrow{P_0P} \) is normal to \( \mathcal{P} \). Moreover

\[
\overrightarrow{P_0P} = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.
\]

(See Figure 8.19.)

**Proof.** The line through \( P_0 \) normal to \( \mathcal{P} \) is given by

\[
P = P_0 + t(ai + bj + ck),
\]

or in terms of coordinates

\[
x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct.
\]

Substituting these formulae in equation 8.32 gives

\[
a(x_0 + at) + b(y_0 + bt) + c(z_0 + ct) = d
\]

\[
t(a^2 + b^2 + c^2) = -(ax_0 + by_0 + cz_0 - d),
\]

so

\[
t = - \left( \frac{ax_0 + by_0 + cz_0 - d}{a^2 + b^2 + c^2} \right).
\]
Then

\[ P_0P = || \overrightarrow{P_0P} || = || t(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) || = |t|\sqrt{a^2 + b^2 + c^2} = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \]

Other interesting geometrical facts about lines and planes are left to the problems at the end of this chapter.

### 8.8 PROBLEMS

1. Find the point where the line through \( A = (3, -2, 7) \) and \( B = (13, 3, -8) \) meets the \( xz \)-plane.
   
   [Ans: (7, 0, 1).]

2. Let \( A, B, C \) be non-collinear points. If \( E \) is the mid-point of the segment \( BC \) and \( F \) is the point on the segment \( EA \) satisfying \( \frac{AF}{EF} = 2 \), prove that
   
   \[ F = \frac{1}{3}(A + B + C). \]
   
   (\( F \) is called the centroid of triangle \( ABC \).)

3. Prove that the points \( (2, 1, 4), (1, -1, 2), (3, 3, 6) \) are collinear.

4. If \( A = (2, 3, -1) \) and \( B = (3, 7, 4) \), find the points \( P \) on the line \( AB \) satisfying \( PA/PB = 2/5 \).
   
   [Ans: \((\frac{16}{7}, \frac{29}{7}, \frac{3}{7})\) and \((\frac{4}{5}, \frac{1}{5}, -\frac{13}{5})\).]

5. Let \( M \) be the line through \( A = (1, 2, 3) \) parallel to the line joining \( B = (-2, 2, 0) \) and \( C = (4, -1, 7) \). Also \( N \) is the line joining \( E = (1, -1, 8) \) and \( F = (10, -1, 11) \). Prove that \( M \) and \( N \) intersect and find the point of intersection.
   
   [Ans: \((7, -1, 10)\).]
6. Prove that the triangle formed by the points \((-3, 5, 6), (-2, 7, 9)\) and \((2, 1, 7)\) is a \(30^\circ, 60^\circ, 90^\circ\) triangle.

7. Find the point on the line \(AB\) closest to the origin, where \(A = (-2, 1, 3)\) and \(B = (1, 2, 4)\). Also find this shortest distance.

   \[\text{Ans: } (-\frac{16}{11}, \frac{13}{11}, \frac{35}{11}) \text{ and } \sqrt{\frac{150}{11}}.\]

8. A line \(N\) is determined by the two planes \(x + y - 2z = 1\), and \(x + 3y - z = 4\).

   Find the point \(P\) on \(N\) closest to the point \(C = (1, 0, 1)\) and find the distance \(PC\).

   \[\text{Ans: } \left(\frac{4}{3}, \frac{17}{15}, \frac{11}{15}\right) \text{ and } \sqrt{\frac{330}{15}}.\]

9. Find a linear equation describing the plane perpendicular to the line of intersection of the planes \(x + y - 2z = 4\) and \(3x - 2y + z = 1\) and which passes through \((6, 0, 2)\).

   \[\text{Ans: } 3x + 7y + 5z = 28.\]

10. Find the length of the projection of the segment \(AB\) on the line \(L\), where \(A = (1, 2, 3)\), \(B = (5, -2, 6)\) and \(L\) is the line \(CD\), where \(C = (7, 1, 9)\) and \(D = (-1, 5, 8)\).

    \[\text{Ans: } \frac{17}{3}.\]

11. Find a linear equation for the plane through \(A = (3, -1, 2)\), perpendicular to the line \(L\) joining \(B = (2, 1, 4)\) and \(C = (-3, -1, 7)\). Also find the point of intersection of \(L\) and the plane and hence determine the distance from \(A\) to \(L\).

    \[\text{Ans: } 5x + 2y - 3z = 7, \left(\frac{111}{13}, \frac{52}{13}, \frac{131}{13}\right), \sqrt{\frac{293}{13}}.\]

12. If \(P\) is a point inside the triangle \(ABC\), prove that

    \[P = rA + sB + tC,\]

    where \(r + s + t = 1\) and \(r > 0, s > 0, t > 0\).

13. If \(B\) is the point where the perpendicular from \(A = (6, -1, 11)\) meets the plane \(3x + 4y + 5z = 10\), find \(B\) and the distance \(AB\).

    \[\text{Ans: } B = \left(\frac{123}{50}, \frac{-286}{50}, \frac{255}{50}\right) \text{ and } AB = \frac{59}{\sqrt{50}}.\]
14. Prove that the triangle with vertices \((-3, 0, 2), (6, 1, 4), (-5, 1, 0)\) has area \(\frac{1}{2}\sqrt{333}\).

15. Find an equation for the plane through \((2, 1, 4), (1, -1, 2), (4, -1, 1)\).
   [Ans: \(2x - 7y + 6z = 21\).]

16. Lines \(\mathcal{L}\) and \(\mathcal{M}\) are non–parallel in 3–dimensional space and are given by equations
   \[
   \mathbf{P} = \mathbf{A} + s\mathbf{X}, \quad \mathbf{Q} = \mathbf{B} + t\mathbf{Y}.
   \]
   (i) Prove that there is precisely one pair of points \(P\) and \(Q\) such that \(\overrightarrow{PQ}\) is perpendicular to \(\mathbf{X}\) and \(\mathbf{Y}\).
   
   (ii) Explain why \(PQ\) is the shortest distance between lines \(\mathcal{L}\) and \(\mathcal{M}\).
   Also prove that
   \[
   PQ = \frac{|(\mathbf{X} \times \mathbf{Y}) \cdot \mathbf{AB}|}{\|\mathbf{X} \times \mathbf{Y}\|}.
   \]

17. If \(\mathcal{L}\) is the line through \(A = (1, 2, 1)\) and \(C = (3, -1, 2)\), while \(\mathcal{M}\)
    is the line through \(B = (1, 0, 2)\) and \(D = (2, 1, 3)\), prove that the shortest distance
    between \(\mathcal{L}\) and \(\mathcal{M}\) equals \(\frac{13}{\sqrt{62}}\).

18. Prove that the volume of the tetrahedron formed by four non–coplanar points \(A_i = (x_i, y_i, z_i), 1 \leq i \leq 4,\) is equal to
   \[
   \frac{1}{6} |(\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3}) \cdot \overrightarrow{A_1A_4}|,
   \]
   which in turn equals the absolute value of the determinant
   \[
   \begin{vmatrix}
   1 & x_1 & y_1 & z_1 \\
   1 & x_2 & y_2 & z_2 \\
   1 & x_3 & y_3 & z_3 \\
   1 & x_4 & y_4 & z_4
   \end{vmatrix}.
   \]

19. The points \(A = (1, 1, 5), B = (2, 2, 1), C = (1, -2, 2)\) and \(D = (-2, 1, 2)\) are the vertices of a tetrahedron. Find the equation of the line through \(A\) perpendicular to the face \(BCD\) and the distance of \(A\) from this face. Also find the shortest distance between the skew lines \(AD\) and \(BC\).
   [Ans: \(\mathbf{P} = (1 + t)(\mathbf{i} + \mathbf{j} + 5\mathbf{k}); 2\sqrt{3}; 3.\)]
Chapter 9

FURTHER READING

Matrix theory has many applications to science, mathematics, economics and engineering. Some of these applications can be found in the books [2, 3, 4, 5, 11, 13, 16, 20, 26, 28].

For the numerical side of matrix theory, [6] is recommended. Its bibliography is also useful as a source of further references.

For applications to:

1. Graph theory, see [7, 13];
2. Coding theory, see [8, 15];
3. Game theory, see [13];
4. Statistics, see [9];
5. Economics, see [10];
6. Biological systems, see [12];
7. Markov non–negative matrices, see [11, 13, 14, 17];
8. The general equation of the second degree in three variables, see [18];
9. Affine and projective geometry, see [19, 21, 22];
10. Computer graphics, see [23, 24].
Bibliography


<table>
<thead>
<tr>
<th>Adjacent</th>
<th>Cosine rule 172</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>De Moivre’s theorem 110</td>
</tr>
<tr>
<td>Between</td>
<td>Dependent unknowns 11</td>
</tr>
<tr>
<td>Vectors</td>
<td>Determinant 74</td>
</tr>
<tr>
<td>Apollonius’</td>
<td>2 × 2 matrix 73</td>
</tr>
<tr>
<td>Circle</td>
<td>Lower triangular 76</td>
</tr>
<tr>
<td>Argand</td>
<td>Scalar matrix 76</td>
</tr>
<tr>
<td>Diagram</td>
<td>Laplace expansion 75</td>
</tr>
<tr>
<td>Asymptotes</td>
<td>Diagonal matrix 76</td>
</tr>
<tr>
<td>Augmented</td>
<td>Diagonal matrix 50</td>
</tr>
<tr>
<td>Matrix</td>
<td>Differential equations 125</td>
</tr>
<tr>
<td>Basis</td>
<td>Direction vector 170</td>
</tr>
<tr>
<td>Left–to–right 64</td>
<td>Distance to a plane 190</td>
</tr>
<tr>
<td>Cauchy-Schwarz inequality 165</td>
<td>Dot product 137, 162, 163</td>
</tr>
<tr>
<td>Centroid</td>
<td>Eigenvalue 122</td>
</tr>
<tr>
<td>Characteristic equation 122</td>
<td>Eigenvector 122</td>
</tr>
<tr>
<td>Coefficient matrix 2, 27</td>
<td>Elementary row matrix 42</td>
</tr>
<tr>
<td>Cofactor 78</td>
<td>Elementary row operations 7</td>
</tr>
<tr>
<td>Column vector 27</td>
<td>Ellipse 143, 144</td>
</tr>
<tr>
<td>Complex number 91</td>
<td>Factor theorem 97</td>
</tr>
<tr>
<td>Imaginary part 91</td>
<td>Field 3</td>
</tr>
<tr>
<td>Real part 91</td>
<td>Additive inverse 4</td>
</tr>
<tr>
<td>Real 91</td>
<td>Multiplicative inverse 4</td>
</tr>
<tr>
<td>Imaginary 92</td>
<td>Gauss’ theorem 97</td>
</tr>
<tr>
<td>Rationalization 93</td>
<td>Gauss-Jordan algorithm 8</td>
</tr>
<tr>
<td>Complex exponential 110</td>
<td>Gram matrix 138</td>
</tr>
<tr>
<td>Modulus 101</td>
<td>Homogeneous system 16</td>
</tr>
<tr>
<td>Square root 94</td>
<td>Hyperbola 143</td>
</tr>
<tr>
<td>Argument 106</td>
<td>Identity matrix 31</td>
</tr>
<tr>
<td>Complex conjugate 99</td>
<td>Imaginary matrix 31</td>
</tr>
<tr>
<td>Modulus-argument form 106</td>
<td>Inconsistent system 1</td>
</tr>
<tr>
<td>Polar representation 106</td>
<td>Independent unknowns 12</td>
</tr>
<tr>
<td>Ratio formulae 103</td>
<td>Cramer’s rule 40, 86</td>
</tr>
</tbody>
</table>
inversion 76
invertible matrix 37
Joachimsthal 168
least squares
  normal equations 48
  residuals 48
linear combination 17
linear equation 1
linear transformation 27
  reflection 29
  rotation 28
  projection 30
linear independence 41, 60
  left–to–right test
lower half plane 98
Markov matrix 54
mathematical induction 32
matrix 23
  addition 23
  additive inverse 24
  equality 23
  inverse 37
  orthogonal 136
  power 32
  product 25
  proper orthogonal 136
  rank 68
  scalar multiple 24
  singular 37
  skew–symmetric 47
  symmetric 47
  subtraction 24
  transpose 46
minor 74
modular addition 4
  multiplication 4
non–singular matrix 37, 50
non–trivial solution 17
orthogonal
  matrix 120, 136
  vectors 174
parabola 143
parallel lines 170
  vectors 170
parallelogram law 156
perpendicular vectors 174
plane 182
  normal form 187
  through 3 points 183, 184
position vector 161
positive octant 160
real axis 97
recurrence relations 33, 35
row–echelon form 6
  reduced 6
row–equivalence 7
rotation equations 28, 120
scalar triple product 179
skew lines 178
subspace 57
  basis 63
  column space 58
  dimension 65
  generated 58
  null space 57
  row space 58
Surveyor’s formula 87
system of equations 1
Three–dimensional space 160
  coordinate axes 160
  coordinate planes 160
  distance 161
  line equation 167
  projection on a line 177
triangle inequality 166
trivial solution 17
unit vectors 28, 164
upper half plane 98
vector
  cross-product 159, 179
equality 156, 171
length 137, 157, 164
scalar multiplication 156, 168
of unknowns 27
zero matrix 24
## CONTENTS

<table>
<thead>
<tr>
<th>Problems</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
<td>1</td>
</tr>
<tr>
<td>2.4</td>
<td>12</td>
</tr>
<tr>
<td>2.7</td>
<td>18</td>
</tr>
<tr>
<td>3.6</td>
<td>32</td>
</tr>
<tr>
<td>4.1</td>
<td>45</td>
</tr>
<tr>
<td>5.8</td>
<td>58</td>
</tr>
<tr>
<td>6.3</td>
<td>69</td>
</tr>
<tr>
<td>7.3</td>
<td>83</td>
</tr>
<tr>
<td>8.8</td>
<td>91</td>
</tr>
</tbody>
</table>
SECTION 1.6

2. (i) \[ \begin{bmatrix} 0 & 0 & 0 \\ 2 & 4 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \frac{1}{2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \]

(ii) \[ \begin{bmatrix} 0 & 1 & 3 \\ 1 & 2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \end{bmatrix}; \]

(iii) \[ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \]

R1 \rightarrow R1 + R3
R3 \rightarrow -R3
R2 \rightarrow R3

R2 \rightarrow R2 - R1
R3 \rightarrow R3 - R1

R1 \rightarrow R1 + R3
R3 \rightarrow R3 + R1

(iv) \[ \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ -4 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \]

R3 \rightarrow R3 + 2R1
R1 \rightarrow \frac{1}{2} R1

R2 \rightarrow R2 - 2R1
R3 \rightarrow R3 - R1

R1 \rightarrow R1 - R2
R3 \rightarrow R3 + 2R2

R1 \rightarrow R1 - 4R3
R2 \rightarrow R2 + 3R3

R1 \rightarrow R1 - 4R3
R2 \rightarrow R2 + 3R3

R1 \rightarrow R1 - 4R3
R2 \rightarrow R2 + 3R3

R3 \rightarrow R3 + 2R2

The augmented matrix has been converted to reduced row–echelon form and we read off the unique solution \( x = -3 \), \( y = \frac{19}{4} \), \( z = \frac{4}{3} \).

3. (a) \[ \begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & -1 & 2 \\ 1 & -1 & -1 & -8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -3 & 4 \\ 0 & -8 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

R1 \rightarrow R1 - R2
R3 \rightarrow R3 + \frac{1}{8} R3

R1 \rightarrow R1 - 4R3
R2 \rightarrow R2 + 3R3

R1 \rightarrow R1 - 4R3
R2 \rightarrow R2 + 3R3

R3 \rightarrow R3 + 2R2

From the last matrix we see that the original system is inconsistent.
\[ \begin{bmatrix} 3 & -1 & 7 & 0 \\ 2 & -1 & 4 & \frac{1}{2} \\ 1 & -1 & 1 & 1 \\ 6 & -4 & 10 & 3 \end{bmatrix} R_1 \leftrightarrow R_3 \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & -1 & 4 & \frac{1}{2} \\ 3 & -1 & 7 & 0 \\ 6 & -4 & 10 & 3 \end{bmatrix} \]

\[
R_2 \rightarrow R_2 - 2R_1 \begin{bmatrix} 1 & -1 & 1 & \frac{1}{2} \\ 0 & 1 & 2 & -\frac{3}{2} \\ 0 & 2 & 4 & -3 \\ 0 & 2 & 4 & -3 \end{bmatrix},
R_3 \rightarrow R_3 - 3R_1 \begin{bmatrix} 1 & 0 & 3 & \frac{1}{3} \\ 0 & 1 & 2 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
R_4 \rightarrow R_4 - 6R_1 \begin{bmatrix} 0 & 1 & -8 & b - a \\ 0 & 0 & 0 & 3b - 2a + c \end{bmatrix}.
\]

The augmented matrix has been converted to reduced row–echelon form and we read off the complete solution $x = -\frac{1}{2} - 3z$, $y = -\frac{3}{2} - 2z$, with $z$ arbitrary.

4. 
\[
\begin{bmatrix} 2 & -1 & 3 & a \\ 3 & 1 & -5 & b \\ -5 & -5 & 21 & c \end{bmatrix} R_2 \rightarrow R_2 - R_1 \begin{bmatrix} 2 & -1 & 3 & a \\ 1 & 2 & -8 & b - a \\ -5 & -5 & 21 & c \end{bmatrix} \]

\[
R_1 \leftrightarrow R_2 \begin{bmatrix} 1 & 2 & -8 & b - a \\ 2 & -1 & 3 & a \\ -5 & -5 & 21 & c \end{bmatrix} R_2 \rightarrow R_2 - 2R_1 \begin{bmatrix} 1 & 2 & -8 & b - a \\ 0 & 1 & -19 & -2b + 3a \\ 0 & 0 & 0 & 3b - 2a + c \end{bmatrix} \]

\[
R_3 \rightarrow R_3 + 5R_1 \begin{bmatrix} 1 & 2 & -8 & b - a \\ 0 & 1 & -19 & -2b + 3a \\ 0 & 0 & 0 & 3b - 2a + c \end{bmatrix} \]

From the last matrix we see that the original system is inconsistent if $3b - 2a + c \neq 0$. If $3b - 2a + c = 0$, the system is consistent and the solution is

\[
x = \frac{(b + a)}{5} + \frac{2}{5}z, \quad y = \frac{(2b - 3a)}{5} + \frac{19}{5}z,
\]

where $z$ is arbitrary.

5. 
\[
\begin{bmatrix} 1 & 1 & 1 \\ t & 1 & t \\ 1 + t & 2 & 3 \end{bmatrix} R_2 \rightarrow R_2 - tR_1 \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -t \\ 0 & 1 & -t \end{bmatrix} R_3 \rightarrow R_3 - (1 + t)R_1 \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -t \\ 0 & 0 & 2 - t \end{bmatrix} R_3 \rightarrow R_3 - R_2 \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -t \\ 0 & 0 & 2 - t \end{bmatrix} = B.
\]

Case 1. $t \neq 2$. No solution.
Case 2. \( t = 2 \). \( B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \).

We read off the unique solution \( x = 1, y = 0 \).

6. **Method 1.**

\[
\begin{bmatrix}
-3 & 1 & 1 & 1 \\
1 & -3 & 1 & 1 \\
1 & 1 & -3 & 1 \\
1 & 1 & 1 & -3
\end{bmatrix}
\begin{align*}
R_1 & \rightarrow R_1 - R_4 \\
R_2 & \rightarrow R_2 - R_4 \\
R_3 & \rightarrow R_3 - R_4
\end{align*}
\]

\[
\begin{bmatrix}
1 & 1 & 1 & -3
\end{bmatrix}
\begin{align*}
R_4 & \rightarrow R_4 - R_3 - R_2 - R_1 \\
R_1 & \rightarrow R_1 - (\lambda - 3)R_2
\end{align*}
\]

Hence the given homogeneous system has complete solution

\[ x_1 = x_4, \ x_2 = x_4, \ x_3 = x_4, \]

with \( x_4 \) arbitrary.

**Method 2.** Write the system as

\[
\begin{align*}
x_1 + x_2 + x_3 + x_4 &= 4x_1 \\
x_1 + x_2 + x_3 + x_4 &= 4x_2 \\
x_1 + x_2 + x_3 + x_4 &= 4x_3 \\
x_1 + x_2 + x_3 + x_4 &= 4x_4.
\end{align*}
\]

Then it is immediate that any solution must satisfy \( x_1 = x_2 = x_3 = x_4 \).
Conversely, if \( x_1, x_2, x_3, x_4 \) satisfy \( x_1 = x_2 = x_3 = x_4 \), we get a solution.

7. \[
\begin{bmatrix}
\lambda - 3 & 1 \\
1 & \lambda - 3
\end{bmatrix}
\begin{align*}
R_1 & \leftrightarrow R_2 \\
R_2 & \rightarrow R_2 - (\lambda - 3)R_1
\end{align*}
\]

\[
\begin{bmatrix}
\lambda - 3 & 1 \\
1 & \lambda - 3
\end{bmatrix}
\begin{align*}
R_2 & \rightarrow R_2 - (\lambda - 3)R_1
\end{align*}
\]

Case 1: \( -\lambda^2 + 6\lambda - 8 \neq 0 \). That is \( - (\lambda - 2)(\lambda - 4) \neq 0 \) or \( \lambda \neq 2, 4 \). Here \( B \) is row equivalent to \( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \):

\[
\begin{bmatrix}
1 & \lambda - 3 \\
0 & 1
\end{bmatrix}
\begin{align*}
R_1 & \rightarrow R_1 - (\lambda - 3)R_2 \\
R_2 & \rightarrow \frac{1}{-\lambda^2 + 6\lambda - 8}R_2
\end{align*}
\]

Hence we get the trivial solution \( x = 0, y = 0 \).
Case 2: \( \lambda = 2 \). Then \( B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \) and the solution is \( x = y \), with \( y \) arbitrary.

Case 3: \( \lambda = 4 \). Then \( B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \) and the solution is \( x = -y \), with \( y \) arbitrary.

8. 

\[
\begin{bmatrix} 3 & 1 & 1 & 1 \\ 5 & -1 & 1 & -1 \end{bmatrix} R_1 \rightarrow \frac{1}{5} R_1 \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \\
R_2 \rightarrow R_2 - 5R_1 \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \\
R_2 \rightarrow \frac{-3}{8} R_2 \begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{5} \\ 0 & 1 & \frac{1}{4} \end{bmatrix} \\
R_1 \rightarrow R_1 - \frac{1}{3} R_2 \begin{bmatrix} 1 & 0 \frac{1}{5} & 0 \end{bmatrix}.
\]

Hence the solution of the associated homogeneous system is

\[ x_1 = -\frac{1}{4} x_3, \ x_2 = -\frac{1}{4} x_3 - x_4, \]

with \( x_3 \) and \( x_4 \) arbitrary.

9. 

\[
A = \begin{bmatrix} 1-n & 1 & \cdots & 1 \\ 1 & 1-n & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-n \end{bmatrix}
\]

\[
R_1 \rightarrow R_1 - R_n \begin{bmatrix} -n & 0 & \cdots & n \\ 0 & -n & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-n \end{bmatrix}
\]

\[
R_2 \rightarrow R_2 - R_n \begin{bmatrix} -n & 0 & \cdots & n \\ 0 & -n & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-n \end{bmatrix}
\]

\[
R_{n-1} \rightarrow R_{n-1} - R_n \begin{bmatrix} -n & 0 & \cdots & n \\ 0 & -n & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1-n \end{bmatrix}
\]

The last matrix is in reduced row–echelon form.

Consequently the homogeneous system with coefficient matrix \( A \) has the solution

\[ x_1 = x_n, \ x_2 = x_n, \ldots, x_{n-1} = x_n, \]

4
with \( x_n \) arbitrary.

Alternatively, writing the system in the form

\[
\begin{align*}
    x_1 + \cdots + x_n &= nx_1 \\
    x_1 + \cdots + x_n &= nx_2 \\
    &\vdots \\
    x_1 + \cdots + x_n &= nx_n
\end{align*}
\]

shows that any solution must satisfy \( nx_1 = nx_2 = \cdots = nx_n \), so \( x_1 = x_2 = \cdots = x_n \). Conversely if \( x_1 = x_n, \ldots, x_{n-1} = x_n \), we see that \( x_1, \ldots, x_n \) is a solution.

10. Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) and assume that \( ad - bc \neq 0 \).

Case 1: \( a \neq 0 \).

\[
\begin{align*}
    \begin{bmatrix} a & b \\ c & d \end{bmatrix} R_1 &\rightarrow \frac{1}{a} R_1 \\
    R_2 &\rightarrow R_2 - cR_1
\end{align*}
\]

\[
\begin{align*}
    R_2 &\rightarrow \frac{a}{ad - bc} R_2 \\
    R_1 &\rightarrow R_1 - \frac{b}{a} R_2
\end{align*}
\]

Case 2: \( a = 0 \). Then \( bc \neq 0 \) and hence \( c \neq 0 \).

\[
\begin{align*}
    A = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} R_1 &\leftrightarrow R_2 \\
    R_2 &\rightarrow R_2 - cR_1 \\
    R_1 &\rightarrow R_1 - \frac{d}{c} R_2
\end{align*}
\]

So in both cases, \( A \) has reduced row–echelon form equal to \( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \).

11. We simplify the augmented matrix of the system using row operations:

\[
\begin{align*}
    \begin{bmatrix} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & a^2 - 14 & a + 2 \end{bmatrix} R_2 &\rightarrow R_2 - 3R_1 \\
    R_3 &\rightarrow R_3 - 4R_1
\end{align*}
\]

\[
\begin{align*}
    R_3 &\rightarrow R_3 - R_2 \\
    R_2 &\rightarrow \frac{1}{7} R_2 \\
    R_1 &\rightarrow R_1 - 2R_2
\end{align*}
\]

Denote the last matrix by \( B \).
Case 1: \(a^2 - 16 \neq 0\). i.e. \(a \neq \pm 4\). Then

\[
\begin{align*}
R_3 &\rightarrow \frac{1}{a^2-16}R_3 \\
R_1 &\rightarrow R_1 - R_3 \\
R_2 &\rightarrow R_2 + 2R_3
\end{align*}
\]

and we get the unique solution

\[
x = \frac{8a + 25}{7(a + 4)}, \quad y = \frac{10a + 54}{7(a + 4)}, \quad z = \frac{1}{a + 4}.
\]

Case 2: \(a = -4\). Then \(B = \begin{bmatrix} 1 & 0 & 1 & 8 \frac{8}{7} \\ 0 & 1 & -2 & 10 \frac{10}{7} \\ 0 & 0 & 0 & -8 \end{bmatrix}\), so our system is inconsistent.

Case 3: \(a = 4\). Then \(B = \begin{bmatrix} 1 & 0 & 1 & 8 \frac{8}{7} \\ 0 & 1 & -2 & 10 \frac{10}{7} \\ 0 & 0 & 0 & 0 \end{bmatrix}\). We read off that the system is consistent, with complete solution \(x = \frac{8}{7} - z, \quad y = \frac{10}{7} + 2z, \) where \(z\) is arbitrary.

12. We reduce the augmented array of the system to reduced row–echelon form:

\[
\begin{align*}
\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} &\rightarrow R_3 + R_1 \\
\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} &\rightarrow R_1 + R_4 \\
\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} &\rightarrow R_3 \leftrightarrow R_4
\end{align*}
\]

The last matrix is in reduced row–echelon form and we read off the solution of the corresponding homogeneous system:

\[
\begin{align*}
x_1 &= -x_4 - x_5 = x_4 + x_5 \\
x_2 &= -x_4 - x_5 = x_4 + x_5 \\
x_3 &= -x_4 = x_4,
\end{align*}
\]
where \( x_4 \) and \( x_5 \) are arbitrary elements of \( \mathbb{Z}_2 \). Hence there are four solutions:

\[
\begin{bmatrix}
  x_1 & x_2 & x_3 & x_4 & x_5 \\
  0 & 0 & 0 & 0 & 0 \\
  1 & 1 & 0 & 0 & 1 \\
  1 & 1 & 1 & 1 & 0 \\
  0 & 0 & 1 & 1 & 1
\end{bmatrix}.
\]

13. (a) We reduce the augmented matrix to reduced row–echelon form:

\[
\begin{bmatrix}
  2 & 1 & 3 & 4 \\
  4 & 1 & 4 & 1 \\
  3 & 1 & 2 & 0
\end{bmatrix}
\]

\[
R_1 \rightarrow 3R_1
\]

\[
\begin{bmatrix}
  1 & 3 & 4 & 2 \\
  4 & 1 & 4 & 1 \\
  3 & 1 & 2 & 0
\end{bmatrix}
\]

\[
R_2 \rightarrow R_2 + R_1
\]

\[
R_3 \rightarrow R_3 + 2R_1
\]

Consequently the system has the unique solution \( x = 1, y = 2, z = 0 \).

(b) Again we reduce the augmented matrix to reduced row–echelon form:

\[
\begin{bmatrix}
  2 & 1 & 3 & 4 \\
  4 & 1 & 4 & 1 \\
  1 & 1 & 0 & 3
\end{bmatrix}
\]

\[
R_1 \leftrightarrow R_3
\]

\[
\begin{bmatrix}
  1 & 1 & 0 & 3 \\
  4 & 1 & 4 & 1 \\
  2 & 1 & 3 & 4
\end{bmatrix}
\]

\[
R_2 \rightarrow R_2 + R_1
\]

\[
R_3 \rightarrow R_3 + 3R_1
\]

\[
R_1 \rightarrow R_1 + 2R_2
\]

\[
R_2 \rightarrow R_2 + 3R_3
\]

We read off the complete solution

\[
x = 1 - 3z = 1 + 2z \\
y = 2 - 2z = 2 + 3z,
\]

where \( z \) is an arbitrary element of \( \mathbb{Z}_5 \).
14. Suppose that \((\alpha_1, \ldots, \alpha_n)\) and \((\beta_1, \ldots, \beta_n)\) are solutions of the system of linear equations

\[
\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad 1 \leq i \leq m.
\]

Then

\[
\sum_{j=1}^{n} a_{ij} \alpha_j = b_i \quad \text{and} \quad \sum_{j=1}^{n} a_{ij} \beta_j = b_i
\]

for \(1 \leq i \leq m\).

Let \(i = (1 - t)\alpha_i + t\beta_i\) for \(1 \leq i \leq m\). Then \((\alpha_1, \ldots, \alpha_n)\) is a solution of the given system. For

\[
\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \{(1 - t)\alpha_j + t\beta_j\}
\]

\[
= \sum_{j=1}^{n} a_{ij} (1 - t)\alpha_j + \sum_{j=1}^{n} a_{ij} t\beta_j
\]

\[
= (1 - t)b_i + tb_i
\]

\[
= b_i.
\]

15. Suppose that \((\alpha_1, \ldots, \alpha_n)\) is a solution of the system of linear equations

\[
\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad 1 \leq i \leq m.
\]  \hspace{1cm} (1)

Then the system can be rewritten as

\[
\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \alpha_j, \quad 1 \leq i \leq m,
\]

or equivalently

\[
\sum_{j=1}^{n} a_{ij} (x_j - \alpha_j) = 0, \quad 1 \leq i \leq m.
\]

So we have

\[
\sum_{j=1}^{n} a_{ij} y_j = 0, \quad 1 \leq i \leq m.
\]

where \(x_j - \alpha_j = y_j\). Hence \(x_j = \alpha_j + y_j\), \(1 \leq j \leq n\), where \((y_1, \ldots, y_n)\) is a solution of the associated homogeneous system. Conversely if \((y_1, \ldots, y_n)\)
is a solution of the associated homogeneous system and $x_j = \alpha_j + y_j$, $1 \leq j \leq n$, then reversing the argument shows that $(x_1, \ldots, x_n)$ is a solution of the system 1.

16. We simplify the augmented matrix using row operations, working towards row-echelon form:

$$
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
a & 1 & 1 & 1 & b \\
3 & 2 & 0 & a & 1 + a
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
1 & 1 & 1 & 1 & b \\
0 & 1 - a & 1 + a & 1 - a & b - a
\end{bmatrix}
\rightarrow
\begin{bmatrix}
0 & 1 & -3 & 3 - a & 2 - a \\
0 & 1 - a & 1 + a & 1 - a & b - a
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 1 & -3 & 3 - a & 2 - a \\
0 & 0 & 4 - 2a & (1 - a)(a - 2) & -a^2 + 2a + b - 2
\end{bmatrix}
= B.
$$

Case 1: $a \neq 2$. Then $4 - 2a \neq 0$ and

$$
B \rightarrow \begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 1 & -3 & 3 - a & 2 - a \\
0 & 0 & 1 & a - 1 & -2a^2 + 4a + b - 2
\end{bmatrix}.
$$

Hence we can solve for $x$, $y$, and $z$ in terms of the arbitrary variable $w$.

Case 2: $a = 2$. Then

$$
B = \begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 1 & -3 & 1 & 0 \\
0 & 0 & 0 & 1 & b - 2
\end{bmatrix}.
$$

Hence there is no solution if $b \neq 2$. However if $b = 2$, then

$$
B = \begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 1 & -3 & 1 & 0 \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & 2 & 0 & 1 \\
0 & 1 & -3 & 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

and we get the solution $x = 1 - 2z$, $y = 3z - w$, where $w$ is arbitrary.

17. (a) We first prove that $1 + 1 + 1 + 1 = 0$. Observe that the elements

$$
1 + 0, \quad 1 + 1, \quad 1 + a, \quad 1 + b
$$

are solutions of the associated homogeneous system.
are distinct elements of $F$ by virtue of the *cancellation law for addition*. For this law states that $1 + x = 1 + y \Rightarrow x = y$ and hence $x \neq y \Rightarrow 1 + x \neq 1 + y$.

Hence the above four elements are just the elements 0, 1, $a$, $b$ in some order. Consequently

$$(1 + 0) + (1 + 1) + (1 + a) + (1 + b) = 0 + 1 + a + b$$

$$(1 + 1 + 1 + 1) + (0 + 1 + a + b) = 0 + (0 + 1 + a + b),$$

so $1 + 1 + 1 + 1 = 0$ after cancellation.

Now $1 + 1 + 1 + 1 = (1 + 1)(1 + 1)$, so we have $x^2 = 0$, where $x = 1 + 1$. Hence $x = 0$. Then $a + a = a(1 + 1) = a \cdot 0 = 0$.

Next $a + b = 1$. For $a + b$ must be one of 0, 1, $a$, $b$. Clearly we can’t have $a + b = a$ or $b$; also if $a + b = 0$, then $a + b = a + a$ and hence $b = a$; hence $a + b = 1$. Then

$$a + 1 = a + (a + b) = (a + a) + b = 0 + b = b.$$  

Similarly $b + 1 = a$. Consequently the addition table for $F$ is

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>$a$</td>
<td>$b$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$b$</td>
<td>$a$</td>
</tr>
<tr>
<td>$a$</td>
<td>$a$</td>
<td>$b$</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$b$</td>
<td>$b$</td>
<td>$a$</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

We now find the multiplication table. First, $ab$ must be one of 1, $a$, $b$; however we can’t have $ab = a$ or $b$, so this leaves $ab = 1$.

Next $a^2 = b$. For $a^2$ must be one of 1, $a$, $b$; however $a^2 = a \Rightarrow a = 0$ or $a = 1$; also

$$a^2 = 1 \Rightarrow a^2 - 1 = 0 \Rightarrow (a - 1)(a + 1) = 0 \Rightarrow (a - 1)^2 = 0 \Rightarrow a = 1;$$

hence $a^2 = b$. Similarly $b^2 = a$. Consequently the multiplication table for $F$ is

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$a$</td>
<td>$b$</td>
</tr>
<tr>
<td>$a$</td>
<td>0</td>
<td>$a$</td>
<td>$b$</td>
<td>1</td>
</tr>
<tr>
<td>$b$</td>
<td>0</td>
<td>$b$</td>
<td>1</td>
<td>$a$</td>
</tr>
</tbody>
</table>

(b) We use the addition and multiplication tables for $F$:

$$A = \begin{bmatrix} 1 & a & b & a \\ a & b & a & 1 \\ 1 & 1 & 1 & a \end{bmatrix}$$

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$a$</td>
<td>0</td>
</tr>
<tr>
<td>$a$</td>
<td>0</td>
<td>$a$</td>
<td>$a$</td>
<td>0</td>
</tr>
<tr>
<td>$b$</td>
<td>0</td>
<td>$b$</td>
<td>$a$</td>
<td>0</td>
</tr>
</tbody>
</table>
\[ R_2 \leftrightarrow R_3 \begin{bmatrix} 1 & a & b & a \\ 0 & b & a & 0 \\ 0 & 0 & a & a \end{bmatrix} , \quad R_2 \rightarrow aR_2 \begin{bmatrix} 1 & a & b & a \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} , \quad R_3 \rightarrow bR_3 \begin{bmatrix} 1 & a & b & a \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} . \]

\[ R_1 \leftrightarrow R_1 + aR_2 \begin{bmatrix} 1 & 0 & a & a \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} , \quad R_1 \rightarrow R_1 + aR_3 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & 1 \end{bmatrix} . \]

The last matrix is in reduced row–echelon form.
Section 2.4

2. Suppose \( B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \) and that \( AB = I_2 \). Then

\[
\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -a + e & -b + f \\ c + e & d + f \end{bmatrix}.
\]

Hence

\[
-a + e = 1 \quad -b + f = 0 \\
\begin{align*}
c + e &= 0 \quad d + f = 1 \quad ; \\
e &= a + 1 \quad f = b \\
c &= -e = -(a + 1) \quad d = 1 - f = 1 - b \quad ;
\end{align*}
\]

\[
B = \begin{bmatrix} a & b \\ -a - 1 & 1 - b \\ a + 1 & b \end{bmatrix}.
\]

Next,

\[
(BA)^2 B = (BA)(BA) B = B(AB)(AB) = BI_2 I_2 = BI_2 = B
\]

4. Let \( p_n \) denote the statement

\[
A^n = \frac{1}{2} \left(3^n - 1\right) A + \frac{3 - 3^n}{2} I_2.
\]

Then \( p_1 \) asserts that \( A = \frac{1}{2} \left(3 - 1\right) A + \frac{3 - 3}{2} I_2 \), which is true. So let \( n \geq 1 \) and assume \( p_n \). Then from (1),

\[
A^{n+1} = A \cdot A^n = A \cdot \frac{1}{2} \left(3^n - 1\right) A + \frac{3 - 3^n}{2} I_2 = \frac{1}{2} A^2 + \frac{3 - 3^n}{2} A
\]

\[
= \frac{1}{2} \left(4A - 3I_2\right) + \frac{3 - 3^n}{2} A = \frac{1}{2} \left(4A - 3I_2 + \left(3 - 3^n\right) A\right)
\]

\[
= \frac{1}{2} \left(4\frac{3^n - 3^n}{2} A + \left(3 - 3^n\right) I_2\right)
\]

\[
= \frac{1}{2} \left(4\frac{3^n - 3^n}{2} A + \left(3 - 3^n\right) I_2\right).
\]

Hence \( p_{n+1} \) is true and the induction proceeds.

5. The equation \( x_{n+1} = ax_n + bx_{n-1} \) is seen to be equivalent to

\[
\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}
\]
or

\[ X_n = AX_{n-1}, \]

where \( X_n = \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} \) and \( A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \). Then

\[ X_n = A^n X_0 \]

if \( n \geq 1 \). Hence by Question 3,

\[
\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \left\{ \frac{(3^n - 1)}{2} A + \frac{(3 - 3^n)}{2} I_2 \right\} \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = \left\{ \frac{(3^n - 1)}{2} \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 3-3^n \\ 0 & 3-3^n \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = \begin{bmatrix} (3^n - 1)2 + \frac{3-3^n}{2} (3^n - 1)(-3) \\ \frac{3^n-1}{2} \frac{3-3^n}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}
\]

Hence, equating the \((2,1)\) elements gives

\[ x_n = \frac{(3^n - 1)}{2} x_1 + \frac{(3 - 3^n)}{2} x_0 \quad \text{if } n \geq 1 \]

7. Note: \( \lambda_1 + \lambda_2 = a + d \) and \( \lambda_1 \lambda_2 = ad - bc \).

Then

\[
(\lambda_1 + \lambda_2)k_n - \lambda_1 \lambda_2 k_{n-1} = (\lambda_1 + \lambda_2)(\lambda_1^{n-1} + \lambda_1^{n-2} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-2} + \lambda_2^{n-1})
\]

\[
-\lambda_1 \lambda_2 (\lambda_1^{n-2} + \lambda_1^{n-3} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-3} + \lambda_2^{n-2})
\]

\[
= (\lambda_1^n + \lambda_1^{n-1} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-1})
\]

\[
+ (\lambda_1^{n-1} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-1} + \lambda_2^n)
\]

\[
- (\lambda_1^{n-1} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-1})
\]

\[
= \lambda_1^n + \lambda_1^{n-1} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-1} + \lambda_2^n = k_{n+1}
\]

If \( \lambda_1 = \lambda_2 \), we see

\[
k_n = \lambda_1^{n-1} + \lambda_1^{n-2} \lambda_2 + \cdots + \lambda_1 \lambda_2^{n-2} + \lambda_2^{n-1}
\]

\[
= \lambda_1^{n-1} + \lambda_1^{n-2} \lambda_1 + \cdots + \lambda_1 \lambda_1^{n-2} + \lambda_1^{n-1}
\]

\[
= n \lambda_1^{n-1}
\]

13
If $\lambda_1 \neq \lambda_2$, we see that

\[
(\lambda_1 - \lambda_2)k_n = (\lambda_1 - \lambda_2)(\lambda_1^{n-1} + \lambda_1^{n-2}\lambda_2 + \cdots + \lambda_1\lambda_2^{n-2} + \lambda_2^{n-1})
\]

\[
= \lambda_1^n + \lambda_1^{n-1}\lambda_2 + \cdots + \lambda_1\lambda_2^{n-1}
\]

\[
- (\lambda_1^{n-1}\lambda_2 + \cdots + \lambda_1\lambda_2^{n-2} + \lambda_2^n)
\]

\[
= \lambda_1^n - \lambda_2^n.
\]

Hence $k_n = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2}$.

We have to prove

\[
A^n = k_n A - \lambda_1\lambda_2 k_{n-1} I_2. \quad *
\]

$n=1$:

\[
A^1 = A; \quad \text{also } k_1 A - \lambda_1\lambda_2 k_0 I_2 = k_1 A - \lambda_1\lambda_2 0 I_2 = A.
\]

Let $n \geq 1$ and assume equation $*$ holds. Then

\[
A^{n+1} = A^n \cdot A = (k_n A - \lambda_1\lambda_2 k_{n-1} I_2)A
\]

\[
= k_n A^2 - \lambda_1\lambda_2 k_{n-1} A.
\]

Now $A^2 = (a + d)A - (ad - bc)I_2 = (\lambda_1 + \lambda_2)A - \lambda_1\lambda_2 I_2$. Hence

\[
A^{n+1} = k_n(\lambda_1 + \lambda_2)A - \lambda_1\lambda_2 k_{n} A - \lambda_1\lambda_2 k_{n-1} A
\]

\[
= \{k_n(\lambda_1 + \lambda_2) - \lambda_1\lambda_2 k_{n-1}\}A - \lambda_1\lambda_2 k_{n} I_2
\]

\[
= k_{n+1} A - \lambda_1\lambda_2 k_{n} I_2,
\]

and the induction goes through.

8. Here $\lambda_1, \lambda_2$ are the roots of the polynomial $x^2 - 2x - 3 = (x - 3)(x + 1)$.

So we can take $\lambda_1 = 3, \lambda_2 = -1$. Then

\[
k_n = \frac{3^n - (-1)^n}{3 - (-1)} = \frac{3^n + (-1)^{n+1}}{4}.
\]

Hence

\[
A^n = \left\{ \frac{3^n + (-1)^{n+1}}{4} \right\} A - \left( -3 \right) \left\{ \frac{3^{n-1} + (-1)^{n}}{4} \right\} I_2
\]

\[
= \frac{3^n + (-1)^{n+1}}{4} \left[ \begin{array}{cc} 1 & 2 \\ 2 & 1 \end{array} \right] + 3 \left\{ \frac{3^{n-1} + (-1)^{n}}{4} \right\} \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right],
\]
which is equivalent to the stated result.

9. In terms of matrices, we have

\[
\begin{bmatrix}
F_{n+1} \\
F_n
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix} \begin{bmatrix}
F_n \\
F_{n-1}
\end{bmatrix} \text{ for } n \geq 1.
\]

\[
\begin{bmatrix}
F_{n+1} \\
F_n
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix}^n \begin{bmatrix}
F_1 \\
F_0
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix} \begin{bmatrix}
1 \\
0
\end{bmatrix}.
\]

Now \(\lambda_1, \lambda_2\) are the roots of the polynomial \(x^2 - x - 1\) here.

Hence \(\lambda_1 = \frac{1 + \sqrt{5}}{2}\) and \(\lambda_2 = \frac{1 - \sqrt{5}}{2}\) and

\[
k_n = \frac{\left(\frac{1 + \sqrt{5}}{2}\right)^{n-1} - \left(\frac{1 - \sqrt{5}}{2}\right)^{n-1}}{\sqrt{5}}.
\]

Hence

\[
A^n = k_n A - \lambda_1 \lambda_2 k_{n-1} I_2
\]

\[
= k_n A + k_{n-1} I_2
\]

So

\[
\begin{bmatrix}
F_{n+1} \\
F_n
\end{bmatrix} = (k_n A + k_{n-1} I_2) \begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

\[
= k_n \begin{bmatrix}
1 \\
1
\end{bmatrix} + k_{n-1} \begin{bmatrix}
1 \\
0
\end{bmatrix} = \begin{bmatrix}
k_n + k_{n-1} \\
k_n
\end{bmatrix}.
\]

Hence

\[
F_n = k_n = \frac{\left(\frac{1 + \sqrt{5}}{2}\right)^{n-1} - \left(\frac{1 - \sqrt{5}}{2}\right)^{n-1}}{\sqrt{5}}.
\]

10. From Question 5, we know that

\[
\begin{bmatrix}
x_n \\
y_n
\end{bmatrix} = \begin{bmatrix}
1 & r \\
1 & 1
\end{bmatrix}^n \begin{bmatrix}
a \\
b
\end{bmatrix}.
\]
Now by Question 7, with \( A = \begin{bmatrix} 1 & r \\ 1 & 1 \end{bmatrix} \),

\[
A^n = k_n A - \lambda_1 \lambda_2 k_{n-1} I_2 = k_n A - (1 - r)k_{n-1} I_2,
\]

where \( \lambda_1 = 1 + \sqrt{r} \) and \( \lambda_2 = 1 - \sqrt{r} \) are the roots of the polynomial \( x^2 - 2x + (1 - r) \) and

\[
k_n = \frac{\lambda_1^n - \lambda_2^n}{2\sqrt{r}}.
\]

Hence

\[
\begin{bmatrix} x_n \\ y_n \end{bmatrix} = (k_n A - (1 - r)k_{n-1} I_2) \begin{bmatrix} a \\ b \end{bmatrix}
\]

\[
= \left( \begin{bmatrix} k_n & k_n r \\ k_n & k_n \end{bmatrix} - \begin{bmatrix} (1 - r)k_{n-1} & 0 \\ 0 & (1 - r)k_{n-1} \end{bmatrix} \right) \begin{bmatrix} a \\ b \end{bmatrix}
\]

\[
= \begin{bmatrix} k_n - (1 - r)k_{n-1} & k_n r \\ k_n & k_n - (1 - r)k_{n-1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}
\]

\[
= \begin{bmatrix} a(k_n - (1 - r)k_{n-1}) + bk_n r \\ ak_n + b(k_n - (1 - r)k_{n-1}) \end{bmatrix}.
\]

Hence, in view of the fact that

\[
\frac{k_n}{k_{n-1}} = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1^{n-1} - \lambda_2^{n-1}} = \frac{\lambda_1^n (1 - (\frac{\lambda_2}{\lambda_1})^n)}{\lambda_1^{n-1} (1 - (\frac{\lambda_2}{\lambda_1})^{n-1})} \to \lambda_1, \quad \text{as} \quad n \to \infty,
\]

we have

\[
\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \frac{a(k_n - (1 - r)k_{n-1}) + bk_n r}{ak_n + b(k_n - (1 - r)k_{n-1})}
\]

\[
= \frac{a(k_n - (1 - r)k_{n-1}) + bk_n r}{ak_n + b(k_n - (1 - r)k_{n-1})}
\]

\[
= \frac{a(\sqrt{r} + r) + b(1 + \sqrt{r})r}{a(1 + \sqrt{r}) + b(\sqrt{r} + r)}
\]

\[
= \sqrt{r} {a(1 + \sqrt{r}) + b(1 + \sqrt{r})\sqrt{r}}
\]

\[
= \sqrt{r}.
\]

16
Section 2.7

1. \[ A | I_2 \] = \[ \frac{1}{13} 4 1 0 1 \] \[ R_2 \rightarrow R_2 + 3R_1 \] \[ \begin{bmatrix} 1 & 4 & 1 & 0 & 1 \\ 0 & 13 & 0 & 1 & 0 \end{bmatrix} \]

\[ R_2 \rightarrow \frac{1}{13} R_2 \] \[ \begin{bmatrix} 1 & 4 & 1 & 0 & 1 \\ 0 & 3/13 & 0 & 1 & 0 \end{bmatrix} \] \[ R_1 \rightarrow R_1 - 4R_2 \] \[ \begin{bmatrix} 1/13 & 0 & 1/13 & -4/13 & 0 \\ 0 & 1/13 & 3/13 & 1/13 & 0 \end{bmatrix} \].

Hence \( A \) is non–singular and \( A^{-1} = \begin{bmatrix} 1/13 & -4/13 \\ 3/13 & 1/13 \end{bmatrix} \).

Moreover

\[ E_{12}(-4)E_2(1/13)E_{21}(3)A = I_2, \]

so

\[ A^{-1} = E_{12}(-4)E_2(1/13)E_{21}(3). \]

Hence

\[ A = \{E_{21}(3)\}^{-1}\{E_2(1/13)\}^{-1}\{E_{12}(-4)\}^{-1} = E_{21}(-3)E_2(13)E_{12}(4). \]

2. Let \( D = [d_{ij}] \) be an \( m \times m \) diagonal matrix and let \( A = [a_{jk}] \) be an \( m \times n \) matrix. Then

\[ (DA)_{ik} = \sum_{j=1}^{n} d_{ij}a_{jk} = d_{ii}a_{ik}, \]

as \( d_{ij} = 0 \) if \( i \neq j \). It follows that the \( i \)th row of \( DA \) is obtained by multiplying the \( i \)th row of \( A \) by \( d_{ii} \).

Similarly, post–multiplication of a matrix by a diagonal matrix \( D \) results in a matrix whose columns are those of \( A \), multiplied by the respective diagonal elements of \( D \).

In particular,

\[ \text{diag}(a_1, \ldots, a_n)\text{diag}(b_1, \ldots, b_n) = \text{diag}(a_1b_1, \ldots, a_nb_n), \]

as the left–hand side can be regarded as pre–multiplication of the matrix \( \text{diag}(b_1, \ldots, b_n) \) by the diagonal matrix \( \text{diag}(a_1, \ldots, a_n) \).

Finally, suppose that each of \( a_1, \ldots, a_n \) is non–zero. Then \( a_1^{-1}, \ldots, a_n^{-1} \) all exist and we have

\[ \text{diag}(a_1, \ldots, a_n)\text{diag}(a_1^{-1}, \ldots, a_n^{-1}) = \text{diag}(a_1a_1^{-1}, \ldots, a_na_n^{-1}) = \text{diag}(1, \ldots, 1) = I_n. \]

Hence \( \text{diag}(a_1, \ldots, a_n) \) is non–singular and its inverse is \( \text{diag}(a_1^{-1}, \ldots, a_n^{-1}) \).
Next suppose that \( a_i = 0 \). Then \( \text{diag}(a_1, \ldots, a_n) \) is row-equivalent to a matrix containing a zero row and is hence singular.

3. \( [A|I_3] = \begin{bmatrix} 0 & 0 & 2 & 1 & 0 \\ 1 & 2 & 6 & 0 & 1 \\ 3 & 7 & 9 & 0 & 0 \end{bmatrix} \)

\[
R_1 \leftrightarrow R_2 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 3 & 7 & 9 & 0 & 1 \end{bmatrix}
\]

\[
R_3 \rightarrow R_3 - 3R_1 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & -9 & 0 & -3 \end{bmatrix} \]

\[
R_2 \leftrightarrow R_3 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 1 & -9 & 0 & -3 \\ 0 & 0 & 2 & 1 & 0 \end{bmatrix}
\]

\[
R_3 \rightarrow \frac{1}{2}R_3 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 1 & -9 & 0 & -3 \\ 0 & 0 & 1 & 1/2 & 0 \end{bmatrix}
\]

\[
R_1 \rightarrow R_1 - 24R_3 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 1 & -9 & 0 & -3 \\ 0 & 0 & 1 & 1/2 & 0 \end{bmatrix}
\]

\[
R_2 \rightarrow R_2 + 9R_3 \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 6 & 0 & 1 \\ 0 & 1 & -9 & 0 & -3 \\ 0 & 0 & 1 & 1/2 & 0 \end{bmatrix}
\]

Hence \( A \) is non-singular and \( A^{-1} = \begin{bmatrix} -12 & 7 & -2 \\ 9/2 & -3 & 1 \\ 1/2 & 0 & 0 \end{bmatrix} \).

Also

\[
E_{23}(9)E_{13}(-24)E_{12}(-2)E_3(1/2)E_{23}E_{31}(-3)E_{12}A = I_3.
\]

Hence

\[
A^{-1} = E_{23}(9)E_{13}(-24)E_{12}(-2)E_3(1/2)E_{23}E_{31}(-3)E_{12},
\]

so

\[
A = E_{12}E_{31}(3)E_{23}E_{32}(2)E_{12}(2)E_{13}(24)E_{23}(-9).
\]

4.

\[
A = \begin{bmatrix} 1 & 2 & k \\ 3 & -1 & 1 \\ 5 & 3 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & k \\ 0 & -7 & 1-3k \\ 0 & -7 & -5-5k \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & k \\ 0 & -7 & 1-3k \\ 0 & 0 & -6-2k \end{bmatrix} = B.
\]

Hence if \(-6-2k \neq 0\), i.e. if \( k \neq -3 \), we see that \( B \) can be reduced to \( I_3 \) and hence \( A \) is non-singular.

If \( k = -3 \), then \( B = \begin{bmatrix} 1 & 2 & -3 \\ 0 & -7 & 10 \\ 0 & 0 & 0 \end{bmatrix} = B \) and consequently \( A \) is singular, as it is row-equivalent to a matrix containing a zero row.
5. \( E_{21}(2) \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \). Hence, as in the previous question, \( \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix} \) is singular.

6. Starting from the equation \( A^2 - 2A + 13I_2 = 0 \), we deduce

\[
A(A - 2I_2) = -13I_2 = (A - 2I_2)A.
\]

Hence \( AB = BA = I_2 \), where \( B = \frac{1}{13}(A - 2I_2) \). Consequently \( A \) is non-singular and \( A^{-1} = B \).

7. We assume the equation \( A^3 = 3A^2 - 3A + I_3 \).

(ii) \( A^4 = A^3A = (3A^2 - 3A + I_3)A = 3A^3 - 3A^2 + A = 3(3A^2 - 3A + I_3) - 3A^2 + A = 6A^2 - 8A + 3I_3 \).

(iii) \( A^3 - 3A^2 + 3A = I_3 \). Hence

\[
A(A^2 - 3A + 3I_3) = I_3 = (A^2 - 3A + 3I_3)A.
\]

Hence \( A \) is non-singular and

\[
A^{-1} = A^2 - 3A + 3I_3 = \begin{bmatrix} -1 & -3 & 1 \\ 2 & 4 & -1 \\ 0 & 1 & 0 \end{bmatrix}.
\]

8. (i) If \( B^3 = 0 \) then

\[
(I_n - B)(I_n + B + B^2) = I_n(I_n + B + B^2) - B(I_n + B + B^2) = (I_n + B + B^2) - (B + B^2 + B^3) = I_n - B^3 = I_n - 0 = I_n.
\]

Similarly \( (I_n + B + B^2)(I_n - B) = I_n \).

Hence \( A = I_n - B \) is non-singular and \( A^{-1} = I_n + B + B^2 \).

It follows that the system \( AX = b \) has the unique solution

\[
X = A^{-1}b = (I_n + B + B^2)b = b + Bb + B^2b.
\]
(ii) Let \( B = \begin{bmatrix} 0 & r & s \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix} \). Then \( B^2 = \begin{bmatrix} 0 & 0 & rt \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \) and \( B^3 = 0 \). Hence from the preceding question

\[
(I_3 - B)^{-1} = I_3 + B + B^2
\]

\[
= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & r & s \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & rt \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\]

\[
= \begin{bmatrix} 1 & r & s + rt \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.
\]

9. (i) Suppose that \( A^2 = 0 \). Then if \( A^{-1} \) exists, we deduce that \( A^{-1}(AA) = A^{-1}0 \), which gives \( A = 0 \) and this is a contradiction, as the zero matrix is singular. We conclude that \( A \) does not have an inverse.

(ii). Suppose that \( A^2 = A \) and that \( A^{-1} \) exists. Then

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

which gives \( A = I_n \). Equivalently, if \( A^2 = A \) and \( A \neq I_n \), then \( A \) does not have an inverse.

10. The system of linear equations

\[
\begin{align*}
x + y - z &= a \\
z &= b \\
2x + y + 2z &= c
\end{align*}
\]

is equivalent to the matrix equation \( AX = B \), where

\[
A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.
\]

By Question 7, \( A^{-1} \) exists and hence the system has the unique solution

\[
X = \begin{bmatrix} -1 & -3 & 1 \\ 2 & 4 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -a - 3b + c \\ -a + 4b - c \\ 2a + 4b - c \end{bmatrix}.
\]

Hence \( x = -a - 3b + c, \ y = 2a + 4b - c, \ z = b \).
12. 

\[ A = E_3(2)E_{14}E_{42}(3) = E_3(2)E_{14} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{bmatrix} \]

\[ = E_3(2) \begin{bmatrix} 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \]

Also

\[ A^{-1} = (E_3(2)E_{14}E_{42}(3))^{-1} \]

\[ = (E_{42}(3))^{-1}E_{14}^{-1}(E_3(2))^{-1} \]

\[ = E_{42}(-3)E_{14}E_3(1/2) \]

\[ = E_{42}(-3)E_{14} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

\[ = E_{42}(-3) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \]

\[ = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 1 & -3 & 0 & 0 \end{bmatrix}. \]

13. (All matrices in this question are over \( \mathbb{Z}_2 \).)

(a) 

\[ \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

21
\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}
\]

Hence \( A \) is non–singular and

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

(b) \( A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1
\end{bmatrix} \)

\( R_4 \rightarrow R_4 + R_1 \)

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

so \( A \) is singular.

14.

(a) \[
\begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 \\
-1 & 1 & 0 & 0 & 1 & 0 \\
2 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\( R_3 \rightarrow \frac{1}{2} R_3 \)

\( R_1 \rightarrow R_1 - R_3 \)

\( R_2 \rightarrow R_2 + R_3 \)

\( R_1 \leftrightarrow R_3 \)

\( R_3 \rightarrow R_3 - R_2 \)

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1/2 \\
0 & 1 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 1 & 1 & -1 & -1
\end{bmatrix}
\]

Hence \( A^{-1} \) exists and

\[
A^{-1} = \begin{bmatrix}
0 & 0 & 1/2 \\
0 & 1 & 1/2 \\
1 & -1 & -1
\end{bmatrix}
\]

(b) \[
\begin{bmatrix}
2 & 2 & 4 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\( R_1 \rightarrow R_1 - 2R_2 \)

\( R_1 \leftrightarrow R_2 \)

\( R_2 \leftrightarrow R_3 \)

\( R_3 \rightarrow R_3 - 2R_2 \)

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

\( R_3 \rightarrow \frac{1}{2} R_3 \)

\[
\begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 1/2 & -1 & -1
\end{bmatrix}
\]

22
\[
R_1 \rightarrow R_1 - R_3 \begin{bmatrix}
1 & 0 & 0 & -1/2 & 2 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 1/2 & -1 & -1
\end{bmatrix}.
\]

Hence \( A^{-1} \) exists and
\[
A^{-1} = \begin{bmatrix}
-1/2 & 2 & 1 \\
0 & 0 & 1 \\
1/2 & -1 & -1
\end{bmatrix}.
\]

(c) \[
\begin{bmatrix}
4 & 6 & -3 \\
0 & 0 & 7 \\
0 & 0 & 5
\end{bmatrix} R_2 \rightarrow \frac{1}{2} R_2 \quad \begin{bmatrix}
4 & 6 & -3 \\
0 & 0 & 1 \\
0 & 0 & 1
\end{bmatrix} R_3 \rightarrow R_3 - R_2 \quad \begin{bmatrix}
4 & 6 & -3 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}.
\]

Hence \( A \) is singular by virtue of the zero row.

(d) \[
\begin{bmatrix}
2 & 0 & 0 & 1 & 0 & 0 \\
0 & -5 & 0 & 0 & 1 & 0 \\
0 & 0 & 7 & 0 & 0 & 1
\end{bmatrix} R_1 \rightarrow R_1 - 2 R_2 \quad \begin{bmatrix}
1 & 0 & 0 & 6 \\
0 & 1 & 2 & 0 \\
0 & 0 & 1 & 2
\end{bmatrix} R_3 \rightarrow R_3 - R_2 \quad \begin{bmatrix}
1 & 0 & 6 & 1 & -2 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

Hence \( A^{-1} \) exists and \( A^{-1} = \text{diag}(1/2, -1/5, 1/7) \).

(Of course this was also immediate from Question 2.)

(e) \[
\begin{bmatrix}
1 & 2 & 4 & 6 & 1 & 0 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 2 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 1
\end{bmatrix} R_1 \rightarrow R_1 - R_2 \quad \begin{bmatrix}
1 & 0 & 0 & 6 \\
0 & 1 & 0 & -4 \\
0 & 0 & 1 & 2 \\
0 & 0 & 0 & 2
\end{bmatrix} R_3 \rightarrow R_3 - 2 R_2 \quad \begin{bmatrix}
1 & 0 & 0 & 6 & 1 & -2 & 0 & 0 \\
0 & 1 & 0 & -4 & 0 & 1 & -2 & 0 \\
0 & 0 & 1 & 2 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

Hence \( A^{-1} \) exists and
\[
A^{-1} = \begin{bmatrix}
1 & -2 & 0 & -3 \\
0 & 1 & -2 & 2 \\
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 1/2
\end{bmatrix}.
\]
\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
5 & 7 & 9
\end{bmatrix}
R_2 \rightarrow R_2 - 4R_1 \\
R_3 \rightarrow R_3 - 5R_1 \\
R_3 \rightarrow R_3 - R_2
\begin{bmatrix}
1 & 2 & 3 \\
0 & -3 & -6 \\
0 & -3 & -6
\end{bmatrix}
\]

Hence \(A\) is singular by virtue of the zero row.

15. Suppose that \(A\) is non–singular. Then
\[
AA^{-1} = I_n = A^{-1}A.
\]

Taking transposes throughout gives
\[
(AA^{-1})^t = I_n^t = (A^{-1}A)^t \\
(A^{-1})^t A^t = I_n = A^t(A^{-1})^t,
\]
so \(A^t\) is non–singular and \((A^t)^{-1} = (A^{-1})^t\).

16. Let \(A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\), where \(ad - bc = 0\). Then the equation
\[
A^2 - (a + d)A + (ad - bc)I_2 = 0
\]
reduces to \(A^2 - (a + d)A = 0\) and hence \(A^2 = (a + d)A\). From the last equation, if \(A^{-1}\) exists, we deduce that \(A = (a + d)I_2\), or
\[
\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a + d & 0 \\ 0 & a + d \end{bmatrix}.
\]
Hence \(a = a + d, b = 0, c = 0, d = a + d\) and \(a = b = c = d = 0\), which contradicts the assumption that \(A\) is non–singular.

17. \(A = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix}\)

\[
R_2 \rightarrow R_2 + aR_1 \\
R_3 \rightarrow R_3 + bR_1
\begin{bmatrix} 1 & a & b \\ 0 & 1 + a^2 & a + ab \\ 0 & ab - c & 1 + b^2 \end{bmatrix}
\]

\[
R_2 \rightarrow \frac{1}{1+a^2}R_2 \\
R_3 \rightarrow (ab - c)R_2
\begin{bmatrix} 1 & a & b \\ 0 & 1 & \frac{c + ab}{1 + a^2} \\ 0 & ab - c & 1 + b^2 \end{bmatrix}
\begin{bmatrix} 1 & a \\ 0 & 1 \\ 0 & 0 \end{bmatrix}
\begin{bmatrix} b \\ \frac{c + ab}{(c - ab)(c + ab)} \end{bmatrix}
\begin{bmatrix} 1 & a \\ 0 & 1 \\ 0 & 0 \end{bmatrix}
\frac{1}{1+a^2}
\begin{bmatrix} b \\ \frac{c + ab}{(c - ab)(c + ab)} \end{bmatrix}
\]

24
Now
\[
1 + b^2 + \frac{(c - ab)(c + ab)}{1 + a^2} = 1 + b^2 + \frac{c^2 - (ab)^2}{1 + a^2} = \frac{1 + a^2 + b^2 + c^2}{1 + a^2} \neq 0.
\]

Hence \(B\) can be reduced to \(I_3\) using four more row operations and consequently \(A\) is non-singular.

18. The proposition is clearly true when \(n = 1\). So let \(n \geq 1\) and assume \((P^{-1}AP)^n = P^{-1}A^nP\). Then
\[
(P^{-1}AP)^{n+1} = (P^{-1}AP)^n(P^{-1}AP) = (P^{-1}A^nP)(P^{-1}AP) = P^{-1}A^n(PP^{-1})AP = P^{-1}A^nIAP = P^{-1}(A^nA)P = P^{-1}A^{n+1}P
\]
and the induction goes through.

19. Let \(A = \begin{bmatrix} 2/3 & 1/4 \\ 1/3 & 3/4 \end{bmatrix}\) and \(P = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}\). Then \(P^{-1} = \frac{1}{7} \begin{bmatrix} 4 & -3 \\ 1 & 1 \end{bmatrix}\).

We then verify that \(P^{-1}AP = \begin{bmatrix} 5/12 & 0 \\ 0 & 1 \end{bmatrix}\). Then from the previous question,
\[
P^{-1}A^nP = (P^{-1}AP)^n = \begin{bmatrix} 5/12 & 0 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} (5/12)^n & 0 \\ 0 & 1^n \end{bmatrix} = \begin{bmatrix} (5/12)^n & 0 \\ 0 & 1 \end{bmatrix}.
\]
Hence
\[
A^n = P \begin{bmatrix} (5/12)^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} (5/12)^n & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{7} \begin{bmatrix} 4 & -3 \\ 1 & 1 \end{bmatrix}
\]
\[
= \frac{1}{7} \begin{bmatrix} 4(5/12)^n + 3 \\ -4(5/12)^n + 4 \end{bmatrix} + \frac{1}{7} \begin{bmatrix} 4(5/12)^n + 3 \\ -4(5/12)^n + 4 \end{bmatrix}.
\]
Notice that $A^n \to \frac{1}{7} \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$ as $n \to \infty$. This problem is a special case of a more general result about Markov matrices.

20. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix whose elements are non-negative real numbers satisfying

$$a \geq 0, b \geq 0, c \geq 0, d \geq 0, a + c = 1 = b + d.$$ 

Also let $P = \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix}$ and suppose that $A \neq I_2$.

(i) $\det P = -(b + c)$. Now $b + c \geq 0$. Also if $b + c = 0$, then we would have $b = c = 0$ and hence $d = a = 1$, resulting in $A = I_2$. Hence $\det P < 0$ and $P$ is non-singular.

Next,

$$P^{-1}AP = \frac{-1}{b+c} \begin{bmatrix} -1 & -1 \\ -c & b \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix} = \frac{-1}{b+c} \begin{bmatrix} -a - c & -b - d \\ -ac + bc & -cb + bd \end{bmatrix} \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix} = \frac{-1}{b+c} \begin{bmatrix} -1 & -1 \\ -ac + bc & -cb + bd \end{bmatrix} \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix} = \frac{-1}{b+c} \begin{bmatrix} -b - c & 0 \\ (-ac + bc)b + (-cb + bd)c & -ac + bc + cb - bd \end{bmatrix}.$$

Now

$$-acb + b^2c - c^2b + bdc = -cb(a + c) + bc(b + d) = -cb + bc = 0.$$ 

Also

$$-(a + d - 1)(b + c) = -ab - ac - db - dc + b + c = -ac + b(1 - a) + c(1 - d) - bd = -ac + bc + cb - bd.$$ 

Hence

$$P^{-1}A = \frac{-1}{b+c} \begin{bmatrix} -(b + c) & 0 \\ 0 & -(a + d - 1)(b + c) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & a + d - 1 \end{bmatrix}.$$
(ii) We next prove that if we impose the extra restriction that $A \neq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $|a + d - 1| < 1$. This will then have the following consequence:

$$A = P \begin{bmatrix} 1 & 0 \\ 0 & a + d - 1 \end{bmatrix} P^{-1}$$

$$A^n = P \begin{bmatrix} 1 & 0 \\ 0 & a + d - 1 \end{bmatrix}^n P^{-1}$$

$$= P \begin{bmatrix} 1 & 0 \\ 0 & (a + d - 1)^n \end{bmatrix} P^{-1}$$

$$\rightarrow P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1}$$

$$= \begin{bmatrix} b & 1 \\ c & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{b+c} \begin{bmatrix} -1 & -1 \\ -c & b \end{bmatrix}$$

$$= -\frac{1}{b+c} \begin{bmatrix} b & 0 \\ c & 0 \end{bmatrix} - \begin{bmatrix} -1 & -1 \\ -c & b \end{bmatrix}$$

$$= -\frac{1}{b+c} \begin{bmatrix} -b & -b \\ -c & -c \end{bmatrix}$$

$$= \frac{1}{b+c} \begin{bmatrix} b & b \\ c & c \end{bmatrix},$$

where we have used the fact that $(a + d - 1)^n \to 0$ as $n \to \infty$.

We first prove the inequality $|a + d - 1| \leq 1$:

$$a + d - 1 \leq 1 + d - 1 = d \leq 1$$

$$a + d - 1 \geq 0 + 0 - 1 = -1.$$  

Next, if $a + d - 1 = 1$, we have $a + d = 2$; so $a = 1 = d$ and hence $c = 0 = b$, contradicting our assumption that $A \neq I_2$. Also if $a + d - 1 = -1$, then $a + d = 0$; so $a = 0 = d$ and hence $c = 1 = b$ and hence $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

22. The system is inconsistent: We work towards reducing the augmented matrix:

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 1 & 5 \\ 3 & 5 & 12 \end{bmatrix} R_2 \rightarrow R_2 - R_1$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 0 & -1 \\ 3 & 5 & 12 \end{bmatrix} R_3 \rightarrow R_3 - 3R_1$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} R_3 \rightarrow R_3 - R_2$$
The last row reveals inconsistency.

The system in matrix form is \( AX = B \), where

\[
A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 5 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 5 \\ 12 \end{bmatrix}.
\]

The normal equations are given by the matrix equation

\[
A^tAX = A^tB.
\]

Now

\[
A^tA = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 11 & 18 \\ 18 & 30 \end{bmatrix}
\]

\[
A^tB = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 12 \end{bmatrix} = \begin{bmatrix} 45 \\ 73 \end{bmatrix}.
\]

Hence the normal equations are

\[
11x + 18y = 45
\]

\[
18x + 30y = 73.
\]

These may be solved, for example, by Cramer’s rule:

\[
x = \frac{\begin{vmatrix} 45 & 18 \\ 73 & 30 \end{vmatrix}}{\begin{vmatrix} 11 & 18 \\ 18 & 30 \end{vmatrix}} = \frac{36}{6} = 6
\]

\[
y = \frac{\begin{vmatrix} 11 & 45 \\ 18 & 73 \end{vmatrix}}{\begin{vmatrix} 11 & 18 \\ 18 & 30 \end{vmatrix}} = \frac{-7}{6}.
\]

23. Substituting the coordinates of the five points into the parabola equation gives the following equations:

\[
a = 0
\]

\[
a + b + c = 0
\]

\[
a + 2b + 4c = -1
\]

\[
a + 3b + 9c = 4
\]

\[
a + 4b + 16c = 8.
\]
The associated normal equations are given by

\[
\begin{bmatrix}
5 & 10 & 30 \\
10 & 30 & 100 \\
30 & 100 & 354
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c
\end{bmatrix}
= 
\begin{bmatrix}
11 \\
42 \\
160
\end{bmatrix},
\]

which have the solution \(a = 1/5, \ b = -2, \ c = 1\).

24. Suppose that \(A\) is symmetric, i.e. \(A^t = A\) and that \(AB\) is defined. Then

\[(B^t AB)^t = B^t A^t (B^t)^t = B^t AB,\]

so \(B^t AB\) is also symmetric.

25. Let \(A\) be \(m \times n\) and \(B\) be \(n \times m\), where \(m > n\). Then the homogeneous system \(BX = 0\) has a non–trivial solution \(X_0\), as the number of unknowns is greater than the number of equations. Then

\[(AB)X_0 = A(BX_0) = A0 = 0\]

and the \(m \times m\) matrix \(AB\) is therefore singular, as \(X_0 \neq 0\).

26. (i) Let \(B\) be a singular \(n \times n\) matrix. Then \(BX = 0\) for some non–zero column vector \(X\). Then \((AB)X = A(BX) = A0 = 0\) and hence \(AB\) is also singular.

(ii) Suppose \(A\) is a singular \(n \times n\) matrix. Then \(A^t\) is also singular and hence by (i) so is \(B^t A^t = (AB)^t\). Consequently \(AB\) is also singular.
Section 3.6

1. (a) Let $S$ be the set of vectors $[x, y]$ satisfying $x = 2y$. Then $S$ is a vector subspace of $\mathbb{R}^2$. For

(i) $[0, 0] \in S$ as $x = 2y$ holds with $x = 0$ and $y = 0$.

(ii) $S$ is closed under addition. For let $[x_1, y_1]$ and $[x_2, y_2]$ belong to $S$. Then $x_1 = 2y_1$ and $x_2 = 2y_2$. Hence

$$x_1 + x_2 = 2y_1 + 2y_2 = 2(y_1 + y_2)$$

and hence

$$[x_1 + x_2, y_1 + y_2] = [x_1, y_1] + [x_2, y_2]$$

belongs to $S$.

(iii) $S$ is closed under scalar multiplication. For let $[x, y] \in S$ and $t \in \mathbb{R}$. Then $x = 2y$ and hence $tx = 2(ty)$. Consequently

$$[tx, ty] = t[x, y] \in S.$$ 

(b) Let $S$ be the set of vectors $[x, y]$ satisfying $x = 2y$ and $2x = y$. Then $S$ is a subspace of $\mathbb{R}^2$. This can be proved in the same way as (a), or alternatively we see that $x = 2y$ and $2x = y$ imply $x = 4x$ and hence $x = 0 = y$. Hence $S = \{[0, 0]\}$, the set consisting of the zero vector. This is always a subspace.

(c) Let $S$ be the set of vectors $[x, y]$ satisfying $x = 2y + 1$. Then $S$ doesn’t contain the zero vector and consequently fails to be a vector subspace.

(d) Let $S$ be the set of vectors $[x, y]$ satisfying $xy = 0$. Then $S$ is not closed under addition of vectors. For example $[1, 0] \in S$ and $[0, 1] \in S$, but $[1, 0] + [0, 1] = [1, 1] \notin S$.

(e) Let $S$ be the set of vectors $[x, y]$ satisfying $x \geq 0$ and $y \geq 0$. Then $S$ is not closed under scalar multiplication. For example $[1, 0] \in S$ and $-1 \in \mathbb{R}$, but $(-1)[1, 0] = [-1, 0] \notin S$.

2. Let $X, Y, Z$ be vectors in $\mathbb{R}^n$. Then by Lemma 3.2.1

$$\langle X + Y, X + Z, Y + Z \rangle \subseteq \langle X, Y, Z \rangle,$$

as each of $X + Y, X + Z, Y + Z$ is a linear combination of $X, Y, Z$. 
Also
\[ X = \frac{1}{2} (X + Y) + \frac{1}{2} (X + Z) - \frac{1}{2} (Y + Z), \]
\[ Y = \frac{1}{2} (X + Y) - \frac{1}{2} (X + Z) + \frac{1}{2} (Y + Z), \]
\[ Z = -\frac{1}{2} (X + Y) + \frac{1}{2} (X + Z) + \frac{1}{2} (Y + Z), \]

so
\[ \langle X, Y, Z \rangle \subseteq \langle X + Y, X + Z, Y + Z \rangle. \]

Hence
\[ \langle X, Y, Z \rangle = \langle X + Y, X + Z, Y + Z \rangle. \]

3. Let \( X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, X_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \) and \( X_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix} \). We have to decide if \( X_1, X_2, X_3 \) are linearly independent, that is if the equation \( xX_1 + yX_2 + zX_3 = 0 \) has only the trivial solution. This equation is equivalent to the following homogeneous system
\[
\begin{align*}
x + 0y + z &= 0 \\
0x + y + z &= 0 \\
x + y + z &= 0 \\
2x + 2y + 3z &= 0.
\end{align*}
\]

We reduce the coefficient matrix to reduced row-echelon form:
\[
\begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 1 \\
2 & 2 & 3
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}
\]

and consequently the system has only the trivial solution \( x = 0, y = 0, z = 0 \). Hence the given vectors are linearly independent.

4. The vectors
\[
X_1 = \begin{bmatrix} \lambda \\ -1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ \lambda \\ -1 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}
\]
are linearly dependent for precisely those values of $\lambda$ for which the equation $xX_1 + yX_2 + zX_3 = 0$ has a non-trivial solution. This equation is equivalent to the system of homogeneous equations

\[
\begin{align*}
\lambda x - y - z &= 0 \\
-x + \lambda y - z &= 0 \\
-x - y + \lambda z &= 0.
\end{align*}
\]

Now the coefficient determinant of this system is

\[
\begin{vmatrix}
\lambda & -1 & -1 \\
-1 & \lambda & -1 \\
-1 & -1 & \lambda
\end{vmatrix} = (\lambda + 1)^2(\lambda - 2).
\]

So the values of $\lambda$ which make $X_1$, $X_2$, $X_3$ linearly independent are those $\lambda$ satisfying $\lambda \neq -1$ and $\lambda \neq 2$.

5. Let $A$ be the following matrix of rationals:

\[
A = \begin{bmatrix}
1 & 1 & 2 & 0 & 1 \\
2 & 2 & 5 & 0 & 3 \\
0 & 0 & 0 & 1 & 3 \\
8 & 11 & 19 & 0 & 11
\end{bmatrix}.
\]

Then $A$ has reduced row–echelon form

\[
B = \begin{bmatrix}
1 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 3
\end{bmatrix}.
\]

From $B$ we read off the following:

(a) The rows of $B$ form a basis for $R(A)$. (Consequently the rows of $A$ also form a basis for $R(A)$.)

(b) The first four columns of $A$ form a basis for $C(A)$.

(c) To find a basis for $N(A)$, we solve $AX = 0$ and equivalently $BX = 0$.

From $B$ we see that the solution is

\[
\begin{align*}
x_1 &= x_5 \\
x_2 &= 0 \\
x_3 &= -x_5 \\
x_4 &= -3x_5.
\end{align*}
\]
with $x_5$ arbitrary. Then

$$X = \begin{bmatrix}
x_5 \\
0 \\
-x_5 \\
-3x_5 \\
x_5
\end{bmatrix} = x_5 \begin{bmatrix}
1 \\
0 \\
-1 \\
-3 \\
1
\end{bmatrix},$$

so $[1, 0, -1, -3, 1]^t$ is a basis for $N(A)$.

6. In Section 1.6, problem 12, we found that the matrix

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

has reduced row–echelon form

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

From $B$ we read off the following:

(a) The three non–zero rows of $B$ form a basis for $R(A)$.

(b) The first three columns of $A$ form a basis for $C(A)$.

(c) To find a basis for $N(A)$, we solve $AX = 0$ and equivalently $BX = 0$.

From $B$ we see that the solution is

$$x_1 = -x_4 - x_5 = x_4 + x_5$$

$$x_2 = -x_4 - x_5 = x_4 + x_5$$

$$x_3 = -x_4 = x_4,$$

with $x_4$ and $x_5$ arbitrary elements of $\mathbb{Z}_2$. Hence

$$X = \begin{bmatrix}
x_4 + x_5 \\
x_4 + x_5 \\
x_4 \\
x_4 \\
x_5
\end{bmatrix} = x_4 \begin{bmatrix}
1 \\
1 \\
1 \\
1 \\
0
\end{bmatrix} + \begin{bmatrix}
1 \\
1 \\
0 \\
0 \\
1
\end{bmatrix}.$$

Hence $[1, 1, 1, 1, 0]^t$ and $[1, 1, 0, 0, 1]^t$ form a basis for $N(A)$.
7. Let \( A \) be the following matrix over \( \mathbb{Z}_5 \):

\[
A = \begin{bmatrix}
1 & 1 & 2 & 0 & 1 & 3 \\
2 & 1 & 4 & 0 & 3 & 2 \\
0 & 0 & 0 & 1 & 3 & 0 \\
3 & 0 & 2 & 4 & 3 & 2
\end{bmatrix}.
\]

We find that \( A \) has reduced row–echelon form \( B \):

\[
B = \begin{bmatrix}
1 & 0 & 0 & 0 & 2 & 4 \\
0 & 1 & 0 & 0 & 4 & 4 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 3 & 0
\end{bmatrix}.
\]

From \( B \) we read off the following:

(a) The four rows of \( B \) form a basis for \( R(A) \). (Consequently the rows of \( A \) also form a basis for \( R(A) \).

(b) The first four columns of \( A \) form a basis for \( C(A) \).

(c) To find a basis for \( N(A) \), we solve \( AX = 0 \) and equivalently \( BX = 0 \). From \( B \) we see that the solution is

\[
\begin{align*}
x_1 &= -2x_5 - 4x_6 = 3x_5 + x_6 \\
x_2 &= -4x_5 - 4x_6 = x_5 + x_6 \\
x_3 &= 0 \\
x_4 &= -3x_5 = 2x_5,
\end{align*}
\]

where \( x_5 \) and \( x_6 \) are arbitrary elements of \( \mathbb{Z}_5 \). Hence

\[
X = x_5 \begin{bmatrix}
3 \\
1 \\
0 \\
2
\end{bmatrix} + x_6 \begin{bmatrix}
1 \\
1 \\
0 \\
0
\end{bmatrix},
\]

so \([3, 1, 0, 2, 1, 0]^t\) and \([1, 1, 0, 0, 0, 1]^t\) form a basis for \( R(A) \).

8. Let \( F = \{0, 1, a, b\} \) be a field and let \( A \) be the following matrix over \( F \):

\[
A = \begin{bmatrix}
1 & a & b & a \\
a & b & b & 1 \\
1 & 1 & 1 & a
\end{bmatrix}.
\]
In Section 1.6, problem 17, we found that $A$ had reduced row–echelon form

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

From $B$ we read off the following:

(a) The rows of $B$ form a basis for $R(A)$. (Consequently the rows of $A$ also form a basis for $R(A)$.

(b) The first three columns of $A$ form a basis for $C(A)$.

(c) To find a basis for $N(A)$, we solve $AX = 0$ and equivalently $BX = 0$. From $B$ we see that the solution is

$$ x_1 = 0 $$
$$ x_2 = -bx_4 = bx_4 $$
$$ x_3 = -x_4 = x_4, $$

where $x_4$ is an arbitrary element of $F$. Hence

$$ X = x_4 \begin{bmatrix} 0 \\ b \\ 1 \\ 1 \end{bmatrix}, $$

so $[0, b, 1, 1]^t$ is a basis for $N(A)$.

9. Suppose that $X_1, \ldots, X_m$ form a basis for a subspace $S$. We have to prove that

$$ X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_m $$

also form a basis for $S$.

First we prove the independence of the family: Suppose

$$ x_1X_1 + x_2(X_1 + X_2) + \cdots + x_m(X_1 + \cdots + X_m) = 0. $$

Then

$$ (x_1 + x_2 + \cdots + x_m)X_1 + \cdots + x_mX_m = 0. $$

Then the linear independence of $X_1, \ldots, X_m$ gives

$$ x_1 + x_2 + \cdots + x_m = 0, \ldots, x_m = 0, $$
form which we deduce that $x_1 = 0, \ldots, x_m = 0$.

Secondly we have to prove that every vector of $S$ is expressible as a linear combination of $X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_m$. Suppose $X \in S$. Then

$$X = a_1 X_1 + \cdots + a_m X_m.$$ 

We have to find $x_1, \ldots, x_m$ such that

$$X = x_1 X_1 + x_2(X_1 + X_2) + \cdots + x_m(X_1 + \cdots + X_m)$$

$$= (x_1 + x_2 + \cdots + x_m)X_1 + \cdots + x_m X_m.$$ 

Then

$$a_1 X_1 + \cdots + a_m X_m = (x_1 + x_2 + \cdots + x_m)X_1 + \cdots + x_m X_m.$$ 

So if we can solve the system

$$x_1 + x_2 + \cdots + x_m = a_1, \ldots, x_m = a_m,$$

we are finished. Clearly these equations have the unique solution

$$x_1 = a_1 - a_2, \ldots, x_{m-1} = a_m - a_{m-1}, x_m = a_m.$$ 

10. Let $A = \begin{bmatrix} a & b & c \\ 1 & 1 & 1 \end{bmatrix}$. If $[a, b, c]$ is a multiple of $[1, 1, 1]$, (that is, $a = b = c$), then rank $A = 1$. For if

$$[a, b, c] = t[1, 1, 1],$$

then

$$R(A) = \langle [a, b, c], [1, 1, 1] \rangle = \langle t[1, 1, 1], [1, 1, 1] \rangle = \langle [1, 1, 1] \rangle,$$

so $[1, 1, 1]$ is a basis for $R(A)$.

However if $[a, b, c]$ is not a multiple of $[1, 1, 1]$, (that is at least two of $a, b, c$ are distinct), then the left–to–right test shows that $[a, b, c]$ and $[1, 1, 1]$ are linearly independent and hence form a basis for $R(A)$. Consequently rank $A = 2$ in this case.

11. Let $S$ be a subspace of $F^n$ with dim $S = m$. Also suppose that $X_1, \ldots, X_m$ are vectors in $S$ such that $S = \langle X_1, \ldots, X_m \rangle$. We have to prove that $X_1, \ldots, X_m$ form a basis for $S$; in other words, we must prove that $X_1, \ldots, X_m$ are linearly independent.
However if $X_1, \ldots, X_m$ were linearly dependent, then one of these vectors would be a linear combination of the remaining vectors. Consequently $S$ would be spanned by $m - 1$ vectors. But there exist a family of $m$ linearly independent vectors in $S$. Then by Theorem 3.3.2, we would have the contradiction $m \leq m - 1$.

12. Let $[x, y, z]^t \in S$. Then $x + 2y + 3z = 0$. Hence $x = -2y - 3z$ and

$$
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} = 
\begin{bmatrix}
  -2y - 3z \\
  y \\
  z
\end{bmatrix} = y \begin{bmatrix}
  -2 \\
  1 \\
  0
\end{bmatrix} + z \begin{bmatrix}
  -3 \\
  0 \\
  1
\end{bmatrix}.
$$

Hence $[-2, 1, 0]^t$ and $[-3, 0, 1]^t$ form a basis for $S$.

Next $(-1) + 2(-1) + 3(1) = 0$, so $[-1, -1, 1]^t \in S$.

To find a basis for $S$ which includes $[-1, -1, 1]^t$, we note that $[-2, 1, 0]^t$ is not a multiple of $[-1, -1, 1]^t$. Hence we have found a linearly independent family of two vectors in $S$, a subspace of dimension equal to 2. Consequently these two vectors form a basis for $S$.

13. Without loss of generality, suppose that $X_1 = X_2$. Then we have the non–trivial dependency relation:

$$1X_1 + (-1)X_2 + 0X_3 + \cdots + 0X_m = 0.$$

14. (a) Suppose that $X_{m+1}$ is a linear combination of $X_1, \ldots, X_m$. Then

$$\langle X_1, \ldots, X_m, X_{m+1} \rangle = \langle X_1, \ldots, X_m \rangle$$

and hence

$$\dim \langle X_1, \ldots, X_m, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle.$$ 

(b) Suppose that $X_{m+1}$ is not a linear combination of $X_1, \ldots, X_m$. If not all of $X_1, \ldots, X_m$ are zero, there will be a subfamily $X_{c_1}, \ldots, X_{c_r}$ which is a basis for $\langle X_1, \ldots, X_m \rangle$.

Then as $X_{m+1}$ is not a linear combination of $X_{c_1}, \ldots, X_{c_r}$, it follows that $X_{c_1}, \ldots, X_{c_r}, X_{m+1}$ are linearly independent. Also

$$\langle X_1, \ldots, X_m, X_{m+1} \rangle = \langle X_{c_1}, \ldots, X_{c_r}, X_{m+1} \rangle.$$ 

Consequently

$$\dim \langle X_1, \ldots, X_m, X_{m+1} \rangle = r + 1 = \dim \langle X_1, \ldots, X_m \rangle + 1.$$
Our result can be rephrased in a form suitable for the second part of the problem:
\[ \dim \langle X_1, \ldots, X_m, X_{m+1} \rangle = \dim \langle X_1, \ldots, X_m \rangle \]
if and only if \( X_{m+1} \) is a linear combination of \( X_1, \ldots, X_m \).

If \( X = [x_1, \ldots, x_n]^t \), then \( AX = B \) is equivalent to
\[ B = x_1A_{s1} + \cdots + x_nA_{sn}. \]

So \( AX = B \) is soluble for \( X \) if and only if \( B \) is a linear combination of the columns of \( A \), that is \( B \in \mathcal{C}(A) \).

15. Let \( a_1, \ldots, a_n \) be elements of \( F \), not all zero. Let \( S \) denote the set of vectors \( [x_1, \ldots, x_n]^t \), where \( x_1, \ldots, x_n \) satisfy
\[ a_1x_1 + \cdots + a_nx_n = 0. \]

Then \( S = N(A) \), where \( A \) is the row matrix \( [a_1, \ldots, a_n] \). Now \( \text{rank } A = 1 \) as \( A \neq 0 \). So by the “rank + nullity” theorem, noting that the number of columns of \( A \) equals \( n \), we have
\[ \dim N(A) = \text{nullity } (A) = n - \text{rank } A = n - 1. \]

16. (a) (Proof of Lemma 3.2.1) Suppose that each of \( X_1, \ldots, X_r \) is a linear combination of \( Y_1, \ldots, Y_s \). Then
\[ X_i = \sum_{j=1}^{s} a_{ij}Y_j, \quad (1 \leq i \leq r). \]

Now let \( X = \sum_{i=1}^{r} x_iX_i \) be a linear combination of \( X_1, \ldots, X_r \). Then
\[ X = x_1(a_{11}Y_1 + \cdots + a_{1s}Y_s) + \cdots + x_r(a_{r1}Y_1 + \cdots + a_{rs}Y_s) = y_1Y_1 + \cdots + y_sY_s, \]
where \( y_j = a_{1j}x_1 + \cdots + a_{rj}x_r \). Hence \( X \) is a linear combination of \( Y_1, \ldots, Y_s \).

Another way of stating Lemma 3.2.1 is
\[ \langle X_1, \ldots, X_r \rangle \subseteq \langle Y_1, \ldots, Y_s \rangle, \quad (1) \]
if each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$.

(b) (Proof of Theorem 3.2.1) Suppose that each of $X_1, \ldots, X_r$ is a linear combination of $Y_1, \ldots, Y_s$ and that each of $Y_1, \ldots, Y_s$ is a linear combination of $X_1, \ldots, X_r$. Then by (a) equation (1) above

$$\langle X_1, \ldots, X_r \rangle \subseteq \langle Y_1, \ldots, Y_s \rangle$$

and

$$\langle Y_1, \ldots, Y_s \rangle \subseteq \langle X_1, \ldots, X_r \rangle.$$

Hence

$$\langle X_1, \ldots, X_r \rangle = \langle Y_1, \ldots, Y_s \rangle.$$

(c) (Proof of Corollary 3.2.1) Suppose that each of $Z_1, \ldots, Z_t$ is a linear combination of $X_1, \ldots, X_r$. Then each of $X_1, \ldots, X_r, Z_1, \ldots, Z_t$ is a linear combination of $X_1, \ldots, X_r$.

Also each of $X_1, \ldots, X_r$ is a linear combination of $X_1, \ldots, X_r, Z_1, \ldots, Z_t$, so by Theorem 3.2.1

$$\langle X_1, \ldots, X_r, Z_1, \ldots, Z_t \rangle = \langle X_1, \ldots, X_r \rangle.$$

(d) (Proof of Theorem 3.3.2) Let $Y_1, \ldots, Y_s$ be vectors in $\langle X_1, \ldots, X_r \rangle$ and assume that $s > r$. We have to prove that $Y_1, \ldots, Y_s$ are linearly dependent. So we consider the equation

$$x_1Y_1 + \cdots + x_sY_s = 0.$$

Now $Y_i = \sum_{j=1}^{r} a_{ij}X_j$, for $1 \leq i \leq s$. Hence

$$x_1Y_1 + \cdots + x_sY_s = x_1(a_{11}X_1 + \cdots + a_{1r}X_r) + \cdots + x_r(a_{s1}X_1 + \cdots + a_{sr}X_r).$$

where $y_j = a_{1j}x_1 + \cdots + a_{sj}x_s$. However the homogeneous system

$$y_1 = 0, \ldots, y_r = 0$$

has a non–trivial solution $x_1, \ldots, x_s$, as $s > r$ and from (1), this results in a non–trivial solution of the equation

$$x_1Y_1 + \cdots + x_sY_s = 0.$$
Hence \( Y_1, \ldots, Y_s \) are linearly dependent.

17. Let \( R \) and \( S \) be subspaces of \( F^n \), with \( R \subseteq S \). We first prove

\[
\dim R \leq \dim S.
\]

Let \( X_1, \ldots, X_r \) be a basis for \( R \). Now by Theorem 3.5.2, because \( X_1, \ldots, X_r \) form a linearly independent family lying in \( S \), this family can be extended to a basis \( X_1, \ldots, X_r, \ldots, X_s \) for \( S \). Then

\[
\dim S = s \geq r = \dim R.
\]

Next suppose that \( \dim R = \dim S \). Let \( X_1, \ldots, X_r \) be a basis for \( R \). Then because \( X_1, \ldots, X_r \) form a linearly independent family in \( S \) and \( S \) is a subspace whose dimension is \( r \), it follows from Theorem 3.4.3 that \( X_1, \ldots, X_r \) form a basis for \( S \). Then

\[
S = \langle X_1, \ldots, X_r \rangle = R.
\]

18. Suppose that \( R \) and \( S \) are subspaces of \( F^n \) with the property that \( R \cup S \) is also a subspace of \( F^n \). We have to prove that \( R \subseteq S \) or \( S \subseteq R \). We argue by contradiction: Suppose that \( R \not\subseteq S \) and \( S \not\subseteq R \). Then there exist vectors \( u \) and \( v \) such that

\[
u \in R \text{ and } v \not\in S, \quad v \in S \text{ and } v \not\in R.
\]

Consider the vector \( u + v \). As we are assuming \( R \cup S \) is a subspace, \( R \cup S \) is closed under addition. Hence \( u + v \in R \cup S \) and so \( u + v \in R \) or \( u + v \in S \). However if \( u + v \in R \), then \( v = (u + v) - u \in R \), which is a contradiction; similarly if \( u + v \in S \).

Hence we have derived a contradiction on the assumption that \( R \not\subseteq S \) and \( S \not\subseteq R \). Consequently at least one of these must be false. In other words \( R \subseteq S \) or \( S \subseteq R \).

19. Let \( X_1, \ldots, X_r \) be a basis for \( S \).

(i) First let

\[
Y_1 = a_{11}X_1 + \cdots + a_{1r}X_r \\
\vdots \\
Y_r = a_{r1}X_1 + \cdots + a_{rr}X_r,
\]
where $A = [a_{ij}]$ is non–singular. Then the above system of equations can be solved for $X_1, \ldots, X_r$ in terms of $Y_1, \ldots, Y_r$. Consequently by Theorem 3.2.1

$$\langle Y_1, \ldots, Y_r \rangle = \langle X_1, \ldots, X_r \rangle = S.$$ 

It follows from problem 11 that $Y_1, \ldots, Y_r$ is a basis for $S$.

(ii) We show that all bases for $S$ are given by equations 2. So suppose that $Y_1, \ldots, Y_r$ forms a basis for $S$. Then because $X_1, \ldots, X_r$ form a basis for $S$, we can express $Y_1, \ldots, Y_r$ in terms of $X_1, \ldots, X_r$ as in 2, for some matrix $A = [a_{ij}]$. We show $A$ is non–singular by demonstrating that the linear independence of $Y_1, \ldots, Y_r$ implies that the rows of $A$ are linearly independent.

So assume

$$x_1[a_{11}, \ldots, a_{1r}] + \cdots + x_r[a_{r1}, \ldots, a_{rr}] = [0, \ldots, 0].$$

Then on equating components, we have

$$a_{11}x_1 + \cdots + a_{r1}x_r = 0$$
$$\vdots$$
$$a_{1r}x_1 + \cdots + a_{rr}x_r = 0.$$

Hence

$$x_1Y_1 + \cdots + x_rY_r = x_1(a_{11}X_1 + \cdots + a_{r1}X_r) + \cdots + x_r(a_{1r}X_1 + \cdots + a_{rr}X_r)$$
$$= (a_{11}x_1 + \cdots + a_{r1}x_r)X_1 + \cdots + (a_{1r}x_1 + \cdots + a_{rr}x_r)X_r$$
$$= 0X_1 + \cdots + 0X_r = 0.$$

Then the linear independence of $Y_1, \ldots, Y_r$ implies $x_1 = 0, \ldots, x_r = 0$.

(We mention that the last argument is reversible and provides an alternative proof of part (i).)
Section 4.1

1. We first prove that the area of a triangle $P_1P_2P_3$, where the points are in anti-clockwise orientation, is given by the formula

$$\frac{1}{2} \left\{ \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix} \right\}.$$

Referring to the above diagram, we have

$$\text{Area } P_1P_2P_3 = \text{Area } OP_1P_2 + \text{Area } OP_2P_3 - \text{Area } OP_1P_3$$

$$= \frac{1}{2} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \frac{1}{2} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \frac{1}{2} \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix},$$

which gives the desired formula.

We now turn to the area of a quadrilateral. One possible configuration occurs when the quadrilateral is convex as in figure (a) below. The interior diagonal breaks the quadrilateral into two triangles $P_1P_2P_3$ and $P_1P_3P_4$. Then

$$\text{Area } P_1P_2P_3P_4 = \text{Area } P_1P_2P_3 + \text{Area } P_1P_3P_4$$

$$= \frac{1}{2} \left\{ \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix} \right\}.$$
Another possible configuration for the quadrilateral occurs when it is not convex, as in figure (b). The interior diagonal $P_2P_4$ then gives two triangles $P_1P_2P_4$ and $P_2P_3P_4$ and we can proceed similarly as before.

2.

\[
\Delta = \begin{vmatrix}
  a+x & b+y & c+z \\
  x+u & y+v & z+w \\
  u+a & v+b & w+c \\
\end{vmatrix} + \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix} = \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix} + \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix}.
\]

Now

\[
\begin{vmatrix}
  a & b & c \\
  x+u & y+v & z+w \\
  u+a & v+b & w+c \\
\end{vmatrix} = \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix} + \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix} + \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix}.
\]

Similarly

\[
\begin{vmatrix}
  x & y & z \\
  x+u & y+v & z+w \\
  u+a & v+b & w+c \\
\end{vmatrix} = -\begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix} = \begin{vmatrix}
  a & b & c \\
  x & y & z \\
  u & v & w \\
\end{vmatrix}.
\]
Hence $\Delta = 2\begin{vmatrix} a & b & c \\ x & y & z \\ u & v & w \end{vmatrix}$.

3. \[
\begin{vmatrix}
\begin{array}{ccc}
(2n+1)^2 & (2n+2)^2 & (2n+3)^2 \\
(2n+2)^2 & (2n+3)^2 & (2n+4)^2 \\
n^2 & (n+1)^2 & (n+2)^2 \\
(n+1)^2 & (n+2)^2 & (n+3)^2 \\
n^2 & 2n+1 & 2n+3 \\
2n+1 & 2n+3 & 2n+5 \\
2n+2 & 2n+4 & 2n+7 \\
\end{array}
\end{vmatrix}
:: \frac{C_3 - C_2}{C_3 - C_2} = \begin{vmatrix}
\begin{array}{ccc}
2n+1 & 2n+3 & 2n+5 \\
2n+2 & 2n+4 & 2n+7 \\
\end{array}
\end{vmatrix}.
\]

\[
\begin{vmatrix}
\begin{array}{ccc}
(2n+1)^2 & (2n+2)^2 & (2n+3)^2 \\
(2n+2)^2 & (2n+3)^2 & (2n+4)^2 \\
n^2 & (n+1)^2 & (n+2)^2 \\
(n+1)^2 & (n+2)^2 & (n+3)^2 \\
n^2 & 2n+1 & 2n+3 \\
2n+1 & 2n+3 & 2n+5 \\
2n+2 & 2n+4 & 2n+7 \\
\end{array}
\end{vmatrix}
:: \frac{R_3 - R_2}{R_2 - R_1} = \begin{vmatrix}
\begin{array}{ccc}
2n+1 & 2n+3 & 2n+5 \\
2n+2 & 2n+4 & 2n+7 \\
\end{array}
\end{vmatrix}.
\]

4. (a)
\[
\begin{vmatrix}
\begin{array}{ccc}
246 & 427 & 327 \\
1014 & 543 & 443 \\
-342 & 721 & 621 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{ccc}
246 & 100 & 327 \\
1014 & 100 & 443 \\
-342 & 100 & 621 \\
\end{array}
\end{vmatrix} = 100\begin{vmatrix}
\begin{array}{ccc}
246 & 1 & 327 \\
768 & 0 & 116 \\
-588 & 0 & 294 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{ccc}
768 & 116 \\
-588 & 294 \\
\end{array}
\end{vmatrix} = 29400000.
\]

(b)
\[
\begin{vmatrix}
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
-2 & 1 & -4 & 3 \\
3 & -4 & -1 & 2 \\
4 & 3 & -2 & -1 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & 5 & 2 & 11 \\
0 & -10 & -10 & -10 \\
0 & -5 & -14 & -17 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{cccc}
5 & 2 & 11 \\
0 & 1 & 1 \\
0 & -5 & -14 & -17 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{cccc}
5 & -3 & 6 \\
1 & 0 & 0 \\
-5 & -9 & -12 \\
\end{array}
\end{vmatrix} = -10\begin{vmatrix}
\begin{array}{cccc}
-3 & 6 \\
-9 & -12 \\
\end{array}
\end{vmatrix} = 900.
\]

5. \[
\begin{vmatrix}
\begin{array}{ccc}
1 & 0 & -2 \\
3 & 1 & 4 \\
5 & 2 & -3 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{ccc}
1 & 0 & 0 \\
3 & 1 & 10 \\
5 & 2 & 7 \\
\end{array}
\end{vmatrix} = \begin{vmatrix}
\begin{array}{ccc}
1 & 0 & 0 \\
10 & 0 & 7 \\
\end{array}
\end{vmatrix} = -13.
\]

44
Hence $A$ is non–singular and

$$A^{-1} = \frac{1}{-13} \text{adj } A = \frac{1}{-13} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = \frac{1}{-13} \begin{bmatrix} -11 & -4 & 2 \\ 29 & 7 & -10 \\ 1 & -2 & 1 \end{bmatrix}.$$  

6. (i) 

$$\begin{vmatrix} 2a & 2b & b-c \\ 2b & 2a & a+c \\ a+b & a+b & b \end{vmatrix} = (a+b) \begin{vmatrix} 2 & 1 \\ 2b & a+c \end{vmatrix} = 2(a+b)(a-b) \begin{vmatrix} 2 & 1 \\ a+b & b \end{vmatrix} = -2(a+b)(a-b)^2.$$  

(ii) 

$$\begin{vmatrix} b+c & b & c \\ c & c+a & a \\ b & a & a+b \end{vmatrix} = \begin{vmatrix} c & b & c \\ -a & c+a & a \\ b-a & a & a+b \end{vmatrix} = -2a \begin{vmatrix} c & b \\ b & c \end{vmatrix} = 2a(c^2 + b^2).$$  

7. Suppose that the curve $y = ax^2 + bx + c$ passes through the points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$, where $x_i \neq x_j$ if $i \neq j$. Then

$$ax_1^2 + bx_1 + c = y_1, \\
ax_2^2 + bx_2 + c = y_2, \\
ax_3^2 + bx_3 + c = y_3.$$  

The coefficient determinant is essentially a Vandermonde determinant:

$$\begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix} = \begin{vmatrix} x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{vmatrix} = -(x_2-x_1)(x_3-x_1)(x_3-x_2).$$
Hence the coefficient determinant is non–zero and by Cramer’s rule, there is a unique solution for \(a, b, c\).

8. Let \(\Delta = \det A = \begin{vmatrix} 1 & 1 & -1 \\ 2 & 3 & k \\ 1 & k & 3 \end{vmatrix}\). Then

\[
\Delta = C_3 \rightarrow C_3 + C_1 \begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & k + 2 \\ 1 & k - 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & k + 2 \\ k - 1 & 4 \end{vmatrix}
\]

\[
= 4 - (k - 1)(k + 2) = - (k^2 - k - 6) = -(k + 3)(k - 2).
\]

Hence \(\det A = 0\) if and only if \(k = -3\) or \(k = 2\).

Consequently if \(k \neq -3\) and \(k \neq 2\), then \(\det A \neq 0\) and the given system

\[
\begin{align*}
x + y - z &= 1 \\
2x + 3y + kz &= 3 \\
x + ky + 3z &= 2
\end{align*}
\]

has a unique solution. We consider the cases \(k = -3\) and \(k = 2\) separately.

\(k = -3\):

\[
AM = \begin{bmatrix}
1 & 1 & -1 & 1 \\
2 & 3 & -3 & 3 \\
1 & -3 & 3 & 2
\end{bmatrix}
\]

\[
R_2 \rightarrow R_2 - 2R_1 \\
R_3 \rightarrow R_3 - R_1
\]

\[
R_3 \rightarrow R_3 + 4R_2
\]

from which we read off inconsistency.

\(k = 2\):

\[
AM = \begin{bmatrix}
1 & 1 & -1 & 1 \\
2 & 3 & 2 & 3 \\
1 & 2 & 3 & 2
\end{bmatrix}
\]

\[
R_2 \rightarrow R_2 - 2R_1 \\
R_3 \rightarrow R_3 - R_1
\]

\[
R_3 \rightarrow R_3 - R_2
\]

We read off the complete solution \(x = 5z, y = 1 - 4z\), where \(z\) is arbitrary.
Finally we have to determine the solution for which \( x^2 + y^2 + z^2 \) is least.

\[
x^2 + y^2 + z^2 = (5z)^2 + (1 - 4z)^2 + z^2 = 42z^2 - 8z + 1
\]

\[
= 42(z^2 - \frac{4}{21}z + \frac{1}{42}) = 42 \left( z - \frac{2}{21} \right)^2 + \frac{1}{42} - \left( \frac{2}{21} \right)^2
\]

\[
= 42 \left( z - \frac{2}{21} \right)^2 + \frac{13}{882}
\]

We see that the least value of \( x^2 + y^2 + z^2 \) is \( 42 \times \frac{13}{882} = \frac{13}{21} \) and this occurs when \( z = \frac{2}{21} \), with corresponding values \( x = \frac{10}{21} \) and \( y = 1 - 4 \times \frac{2}{21} = \frac{13}{21} \).

9. Let \( \Delta = \begin{vmatrix} 1 & -2 & b \\ a & 0 & 2 \\ 5 & 2 & 0 \end{vmatrix} \) be the coefficient determinant of the given system. Then expanding along column 2 gives

\[
\Delta = 2 \begin{vmatrix} a & 2 \\ 5 & 0 \end{vmatrix} - 2 \begin{vmatrix} 1 & b \\ 5 & 2 \end{vmatrix} = -20 - 2(2 - ab)
\]

\[
= 2ab - 24 = 2(ab - 12).
\]

Hence \( \Delta = 0 \) if and only if \( ab = 12 \). Hence if \( ab \neq 12 \), the given system has a unique solution.

If \( ab = 12 \) we must argue with care:

\[
AM = \begin{bmatrix} 1 & -2 & b \\ a & 0 & 2 \\ 5 & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & b & 3 \\ 0 & 2a & 2 - ab & 2 - 3a \\ 0 & 12 & -5b & -14 \end{bmatrix}
\]

\[
\rightarrow \begin{bmatrix} 1 & -2 & b & 3 \\ 0 & 1 & \frac{-5b}{12} & \frac{-7}{6} \\ 0 & 2a & 2 - ab & 2 - 3a \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & b & 3 \\ 0 & 1 & \frac{-5b}{12} & \frac{-7}{6} \\ 0 & 0 & \frac{-7}{6} - \frac{2a}{3} \end{bmatrix}
\]

\[
= \begin{bmatrix} 1 & -2 & b & 3 \\ 0 & 1 & \frac{-5b}{12} & \frac{-7}{6} \\ 0 & 0 & \frac{-7}{6} - \frac{2a}{3} \end{bmatrix} = B.
\]

Hence if \( 6 - 2a \neq 0 \), i.e. \( a \neq 3 \), the system has no solution.

If \( a = 3 \) (and hence \( b = 4 \)), then

\[
B = \begin{bmatrix} 1 & -2 & 4 & 3 \\ 0 & 1 & \frac{-5}{3} & \frac{-7}{6} \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2/3 & 2/3 \\ 0 & 1 & \frac{-5}{3} & \frac{-7}{6} \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\]
Consequently the complete solution of the system is \( x = \frac{2}{3} + \frac{2}{3}z, \ y = \frac{-7}{6} + \frac{5}{3}z, \) where \( z \) is arbitrary. Hence there are infinitely many solutions.

10. \[
\Delta = \begin{vmatrix}
1 & 1 & 2 & 1 \\
1 & 2 & 3 & 4 \\
2 & 4 & 7 & 2t + 6 \\
2 & 2 & 6 & t - t
\end{vmatrix} = R_4 \rightarrow R_4 - 2R_1 \\
\begin{vmatrix}
1 & 1 & 2 & 1 \\
0 & 1 & 3 \\
0 & 2 - t & t - 2 \\
0 & 2 - t & t - 2
\end{vmatrix} = R_2 \rightarrow R_2 - 2R_1 \\
\begin{vmatrix}
1 & 1 & 3 \\
2 & 3 & 2t + 4 \\
0 & 2 - t & t - 2 \\
0 & 2 - t & t - 2
\end{vmatrix} = R_3 \rightarrow R_3 - 2R_1 \\
\begin{vmatrix}
1 & 1 & 2t - 2 \\
2 & 2t - 2 \\
2 - t & t - 2 \\
2 - t & t - 2
\end{vmatrix} = (t - 2) \begin{vmatrix}
1 & 2t - 2 \\
-1 & 1
\end{vmatrix} = (t - 2)(2t - 1).
\]

Hence \( \Delta = 0 \) if and only if \( t = 2 \) or \( t = \frac{1}{2} \). Consequently the given matrix \( B \) is non–singular if and only if \( t \neq 2 \) and \( t \neq \frac{1}{2} \).

11. Let \( A \) be a \( 3 \times 3 \) matrix with \( \det A \neq 0 \). Then

(i)

\[
A \text{adj} A = (\det A)I_3 \quad (1)
\]

\[
(\det A) \det (\text{adj} A) = \det (\det A \cdot I_3) = (\det A)^3.
\]

Hence, as \( \det A \neq 0 \), dividing out by \( \det A \) in the last equation gives

\[
\det (\text{adj} A) = (\det A)^2.
\]

(ii). Also from equation (1)

\[
\left( \frac{1}{\det A} \right) A \text{adj} A = I_3,
\]

so \( \text{adj} A \) is non–singular and

\[
(\text{adj} A)^{-1} = \frac{1}{\det A} A.
\]

Finally

\[
A^{-1} \text{adj} (A^{-1}) = (\det A^{-1})I_3
\]

and multiplying both sides of the last equation by \( A \) gives

\[
\text{adj} (A^{-1}) = A(\det A^{-1})I_3 = \frac{1}{\det A} A.
\]
12. Let $A$ be a real $3 \times 3$ matrix satisfying $A^tA = I_3$. Then

\[(i) \quad A^t(A - I_3) = A^tA - A^t = I_3 - A^t = -(A^t - I_3) = -(A - I_3)^t.\]

Taking determinants of both sides then gives

\[
\det A^t \det (A - I_3) = \det (-(A - I_3)^t)\\
\det A \det (A - I_3) = (-1)^3 \det (A - I_3)^t = - \det (A - I_3) (1).
\]

(ii) Also $\det AA^t = \det I_3$, so

\[
\det A^t \det A = 1 = (\det A)^2.
\]

Hence $\det A = \pm 1$.

(iii) Suppose that $\det A = 1$. Then equation (1) gives

\[
\det (A - I_3) = - \det (A - I_3),
\]

so $(1 + 1) \det (A - I_3) = 0$ and hence $\det (A - I_3) = 0$.

13. Suppose that column 1 is a linear combination of the remaining columns:

\[A_{*1} = x_2A_{*2} + \cdots + x_nA_{*n}.\]

Then

\[
\det A = \begin{vmatrix}
  x_2a_{12} + \cdots + x_na_{1n} & a_{12} & \cdots & a_{1n} \\
  x_2a_{22} + \cdots + x_na_{2n} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_2a_{n2} + \cdots + x_na_{nn} & a_{n2} & \cdots & a_{nn}
\end{vmatrix}.
\]

Now $\det A$ is unchanged in value if we perform the operation

\[C_1 \rightarrow C_1 - x_2C_2 - \cdots - x_nC_n: \]

\[
\det A = \begin{vmatrix}
  0 & a_{12} & \cdots & a_{1n} \\
  0 & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & a_{n2} & \cdots & a_{nn}
\end{vmatrix} = 0.
\]
Conversely, suppose that \( \det A = 0 \). Then the homogeneous system \( AX = 0 \) has a non-trivial solution \( X = [x_1, \ldots, x_n]^t \). So

\[
x_1 A_{*1} + \cdots + x_n A_{*n} = 0.
\]

Suppose for example that \( x_1 \neq 0 \). Then

\[
A_{*1} = \left( -\frac{x_2}{x_1} \right) + \cdots + \left( -\frac{x_n}{x_1} \right) A_{*n}
\]

and the first column of \( A \) is a linear combination of the remaining columns.

14. Consider the system

\[
\begin{align*}
-2x + 3y - z &= 1 \\
x + 2y - z &= 4 \\
-2x - y + z &= -3
\end{align*}
\]

Let \( \Delta = \left| \begin{array}{ccc}
-2 & 3 & -1 \\
1 & 2 & -1 \\
-2 & -1 & 1
\end{array} \right| = \left| \begin{array}{ccc}
0 & 7 & -3 \\
1 & 2 & -1 \\
0 & 3 & -1
\end{array} \right| = -7 - 3 = -2 \neq 0.
\]

Hence the system has a unique solution which can be calculated using Cramer's rule:

\[
x = \frac{\Delta_1}{\Delta}, \quad y = \frac{\Delta_2}{\Delta}, \quad z = \frac{\Delta_3}{\Delta},
\]

where

\[
\Delta_1 = \left| \begin{array}{ccc}
1 & 3 & -1 \\
4 & 2 & -1 \\
-3 & -1 & 1
\end{array} \right| = -4,
\]

\[
\Delta_2 = \left| \begin{array}{ccc}
-2 & 1 & -1 \\
1 & 4 & -1 \\
-2 & -3 & 1
\end{array} \right| = -6,
\]

\[
\Delta_3 = \left| \begin{array}{ccc}
-2 & 3 & 1 \\
1 & 2 & 4 \\
-2 & -1 & -3
\end{array} \right| = -8.
\]

Hence \( x = -\frac{4}{2} = 2, \ y = -\frac{6}{2} = 3, \ z = -\frac{8}{2} = 4. \)

15. In Remark 4.0.4, take \( A = I_n \). Then we deduce

(a) \( \det E_{ij} = -1; \)

(b) \( \det E_i(t) = t; \)
(c) \( \det E_{ij}(t) = 1. \)

Now suppose that \( B \) is a non–singular \( n \times n \) matrix. Then we know that \( B \) is a product of elementary row matrices:

\[
B = E_1 \cdots E_m.
\]

Consequently we have to prove that

\[
\det E_1 \cdots E_m A = \det E_1 \cdots E_m \det A.
\]

We prove this by induction on \( m \).

First the case \( m = 1 \). We have to prove \( \det E_1 A = \det E_1 \det A \) if \( E_1 \) is an elementary row matrix. This follows from Remark 4.0.4:

(a) \( \det E_{ij} A = - \det A = \det E_{ij} \det A; \)

(b) \( \det E_i(t) A = t \det A = \det E_i(t) \det A; \)

(c) \( \det E_{ij}(t) A = \det A = \det E_{ij}(t) \det A. \)

Let \( m \geq 1 \) and assume the proposition holds for products of \( m \) elementary row matrices. Then

\[
\det E_1 \cdots E_m E_{m+1} A = \det (E_1 \cdots E_m) (E_{m+1} A)
= \det (E_1 \cdots E_m) \det (E_{m+1} A)
= \det (E_1 \cdots E_m) \det E_{m+1} \det A
= \det ((E_1 \cdots E_m) E_{m+1}) \det A
\]

and the induction goes through.

Hence \( \det BA = \det B \det A \) if \( B \) is non–singular.

If \( B \) is singular, problem 26, Chapter 2.7 tells us that \( BA \) is also singular. However singular matrices have zero determinant, so

\[
\det B = 0 \quad \det BA = 0,
\]

so the equation \( \det BA = \det B \det A \) holds trivially in this case.

16.

\[
\begin{vmatrix}
  a + b + c & a + b & a & a \\
  a + b & a + b + c & a & a \\
  a & a & a + b + c & a + b \\
  a & a & a + b & a + b + c
\end{vmatrix}
\]

51
\[
\begin{array}{c|cccc}
R_1 & R_1 - R_2 & c & -c & 0 & 0 \\
R_2 & R_2 - R_3 & b & b + c & -b - c & -b \\
R_3 & R_3 - R_4 & 0 & 0 & c & -c \\
\end{array} = \begin{array}{c|cccc}
a & a & a + b & a + b + c \\
\end{array}
\]

\[
C_2 \rightarrow C_2 + C_1 = \begin{array}{c|cccc}
c & 0 & 0 & 0 \\
b & 2b + c & -b - c & -b \\
a & 2a & a + b & a + b + c \\
\end{array} = c \begin{array}{c|cccc}
2b + c & -b - c & -b \\
0 & c & -c \\
2a & a + b & a + b + c \\
\end{array}
\]

\[
C_3 \rightarrow C_3 + C_2 = \begin{array}{c|cccc}
2b + c & -b - c & -2b - c \\
0 & c & 0 \\
2a & a + b & 2a + 2b + c \\
\end{array} = c^2 \begin{array}{c|cccc}
2b + c & -2b - c \\
0 & 2a & 2a + 2b + c \\
\end{array}
\]

\[
= c^2(2b + c) \begin{array}{c|c}
1 & -1 \\
2a & 2a + 2b + c \\
\end{array} = c^2(2b + c)(4a + 2b + c).
\]

17. Let \(\Delta = \begin{array}{cccc}
1 + u_1 & u_1 & u_1 & u_1 \\
u_2 & 1 + u_2 & u_2 & u_2 \\
u_3 & u_3 & 1 + u_3 & u_3 \\
u_4 & u_4 & u_4 & 1 + u_4 \\
\end{array}\). Then using the operation \(R_1 \rightarrow R_1 + R_2 + R_3 + R_4\), we have

\[
\Delta = \begin{array}{cccc}
t & t & t & t \\
u_2 & 1 + u_2 & u_2 & u_2 \\
u_3 & u_3 & 1 + u_3 & u_3 \\
u_4 & u_4 & u_4 & 1 + u_4 \\
\end{array}
\]

(\text{where } t = 1 + u_1 + u_2 + u_3 + u_4)

\[
= (1 + u_1 + u_2 + u_3 + u_4) \begin{array}{cccc}
1 & 1 & 1 & 1 \\
u_2 & 1 + u_2 & u_2 & u_2 \\
u_3 & u_3 & 1 + u_3 & u_3 \\
u_4 & u_4 & u_4 & 1 + u_4 \\
\end{array}
\]

The last determinant equals

\[
\begin{array}{c|cccc}
C_2 & C_2 - C_1 & 1 & 0 & 0 \\
C_3 & C_3 - C_1 & u_2 & 1 & 0 \\
C_4 & C_4 - C_1 & u_3 & 0 & 1 \\
\end{array} = 1.
\]

52
18. Suppose that \( A^t = -A \), that \( A \in M_{n \times n}(F) \), where \( n \) is odd. Then
\[
\begin{align*}
\det A^t &= \det(-A) \\
\det A &= (-1)^n \det A = - \det A.
\end{align*}
\]
Hence \((1 + 1) \det A = 0\) and consequently \( \det A = 0 \) if \( 1 + 1 \neq 0 \) in \( F \).

19.
\[
\begin{vmatrix}
1 & 1 & 1 & 1 \\
r & 1 & 1 & 1 \\
r & r & 1 & 1 \\
r & r & r & 1
\end{vmatrix} = \begin{vmatrix}
C_4 & C_4 - C_3 \\
C_3 & C_3 - C_2 \\
C_2 & C_2 - C_1 \\
= \begin{vmatrix}
1 & 0 & 0 & 0 \\
r & 1 - r & 0 & 0 \\
r & 0 & 1 - r & 0 \\
r & 0 & 0 & 1 - r
\end{vmatrix}
\end{vmatrix} = (1 - r)^3.
\]

20.
\[
\begin{vmatrix}
1 & a^2 - bc & a^4 \\
1 & b^2 - ca & b^4 \\
1 & c^2 - ab & c^4
\end{vmatrix} = \begin{vmatrix}
R_2 & R_2 - R_1 \\
R_3 & R_3 - R_1 \\
= \begin{vmatrix}
1 & a^2 - bc & a^4 \\
0 & b^2 - ca - a^2 + bc & b^4 - a^4 \\
0 & c^2 - ab - a^2 + bc & c^4 - a^4
\end{vmatrix}
\end{vmatrix}
\]
\[
= \begin{vmatrix}
b^2 - ca - a^2 + bc & b^4 - a^4 \\
c^2 - ab - a^2 + bc & c^4 - a^4
\end{vmatrix}
= \begin{vmatrix}
(b - a)(b + a) + c(b - a) & (b - a)(b + a)(b^2 + a^2) \\
(c - a)(c + a) + b(c - a) & (c - a)(c + a)(c^2 + a^2)
\end{vmatrix}
= \begin{vmatrix}
(b - a)(b + a + c) & (b - a)(b + a)(b^2 + a^2) \\
(c - a)(c + a + b) & (c - a)(c + a)(c^2 + a^2)
\end{vmatrix}
= \begin{vmatrix}
(b - a)(c - a) & (b + a)(b^2 + a^2) \\
c + a + b & (c + a)(c^2 + a^2)
\end{vmatrix}
= \begin{vmatrix}
(b - a)(c - a)(a + b + c) & 1 \ (b + a)(b^2 + a^2) \\
1 & 1 \ (c + a)(c^2 + a^2)
\end{vmatrix}.
\]

Finally
\[
\begin{vmatrix}
1 & (b + a)(b^2 + a^2) \\
1 & (c + a)(c^2 + a^2)
\end{vmatrix} = (c^3 + ac^2 + ca^2 + a^3) - (b^3 + ab^2 + ba^2 + a^3)
\]
\[
= (c^3 - b^3) + a(c^2 - b^2) + a^2(c - b)
= (c - b)(c^2 + cb + b^2 + a(c + b) + a^2)
= (c - b)(c^2 + cb + b^2 + ac + ab + a^2).
\]

53
Section 5.8

1.

(i) \((-3 + i)(14 - 2i)\) = \((-3)(14 - 2i) + i(14 - 2i)\)
= \{(-3)14 - (-3)(2i)\} + i(14) - i(2i)
= \(-42 + 6i + (14i + 2)\) = \(-40 + 20i\).

(ii) \(\frac{2 + 3i}{1 - 4i}\) = \(\frac{(2 + 3i)(1 + 4i)}{(1 - 4i)(1 + 4i)}\)
= \(\frac{(2 + 3i) + (2 + 3i)(4i)}{1^2 + 4^2}\)
= \(-\frac{10 + 11i}{17}\) = \(-\frac{10}{17} + \frac{11}{17}i\).

(iii) \(\frac{(1 + 2i)^2}{1 - i}\) = \(\frac{1 + 4i + (2i)^2}{1 - i}\)
= \(\frac{1 + 4i - 4}{1 - i}\) = \(-3 + 4i\)
= \(\frac{-3 + 4i(1 + i)}{2}\) = \(-\frac{7 + i}{2}\) = \(-\frac{7}{2} + \frac{1}{2}i\).

2. (i)

\(iz + (2 - 10i)z = 3z + 2i\) \(\Leftrightarrow\) \(z(i + 2 - 10i - 3) = 2i\)
\(=\) \(z(-1 - 9i) = 2i\) \(\Leftrightarrow\) \(z = \frac{-2i}{1 + 9i}\)
= \(\frac{-2i(1 - 9i)}{1 + 81}\) = \(-\frac{18 - 2i}{82}\) = \(-\frac{9}{41} - \frac{i}{41}\).

(ii) The coefficient determinant is
\[
\begin{vmatrix}
1 + i & 2 - i \\
1 + 2i & 3 + i
\end{vmatrix}
= (1 + i)(3 + i) - (2 - i)(1 + 2i) = -2 + i \neq 0.
\]
Hence Cramer’s rule applies: there is a unique solution given by

\[
z = \frac{\begin{vmatrix}
-3i & 2 - i \\
2 + 2i & 3 + i
\end{vmatrix}}{\begin{vmatrix}
-2 + i & 2 + 2i \\
1 + 2i & 2 + 2i
\end{vmatrix}} = \frac{-3 - 11i}{-2 + i} = -1 + 5i
\]
\[
w = \frac{\begin{vmatrix}
1 + i & -3i \\
1 + 2i & 2 + 2i
\end{vmatrix}}{\begin{vmatrix}
-2 + i & 2 + 2i \\
1 + 2i & 2 + 2i
\end{vmatrix}} = \frac{-6 + 7i}{-2 + i} = \frac{19 - 8i}{5}.
\]
3.

\[ 1 + (1 + i) + \cdots + (1 + i)^{99} = \frac{(1 + i)^{100} - 1}{(1 + i) - 1} = \frac{(1 + i)^{100} - 1}{i} = -i \{(1 + i)^{100} - 1\}. \]

Now \((1 + i)^2 = 2i\). Hence 

\[ (1 + i)^{100} = (2i)^{50} = 2^{50}i^{50} = 2^{50}(-1)^{25} = -2^{50}. \]

Hence 

\[ -i \{(1 + i)^{100} - 1\} = -i(-2^{50} - 1) = (2^{50} + 1)i. \]

4. (i) Let \(z^2 = -8 - 6i\) and write \(z = x + iy\), where \(x\) and \(y\) are real. Then 

\[ z^2 = x^2 - y^2 + 2xyi = -8 - 6i, \]

so \(x^2 - y^2 = -8\) and \(2xy = -6\). Hence 

\[ y = -3/x, \quad x^2 - \left(\frac{-3}{x}\right)^2 = -8, \]

so \(x^4 + 8x^2 - 9 = 0\). This is a quadratic in \(x^2\). Hence \(x^2 = 1\) or \(-9\) and consequently \(x^2 = 1\). Hence \(x = 1, y = -3\) or \(x = -1\) and \(y = 3\). Hence \(z = 1 - 3i\) or \(z = -1 + 3i\).

(ii) \(z^2 - (3 + i)z + 4 + 3i = 0\) has the solutions \(z = (3 + i \pm d)/2\), where \(d\) is any complex number satisfying 

\[ d^2 = (3 + i)^2 - 4(4 + 3i) = -8 - 6i. \]

Hence by part (i) we can take \(d = 1 - 3i\). Consequently 

\[ z = \frac{3 + i \pm (1 - 3i)}{2} = 2 - i \quad \text{or} \quad 1 + 2i. \]

(i) The number lies in the first quadrant of the complex plane. 

\[ |4 + i| = \sqrt{4^2 + 1^2} = \sqrt{17}. \]

Also \(\text{Arg}(4 + i) = \alpha\), where \(\tan \alpha = 1/4\) and \(0 < \alpha < \pi/2\). Hence \(\alpha = \tan^{-1}(1/4)\).
(ii) The number lies in the third quadrant of the complex plane.

\[ |\frac{-3-i}{2}| = \frac{|-3-i|}{2} \]

\[ = \frac{1}{2} \sqrt{(-3)^2 + (-1)^2} = \frac{1}{2} \sqrt{9 + 1} = \frac{\sqrt{10}}{2}. \]

Also \( \text{Arg} \left( \frac{-3-i}{2} \right) = -\pi + \alpha \), where \( \tan \alpha = \frac{\frac{1}{2}}{\frac{1}{3}} = 1/3 \) and \( 0 < \alpha < \pi/2 \). Hence \( \alpha = \tan^{-1}(1/3) \).

(iii) The number lies in the second quadrant of the complex plane.

\[ |-1+2i| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}. \]

Also \( \text{Arg} (-1+2i) = \pi - \alpha \), where \( \tan \alpha = 2 \) and \( 0 < \alpha < \pi/2 \). Hence \( \alpha = \tan^{-1}2 \).

(iv) The number lies in the second quadrant of the complex plane.

\[ \left| \frac{-1+i\sqrt{3}}{2} \right| = \frac{|-1+i\sqrt{3}|}{2} \]

\[ = \frac{1}{2} \sqrt{(-1)^2 + (\sqrt{3})^2} = \frac{1}{2} \sqrt{1 + 3} = 1. \]

Also \( \text{Arg} \left( \frac{-1+i\sqrt{3}}{2} \right) = \pi - \alpha \), where \( \tan \alpha = \frac{\sqrt{3}}{\frac{1}{2}} = \sqrt{3} \) and \( 0 < \alpha < \pi/2 \). Hence \( \alpha = \pi/3 \).

6. (i) Let \( z = (1+i)(1+\sqrt{3}i)(\sqrt{3} - i) \). Then

\[ |z| = |1+i||1+\sqrt{3}i||\sqrt{3} - i| \]

\[ = \sqrt{1^2 + 1^2} \sqrt{1^2 + (\sqrt{3})^2} \sqrt{(\sqrt{3})^2 + (-1)^2} \]

\[ = \sqrt{2} \sqrt{4} = 2\sqrt{2}. \]

\[ \text{Arg} \ z \equiv \text{Arg} \ (1+i) + \text{Arg} \ (1+\sqrt{3}) + \text{Arg} \ (\sqrt{3} - i) \pmod{2\pi} \]
\[ \equiv \frac{\pi}{4} + \frac{\pi}{3} - \frac{\pi}{6} \equiv \frac{5}{12}. \]

Hence \( \text{Arg } z = \frac{5}{12} \) and the polar decomposition of \( z \) is

\[ z = 4\sqrt{2} \left( \cos \frac{5\pi}{12} + i \sin \frac{5\pi}{12} \right). \]

(ii) Let \( z = \frac{(1+i)^5(1-i\sqrt{3})^5}{(\sqrt{3}+i)^4} \). Then

\[ |z| = \frac{|(1+i)|^5|(1-i\sqrt{3})|^5}{|\sqrt{3}+i|^4} = \frac{(\sqrt{2})^5 2^5}{2^4} = 2^{7/2}. \]

\[ \text{Arg } z \equiv \text{Arg} (1+i)^5 + \text{Arg} (1-i\sqrt{3})^5 - \text{Arg} (\sqrt{3}+i)^4 \pmod{2\pi} \]

\[ \equiv 5\text{Arg} (1+i) + 5\text{Arg} (1-i\sqrt{3}) - 4\text{Arg} (\sqrt{3}+i) \]

\[ \equiv \frac{5\pi}{4} + 5\left( -\frac{\pi}{3} \right) - 4\frac{\pi}{6} \equiv -\frac{13\pi}{12} \equiv \frac{11\pi}{12}. \]

Hence \( \text{Arg } z = \frac{11\pi}{12} \) and the polar decomposition of \( z \) is

\[ z = 2^{7/2} \left( \cos \frac{11\pi}{12} + i \sin \frac{11\pi}{12} \right). \]

7. (i) Let \( z = 2(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}) \) and \( w = 3(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6}) \). (Both of these numbers are already in polar form.)

(a) \( zw = 6(\cos \frac{\pi}{4} + \frac{\pi}{6}) + i \sin \left( \frac{\pi}{4} + \frac{\pi}{6} \right) \)

\[ = 6(\cos \frac{5\pi}{12} + i \sin \frac{5\pi}{12}). \]

(b) \( \frac{z}{w} = \frac{2}{3}(\cos \frac{\pi}{4} - \frac{\pi}{6}) + i \sin \left( \frac{\pi}{4} - \frac{\pi}{6} \right) \)

\[ = \frac{2}{3}(\cos \frac{\pi}{12} + i \sin \frac{\pi}{12}). \]

(c) \( \frac{w}{z} = \frac{3}{2}(\cos \frac{\pi}{6} - \frac{\pi}{4}) + i \sin \left( \frac{\pi}{6} - \frac{\pi}{4} \right) \)

\[ = \frac{3}{2}(\cos \frac{5\pi}{12} + i \sin \frac{5\pi}{12}). \]

(d) \( \frac{w^2}{z^2} = \frac{32}{9}(\cos \frac{5\pi}{12} - \frac{2\pi}{6}) + i \sin \left( \frac{5\pi}{12} - \frac{2\pi}{6} \right) \)

\[ = \frac{32}{9}(\cos \frac{11\pi}{12} + i \sin \frac{11\pi}{12}). \]

(a) \( (1+i)^2 = 2i \), so

\[ (1+i)^{12} = (2i)^6 = 2^6 i^6 = 64(i^2)^3 = 64(-1)^3 = -64. \]
(b) \((1 - i)^2 = -i\), so

\[
\left(\frac{1 - i}{\sqrt{2}}\right)^{-6} = \left(\left(\frac{1 - i}{\sqrt{2}}\right)^2\right)^{-3}
= (-i)^{-3} = \frac{-1}{i^3} = \frac{-1}{i} = \frac{1}{i} = -i.
\]

8. (i) To solve the equation \(z^2 = 1 + \sqrt{3}i\), we write \(1 + \sqrt{3}i\) in modulus–argument form:

\[
1 + \sqrt{3}i = 2\left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3}\right).
\]

Then the solutions are

\[
z_k = \sqrt{2} \left(\cos \left(\frac{\pi}{3} + \frac{2k\pi}{2}\right) + i \sin \left(\frac{\pi}{3} + \frac{2k\pi}{2}\right)\right), \quad k = 0, 1.
\]

Now \(k = 0\) gives the solution

\[
z_0 = \sqrt{2} \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6}\right) = \sqrt{2} \left(\frac{\sqrt{3}}{2} + i \frac{1}{2}\right) = \frac{\sqrt{3}}{\sqrt{2}} + \frac{i}{\sqrt{2}}.
\]

Clearly \(z_1 = -z_0\).

(ii) To solve the equation \(z^4 = i\), we write \(i\) in modulus–argument form:

\[
i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2}.
\]

Then the solutions are

\[
z_k = \cos \left(\frac{\pi}{2} + \frac{2k\pi}{4}\right) + i \sin \left(\frac{\pi}{2} + \frac{2k\pi}{4}\right), \quad k = 0, 1, 2, 3.
\]

Now \(\cos \left(\frac{\pi}{2} + \frac{2k\pi}{4}\right) = \cos \left(\frac{\pi}{8} + \frac{k\pi}{2}\right)\), so

\[
z_k = \cos \left(\frac{\pi}{8} + \frac{k\pi}{2}\right) + i \sin \left(\frac{\pi}{8} + \frac{k\pi}{2}\right)
= \left(\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}\right)^k \cos \frac{\pi}{8} + i \sin \frac{\pi}{8}
= i^k \left(\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}\right).
\]
Geometrically, the solutions lie equi–spaced on the unit circle at arguments
\[
\frac{\pi}{8}, \frac{\pi}{8} + \frac{\pi}{2} = \frac{5\pi}{8}, \frac{\pi}{8} + \pi = \frac{9\pi}{8}, \frac{\pi}{8} + 3\frac{\pi}{2} = \frac{13\pi}{8}.
\]
Also \(z_2 = -z_0\) and \(z_3 = -z_1\).

(iii) To solve the equation \(z^3 = -8i\), we rewrite the equation as
\[
\left(\frac{z}{-2i}\right)^3 = 1.
\]
Then
\[
\left(\frac{z}{-2i}\right) = 1, \quad \frac{-1 + \sqrt{3}i}{2}, \quad \text{or} \quad \frac{-1 - \sqrt{3}i}{2}.
\]
Hence \(z = -2i, \sqrt{3} + i\) or \(-\sqrt{3} + i\).
Geometrically, the solutions lie equi–spaced on the circle \(|z| = 2\), at arguments
\[
\frac{\pi}{6}, \frac{\pi}{6} + \frac{2\pi}{3} = \frac{5\pi}{6}, \frac{\pi}{6} + 2\frac{2\pi}{3} = \frac{3\pi}{2}.
\]

(iv) To solve \(z^4 = 2 - 2i\), we write \(2 - 2i\) in modulus–argument form:
\[
2 - 2i = 2^{3/2} \left(\cos \frac{-\pi}{4} + i \sin \frac{-\pi}{4}\right).
\]
Hence the solutions are
\[
z_k = 2^{3/8} \cos \left(\frac{-\pi}{4} + 2k\pi\right) + i \sin \left(\frac{-\pi}{4} + 2k\pi\right), \quad k = 0, 1, 2, 3.
\]
We see the solutions can also be written as
\[
z_k = 2^{3/8} i^k \left(\cos \frac{-\pi}{16} + i \sin \frac{-\pi}{16}\right)
\]
\[
= 2^{3/8} i^k \left(\cos \frac{\pi}{16} - i \sin \frac{\pi}{16}\right).
\]
Geometrically, the solutions lie equi–spaced on the circle \(|z| = 2^{3/8}\), at arguments
\[
\frac{-\pi}{16}, \frac{-\pi}{16} + \frac{\pi}{2} = \frac{7\pi}{16}, \frac{-\pi}{16} + 2\frac{\pi}{2} = \frac{15\pi}{16}, \frac{-\pi}{16} + 3\frac{\pi}{2} = \frac{23\pi}{16}.
\]
Also \(z_2 = -z_0\) and \(z_3 = -z_1\).
9.

\[
\begin{bmatrix}
2 + i & -1 + 2i & 2 \\
1 + i & -1 + i & 1 \\
1 + 2i & -2 + i & 1 + i
\end{bmatrix}
\begin{array}{c}
R_1 \to R_1 - R_2 \\
R_3 \to R_3 - R_2
\end{array}
\begin{bmatrix}
1 & i & 1 \\
1 + i & -1 + i & 1 \\
i & -1 & i
\end{bmatrix}
\]
\[
R_2 \to R_2 - (1 + i)R_1 \\
R_3 \to R_3 - iR_1
\begin{bmatrix}
1 & i & 1 \\
0 & 0 & -i \\
0 & 0 & 0
\end{bmatrix}
\begin{array}{c}
R_2 \to iR_2 \\
R_3 \to R_3 - R_2
\end{array}
\begin{bmatrix}
1 & i & 1 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}
\]

The last matrix is in reduced row–echelon form.

10. (i) Let \( p = l + im \) and \( z = x + iy \). Then

\[
\overline{pz} + p\overline{z} = (l - im)(x + iy) + (l + im)(x - iy)
\]
\[
= (lx + liy - imx + my) + (lx - liy + imx + my)
\]
\[
= 2(lx + my).
\]

Hence \( \overline{pz} + p\overline{z} = 2n \iff lx + my = n. \)

(ii) Let \( w \) be the complex number which results from reflecting the complex number \( z \) in the line \( lx + my = n \). Then because \( p \) is perpendicular to the given line, we have

\[
w - z = tp, \quad t \in \mathbb{R}. \quad (a)
\]

Also the midpoint \( \frac{w + z}{2} \) of the segment joining \( w \) and \( z \) lies on the given line, so

\[
\overline{p} \left( \frac{w + z}{2} \right) + p \left( \frac{w + z}{2} \right) = n,
\]
\[
\overline{p} \left( \frac{w + z}{2} \right) + p \left( \frac{\overline{w} + \overline{z}}{2} \right) = n. \quad (b)
\]

Taking conjugates of equation (a) gives

\[
\overline{w} - \overline{z} = t\overline{p}. \quad (c)
\]

Then substituting in (b), using (a) and (c), gives

\[
\overline{p} \left( \frac{2w - tp}{2} \right) + p \left( \frac{2\overline{w} + t\overline{p}}{2} \right) = n
\]
and hence
\[ p w + p \overline{z} = n. \]

(iii) Let \( p = b - a \) and \( n = |b|^2 - |a|^2 \). Then
\[
|z - a| = |z - b| \iff |z - a|^2 = |z - b|^2
\]
\[
\iff (z - a)(\overline{z} - a) = (z - b)(\overline{z} - b)
\]
\[
\iff z \overline{z} - a \overline{z} - z a + \overline{a} a = z \overline{z} - b \overline{z} - z b + \overline{b} b
\]
\[
\iff (b - \overline{a})z + (b - a) \overline{z} = |b|^2 - |a|^2
\]
\[
\iff p w + p \overline{z} = n.
\]

Suppose \( z \) lies on the circle \( \left| \frac{z - a}{z - b} \right| \) and let \( w \) be the reflection of \( z \) in the line \( p w + p \overline{z} = n \). Then by part (ii)
\[ p w + p \overline{z} = n. \]

Taking conjugates gives \( p \overline{w} + p \overline{z} = n \) and hence
\[ z = \frac{n - p \overline{w}}{\overline{p}} \quad (a) \]

Substituting for \( z \) in the circle equation, using (a) gives
\[
\lambda = \left| \frac{n - p \overline{w} - a}{n - p \overline{w} - b} \right| = \left| \frac{n - \overline{p} \overline{w} - \overline{p} a}{n - \overline{p} \overline{w} - \overline{p} b} \right|. \quad (b)
\]

However
\[
n - p a = |b|^2 - |a|^2 - (\overline{b} - \overline{a}) a
\]
\[
= \overline{b} b - \overline{a} a - \overline{b} a + \overline{a} a
\]
\[
= b b - \overline{a} a = \overline{b} p.
\]

Similarly \( n - p \overline{b} = \overline{a} p \). Consequently (b) simplifies to
\[
\lambda = \frac{\overline{b} p - p \overline{w}}{\overline{a} p - p \overline{w}} = \frac{\overline{b} - \overline{w}}{\overline{a} - \overline{w}} = \frac{|w - b|}{|w - a|},
\]

which gives \( \left| \frac{w - a}{w - b} \right| = \frac{1}{\lambda} \).
11. Let $a$ and $b$ be distinct complex numbers and $0 < \alpha < \pi$.

(i) When $z_1$ lies on the circular arc shown, it subtends a constant angle $\alpha$. This angle is given by $\text{Arg} \left( \frac{z_1 - a}{z_1 - b} \right)$. However

$$\text{Arg} \left( \frac{z_1 - a}{z_1 - b} \right) = \text{Arg} \left( \frac{z_1 - a}{z_1 - b} \right) + 2k\pi = \alpha + 2k\pi.$$ 

It follows that $k = 0$, as $0 < \alpha < \pi$ and $-\pi < \text{Arg} \theta \leq \pi$. Hence

$$\text{Arg} \left( \frac{z_1 - a}{z_1 - b} \right) = \alpha.$$ 

Similarly if $z_2$ lies on the circular arc shown, then

$$\text{Arg} \left( \frac{z_2 - a}{z_2 - b} \right) = -\alpha = -\left( \pi - \alpha \right) = \alpha - \pi.$$ 

Replacing $\alpha$ by $\pi - \alpha$, we deduce that if $z_4$ lies on the circular arc shown, then

$$\text{Arg} \left( \frac{z_4 - a}{z_4 - b} \right) = \pi - \alpha,$$

while if $z_3$ lies on the circular arc shown, then

$$\text{Arg} \left( \frac{z_3 - a}{z_3 - b} \right) = -\alpha.$$ 

The straight line through $a$ and $b$ has the equation

$$z = (1 - t)a + tb,$$
where \( t \) is real. Then \( 0 < t < 1 \) describes the segment \( ab \). Also
\[
\frac{z - a}{z - b} = \frac{t}{t - 1}.
\]
Hence \( \frac{z - a}{z - b} \) is real and negative if \( z \) is on the segment \( a \), but is real and positive if \( z \) is on the remaining part of the line, with corresponding values
\[
\text{Arg} \left( \frac{z - a}{z - b} \right) = \pi, \ 0,
\]
respectively.

(ii) Case (a) Suppose \( z_1, \ z_2 \) and \( z_3 \) are not collinear. Then these points determine a circle. Now \( z_1 \) and \( z_2 \) partition this circle into two arcs. If \( z_3 \) and \( z_4 \) lie on the same arc, then
\[
\text{Arg} \left( \frac{z_3 - z_1}{z_3 - z_2} \right) = \text{Arg} \left( \frac{z_4 - z_1}{z_4 - z_2} \right);
\]
whereas if \( z_3 \) and \( z_4 \) lie on opposite arcs, then
\[
\text{Arg} \left( \frac{z_3 - z_1}{z_3 - z_2} \right) = \alpha
\]
and
\[
\text{Arg} \left( \frac{z_4 - z_1}{z_4 - z_2} \right) = \alpha - \pi.
\]
Hence in both cases
\[
\text{Arg} \left( \frac{z_3 - z_1}{z_3 - z_2} \right) \equiv \text{Arg} \left( \frac{z_3 - z_1}{z_3 - z_2} \right) - \text{Arg} \left( \frac{z_4 - z_1}{z_4 - z_2} \right) \pmod{2\pi}
\]
\[
\equiv 0 \text{ or } \pi.
\]
In other words, the cross–ratio
\[
\frac{z_3 - z_1}{z_3 - z_2} / \frac{z_4 - z_1}{z_4 - z_2}
\]
is real.
(b) If \( z_1, \ z_2 \) and \( z_3 \) are collinear, then again the cross–ratio is real.
The argument is reversible.

(iii) Assume that \( A, B, C, D \) are distinct points such that the cross–ratio
\[
r = \frac{z_3 - z_1}{z_3 - z_2} / \frac{z_4 - z_1}{z_4 - z_2}
\]
is real. Now \( r \) cannot be 0 or 1. Then there are three cases:
(i) $0 < r < 1$;
(ii) $r < 0$;
(iii) $r > 1$.

Case (i). Here $|r| + |1 - r| = 1$. So

$$\left| \frac{z_4 - z_1}{z_4 - z_2} \cdot \frac{z_3 - z_2}{z_3 - z_1} \right| + \left| 1 - \left( \frac{z_4 - z_1}{z_4 - z_2} \cdot \frac{z_3 - z_2}{z_3 - z_1} \right) \right| = 1.$$ 

Multiplying both sides by the denominator $|z_4 - z_2||z_3 - z_1|$ gives after simplification

$$|z_4 - z_1||z_3 - z_2| + |z_2 - z_1||z_4 - z_3| = |z_4 - z_2||z_3 - z_1|,$$

or

(a) \[ AD \cdot BC + AB \cdot CD = BD \cdot AC. \]

Case (ii). Here $1 + |r| = |1 - r|$. This leads to the equation

(b) \[ BD \cdot AC + AD \cdot BC = AB \cdot CD. \]

Case (iii). Here $1 + |1 - r| = |r|$. This leads to the equation

(c) \[ BD \cdot AC + AB \cdot CD = AD \cdot BC. \]

Conversely if (a), (b) or (c) hold, then we can reverse the argument to deduce that $r$ is a complex number satisfying one of the equations

$$|r| + |1 - r| = 1, \quad 1 + |r| = |1 - r|, \quad 1 + |1 - r| = |r|,$$

from which we deduce that $r$ is real.
1. Let $A = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$. Then $A$ has characteristic equation $\lambda^2 - 4\lambda + 3 = 0$ or $(\lambda - 3)(\lambda - 1) = 0$. Hence the eigenvalues of $A$ are $\lambda_1 = 3$ and $\lambda_2 = 1$.

$\lambda_1 = 3$. The corresponding eigenvectors satisfy $(A - \lambda_1 I_2)X = 0$, or

$$\begin{bmatrix} 1 & -3 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

or equivalently $x - 3y = 0$. Hence

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3y \\ y \end{bmatrix} = y \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

and we take $X_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

Similarly for $\lambda_2 = 1$ we find the eigenvector $X_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Hence if $P = [X_1 | X_2] = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$, then $P$ is non–singular and

$$P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hence

$$A = P \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} P^{-1}$$

and consequently

$$A^n = P \begin{bmatrix} 3^n & 0 \\ 0 & 1^n \end{bmatrix} P^{-1}$$

$$= \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3^n & 0 \\ 0 & 1^n \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 3^{n+1} & 1 \\ 3^n & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 3^{n+1} - 1 & -3^{n+1} + 3 \\ 3^n - 1 & -3^n + 3 \end{bmatrix}$$

$$= \frac{3^n - 1}{2} A + \frac{3 - 3^n}{2} I_2.$$
2. Let \( A = \begin{bmatrix} 3/5 & 4/5 \\ 2/5 & 1/5 \end{bmatrix} \). Then we find that the eigenvalues are \( \lambda_1 = 1 \) and \( \lambda_2 = -1/5 \), with corresponding eigenvectors

\[
X_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.
\]

Then if \( P = [X_1|X_2] \), \( P \) is non–singular and

\[
P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & -1/5 \end{bmatrix} \quad \text{and} \quad A = P \begin{bmatrix} 1 & 0 \\ 0 & -1/5 \end{bmatrix} P^{-1}.
\]

Hence

\[
A^n = P \begin{bmatrix} 1 & 0 \\ 0 & (-1/5)^n \end{bmatrix} P^{-1} \\
\rightarrow P \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} P^{-1} \\
= \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3/5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \\
= \begin{bmatrix} 1 & 2 & 0 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \\
= \begin{bmatrix} 1 & 2 & 2 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix}.
\]

3. The given system of differential equations is equivalent to \( \dot{X} = AX \), where

\[
A = \begin{bmatrix} 3 & -2 \\ 5 & -4 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} x \\ y \end{bmatrix}.
\]

The matrix \( P = \begin{bmatrix} 2 & 1 \\ 5 & 1 \end{bmatrix} \) is a non-singular matrix of eigenvectors corresponding to eigenvalues \( \lambda_1 = -2 \) and \( \lambda_2 = 1 \). Then

\[
P^{-1}AP = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}.
\]

The substitution \( X = PY \), where \( Y = [x_1, y_1]^t \), gives

\[
\dot{Y} = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} Y,
\]

66
or equivalently $x_1 = -2x_1$ and $y_1 = y_1$.

Hence $x_1 = x_1(0)e^{-2t}$ and $y_1 = y_1(0)e^t$. To determine $x_1(0)$ and $y_1(0)$, we note that

$$
\begin{bmatrix}
    x_1(0) \\
y_1(0)
\end{bmatrix}
= P^{-1}
\begin{bmatrix}
x(0) \\
y(0)
\end{bmatrix}
= \frac{1}{3}
\begin{bmatrix}
    1 & -1 \\
    -5 & 2
\end{bmatrix}
\begin{bmatrix}
    13 \\
    22
\end{bmatrix}
= \begin{bmatrix}
    3 \\
    7
\end{bmatrix}.
$$

Hence $x_1 = 3e^{-2t}$ and $y_1 = 7e^t$. Consequently

$$x = 2x_1 + y_1 = 6e^{-2t} + 7e^t \quad \text{and} \quad y = 5x_1 + y_1 = 15e^{-2t} + 7e^t.$$

4. Introducing the vector $X_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$, the system of recurrence relations

$$
x_{n+1} = 3x_n - y_n \\
y_{n+1} = -x_n + 3y_n,
$$

becomes $X_{n+1} = AX_n$, where $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$. Hence $X_n = A^nX_0$, where $X_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

To find $A^n$ we can use the eigenvalue method. We get

$$A^n = \frac{1}{2}
\begin{bmatrix}
    2^n + 4^n & 2^n - 4^n \\
    2^n - 4^n & 2^n + 4^n
\end{bmatrix}.$$

Hence

$$
X_n = \frac{1}{2}
\begin{bmatrix}
    2^n + 4^n & 2^n - 4^n \\
    2^n - 4^n & 2^n + 4^n
\end{bmatrix}
\begin{bmatrix}
    1 \\
    2
\end{bmatrix}
= \frac{1}{2}
\begin{bmatrix}
    2^n + 4^n + 2(2^n - 4^n) \\
    2^n - 4^n + 2(2^n + 4^n)
\end{bmatrix}
= \frac{1}{2}
\begin{bmatrix}
    3 \times 2^n - 4^n \\
    3 \times 2^n + 4^n
\end{bmatrix}
= \begin{bmatrix}
    (3 \times 2^n - 4^n)/2 \\
    (3 \times 2^n + 4^n)/2
\end{bmatrix}.
$$

Hence $x_n = \frac{1}{2}(3 \times 2^n - 4^n)$ and $y_n = \frac{1}{2}(3 \times 2^n + 4^n)$.

5. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real or complex matrix with distinct eigenvalues $\lambda_1, \lambda_2$ and corresponding eigenvectors $X_1, X_2$. Also let $P = [X_1|X_2]$.

(a) The system of recurrence relations

$$
x_{n+1} = ax_n + by_n \\
y_{n+1} = cx_n + dy_n
$$

67
has the solution

\[
\begin{bmatrix}
x_n \\
y_n
\end{bmatrix}
= A^n
\begin{bmatrix}
x_0 \\
y_0
\end{bmatrix}
= \left(P
\begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}
P^{-1}\right)^n
\begin{bmatrix}
x_0 \\
y_0
\end{bmatrix}
= P
\begin{bmatrix}
\lambda_1^n & 0 \\
0 & \lambda_2^n
\end{bmatrix}
P^{-1}
\begin{bmatrix}
x_0 \\
y_0
\end{bmatrix}
= [X_1|X_2]
\begin{bmatrix}
\lambda_1^n & 0 \\
0 & \lambda_2^n
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}
= [X_1|X_2]
\begin{bmatrix}
\lambda_1^n\alpha \\
\lambda_2^n\beta
\end{bmatrix}
= \lambda_1^n\alpha X_1 + \lambda_2^n\beta X_2,
\]

where

\[
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}
= P^{-1}
\begin{bmatrix}
x_0 \\
y_0
\end{bmatrix}.
\]

(b) In matrix form, the system is \(\dot{X} = AX\), where \(X = \begin{bmatrix} x \\ y \end{bmatrix}\). We substitute \(X = PY\), where \(Y = [x_1, y_1]^t\). Then

\(\dot{X} = P\dot{Y} = AX = A(PY)\),

so

\(\dot{Y} = (P^{-1}AP)Y = \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}
\begin{bmatrix}
x_1 \\
y_1
\end{bmatrix}.
\)

Hence \(x_1 = \lambda_1 x_1\) and \(y_1 = \lambda_2 y_1\). Then

\(x_1 = x_1(0)e^{\lambda_1 t}\) and \(y_1 = y_1(0)e^{\lambda_2 t}\).

But

\[
\begin{bmatrix}
x(0) \\
y(0)
\end{bmatrix}
= P
\begin{bmatrix}
x_1(0) \\
y_1(0)
\end{bmatrix},
\]

so

\[
\begin{bmatrix}
x_1(0) \\
y_1(0)
\end{bmatrix}
= P^{-1}
\begin{bmatrix}
x(0) \\
y(0)
\end{bmatrix}
= \begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}.
\]

Consequently \(x_1(0) = \alpha\) and \(y_1(0) = \beta\) and

\[
\begin{bmatrix}
x \\
y
\end{bmatrix}
= P
\begin{bmatrix}
x_1 \\
y_1
\end{bmatrix}
= [X_1|X_2]
\begin{bmatrix}
\alpha e^{\lambda_1 t} \\
\beta e^{\lambda_2 t}
\end{bmatrix}
= \alpha e^{\lambda_1 t}X_1 + \beta e^{\lambda_2 t}X_2.
\]
6. Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) be a real matrix with non–real eigenvalues \( \lambda = a + ib \) and \( \bar{\lambda} = a - ib \), with corresponding eigenvectors \( X = U + iV \) and \( \bar{X} = U - iV \), where \( U \) and \( V \) are real vectors. Also let \( P \) be the real matrix defined by \( P = [U | V] \). Finally let \( a + ib = re^{i\theta}, \) where \( r > 0 \) and \( \theta \) is real.

(a) As \( X \) is an eigenvector corresponding to the eigenvalue \( \lambda \), we have \( AX = \lambda X \) and hence
\[
A(U + iV) = (a + ib)(U + iV) \\
AU + iAV = aU - bV + i(bU + aV).
\]
Equating real and imaginary parts then gives
\[
AU = aU - bV \\
AV = bU + aV.
\]

(b)
\[
AP = A[U|V] = [AU|AV] = [aU - bV|bU + aV] = [U|V] \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = P \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.
\]
Hence, as \( P \) can be shown to be non–singular,
\[
P^{-1}AP = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.
\]
(The fact that \( P \) is non–singular is easily proved by showing the columns of \( P \) are linearly independent: Assume \( xU + yV = 0 \), where \( x \) and \( y \) are real. Then we find
\[
(x + iy)(U - iV) + (x - iy)(U + iV) = 0.
\]
Consequently \( x + iy = 0 \) as \( U - iV \) and \( U + iV \) are eigenvectors corresponding to distinct eigenvalues \( a - ib \) and \( a + ib \) and are hence linearly independent. Hence \( x = 0 \) and \( y = 0 \).)

(c) The system of recurrence relations
\[
x_{n+1} = ax_n + by_n \\
y_{n+1} = cx_n + dy_n
\]
has solution
\[
\begin{bmatrix}
x_n \\
y_n
\end{bmatrix} = A^n \begin{bmatrix}
x_0 \\
y_0
\end{bmatrix} = P \left[ \begin{array}{cc}
a & b \\-b & a
\end{array} \right]^n P^{-1} \begin{bmatrix}
x_0 \\
y_0
\end{bmatrix} = P \left[ \begin{array}{cc}
r \cos \theta & r \sin \theta \\-r \sin \theta & r \cos \theta
\end{array} \right]^n \begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = P r^n \left[ \begin{array}{cc}
\cos \theta & \sin \theta \\-\sin \theta & \cos \theta
\end{array} \right]^n \begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = r^n \begin{bmatrix}
\cos n\theta & \sin n\theta \\-\sin n\theta & \cos n\theta
\end{bmatrix} \begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} = r^n \left[ \begin{array}{c}
\alpha \cos n\theta + \beta \sin n\theta U + (-\alpha \sin n\theta + \beta \cos n\theta)V \\
-\alpha \sin n\theta + \beta \cos n\theta
\end{array} \right] = r^n \{(\alpha \cos n\theta)(\alpha U + \beta V) + (\sin n\theta)(\beta U - \alpha V)\}.
\]

(d) The system of differential equations
\[
\begin{align*}
\frac{dx}{dt} &= ax + by \\
\frac{dy}{dt} &= cx + dy
\end{align*}
\]
is attacked using the substitution \(X = PY\), where \(Y = [x_1, y_1]^t\). Then
\[
\dot{Y} = (P^{-1}AP)Y,
\]
so
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{y}_1
\end{bmatrix} = \begin{bmatrix}
a & b \\-b & a
\end{bmatrix} \begin{bmatrix}
x_1 \\
y_1
\end{bmatrix}.
\]
Equating components gives
\[
\begin{align*}
\dot{x}_1 &= ax_1 + by_1 \\
\dot{y}_1 &= -bx_1 + ay_1.
\end{align*}
\]
Now let \(z = x_1 + iy_1\). Then
\[
\dot{z} = \dot{x}_1 + iy_1 = (ax_1 + by_1) + i(-bx_1 + ay_1) = (a - ib)(x_1 + iy_1) = (a - ib)z.
\]
Hence
\[ z = z(0)e^{(a - ib)t} \]
\[ x_1 + iy_1 = (x_1(0) + iy_1(0))e^{at}(\cos bt - i \sin bt). \]

Equating real and imaginary parts gives
\[ x_1 = e^{at}\{x_1(0) \cos bt + y_1(0) \sin bt\} \]
\[ y_1 = e^{at}\{y_1(0) \cos bt - x_1(0) \sin bt\}. \]

Now if we define \( \alpha \) and \( \beta \) by
\[
\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix},
\]
we see that \( \alpha = x_1(0) \) and \( \beta = y_1(0) \). Then
\[
\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \left[U | V\right] \begin{bmatrix} \alpha \cos bt + \beta \sin bt \\ \alpha \cos bt - \beta \sin bt \end{bmatrix} = e^{at}\{(\alpha \cos bt + \beta \sin bt)U + (\beta \cos bt - \alpha \sin bt)V\} = e^{at}\{\cos bt(\alpha U + \beta V) + \sin bt(\beta U - \alpha V)\}.
\]

7. (The case of repeated eigenvalues.) Let \( A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) and suppose that the characteristic polynomial of \( A, \lambda^2 - (a + d)\lambda + (ad - bc) \), has a repeated root \( \alpha \). Also assume that \( A \neq \alpha I_2 \).

(i)
\[
\lambda^2 - (a + d)\lambda + (ad - bc) = (\lambda - \alpha)^2 = \lambda^2 - 2\alpha \lambda + \alpha^2.
\]

Hence \( a + d = 2\alpha \) and \( ad - bc = \alpha^2 \) and
\[
(a + d)^2 = 4(ad - bc), \\
a^2 + 2ad + d^2 = 4ad - 4bc, \\
a^2 - 2ad + d^2 + 4bc = 0, \\
(a - d)^2 + 4bc = 0.
\]

71
(ii) Let $B - A - \alpha I$. Then

$$B^2 = (A - \alpha I)^2 = A^2 - 2\alpha A + \alpha^2 I_2$$

$$= A^2 - (a + d)A + (ad - bc)I_2,$$

But by problem 3, chapter 2.4, $A^2 - (a + d)A + (ad - bc)I_2 = 0$, so $B^2 = 0$.

(iii) Now suppose that $B \neq 0$. Then $BE_1 \neq 0$ or $BE_2 \neq 0$, as $BE_i$ is the $i$–th column of $B$. Hence $BX_i \neq 0$, where $X_2 = E_1$ or $X_2 = E_2$.

(iv) Let $X_1 = BX_2$ and $P = [X_1|X_2]$. We prove $P$ is non–singular by demonstrating that $X_1$ and $X_2$ are linearly independent.

Assume $xX_1 + yX_2 = 0$. Then

$$xBX_2 + yX_2 = 0$$

$$B(xBX_2 + yX_2) = B0 = 0$$

$$xB^2X_2 + yBX_2 = 0$$

$$x0X_2 + yBX_2 = 0$$

$$yBX_2 = 0.$$

Hence $y = 0$ as $BX_2 \neq 0$. Hence $xBX_2 = 0$ and so $x = 0$.

Finally, $BX_1 = B(BX_2) = B^2X_2 = 0$, so $(A - \alpha I)X_1 = 0$ and

$$AX_1 = \alpha X_1. \quad (2)$$

Also

$$X_1 = BX_2 = (A - \alpha I)X_2 = AX_2 - \alpha X_2.$$

Hence

$$AX_2 = X_1 + \alpha X_2. \quad (3)$$

Then, using (2) and (3), we have

$$AP = A[X_1|X_2] = [AX_1|AX_2]$$

$$= [\alpha X_1 | X_1 + \alpha X_2]$$

$$= [X_1 | X_2] \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix}.$$
\[ P^{-1}AP = \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix}. \]

8. The system of differential equations is equivalent to the single matrix equation \( \dot{X} = AX \), where \( A = \begin{bmatrix} 4 & -1 \\ 4 & 8 \end{bmatrix} \).

The characteristic polynomial of \( A \) is \( \lambda^2 - 12\lambda + 36 = (\lambda - 6)^2 \), so we can use the previous question with \( \alpha = 6 \). Let

\[ B = A - 6I_2 = \begin{bmatrix} -2 & -1 \\ 4 & 2 \end{bmatrix}. \]

Then \( BX_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \), if \( X_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \). Also let \( X_1 = BX_2 \). Then if \( P = [X_1|X_2] \), we have \( P^{-1}AP = \begin{bmatrix} 6 & 1 \\ 0 & 6 \end{bmatrix} \).

Now make the change of variables \( X = PY \), where \( Y = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \). Then

\[ \dot{Y} = (P^{-1}AP)Y = \begin{bmatrix} 6 & 1 \\ 0 & 6 \end{bmatrix}Y, \]

or equivalently \( \dot{x}_1 = 6x_1 + y_1 \) and \( \dot{y}_1 = 6y_1 \).

Solving for \( y_1 \) gives \( y_1 = y_1(0)e^{6t} \). Consequently \( \dot{x}_1 = 6x_1 + y_1(0)e^{6t} \).

Multiplying both side of this equation by \( e^{-6t} \) gives

\[ \frac{d}{dt}(e^{-6t}x_1) = e^{-6t}\dot{x}_1 - 6e^{-6t}x_1 = y_1(0) \]

\[ e^{-6t}x_1 = y_1(0)t + c, \]

where \( c \) is a constant. Substituting \( t = 0 \) gives \( c = x_1(0) \). Hence \( e^{-6t}x_1 = y_1(0)t + x_1(0) \)

and hence \( x_1 = e^{6t}(y_1(0)t + x_1(0)) \).
However, since we are assuming $x(0) = 1 = y(0)$, we have

\[
\begin{bmatrix}
    x_1(0) \\
y_1(0)
\end{bmatrix} = P^{-1} \begin{bmatrix}
    x(0) \\
y(0)
\end{bmatrix} = \frac{1}{-4} \begin{bmatrix}
    0 & -1 \\
    -4 & -2
\end{bmatrix} \begin{bmatrix}
    1 \\
    1
\end{bmatrix} = \frac{1}{-4} [1 \ -1] = [1/4]
\]

Hence $x_1 = e^{6t}(\frac{3}{2}t + \frac{1}{4})$ and $y_1 = \frac{3}{2}e^{6t}$.

Finally, solving for $x$ and $y$,

\[
\begin{bmatrix}
x \\
y
\end{bmatrix} = \begin{bmatrix}
    -2 & 1 \\
    4 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
y_1
\end{bmatrix} = \begin{bmatrix}
    -2 & 1 \\
    4 & 0
\end{bmatrix} \begin{bmatrix}
e^{6t}(\frac{3}{2}t + \frac{1}{4}) \\
e^{6t}/2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    (-2)e^{6t}(\frac{3}{2}t + \frac{1}{4}) + \frac{3}{2}e^{6t} \\
4e^{6t}(\frac{3}{2}t + \frac{1}{4})
\end{bmatrix}
\]

\[
= \begin{bmatrix}
e^{6t}(1 - 3t) \\
e^{6t}(6t + 1)
\end{bmatrix}.
\]

Hence $x = e^{6t}(1 - 3t)$ and $y = e^{6t}(6t + 1)$.

9. Let

\[
A = \begin{bmatrix}
    1/2 & 1/2 & 0 \\
    1/4 & 1/4 & 1/2 \\
    1/4 & 1/4 & 1/2
\end{bmatrix}.
\]

(a) We first determine the characteristic polynomial $\text{ch}_A(\lambda)$.

\[
\text{ch}_A(\lambda) = \det (\lambda I_3 - A) = \begin{vmatrix}
    \lambda - 1/2 & -1/2 & 0 \\
    1/4 & \lambda - 1/4 & -1/2 \\
    -1/4 & -1/4 & \lambda - 1/2
\end{vmatrix}
\]

\[
= \begin{vmatrix}
    \lambda - 1/2 & \lambda - 1/4 & -1/2 \\
    1/4 & \lambda - 1/2 & -1/2 \\
    -1/4 & -1/4 & \lambda - 1/2
\end{vmatrix}
\]

\[
= \lambda \begin{vmatrix}
    \lambda - 1/2 & \lambda - 1/2 \\
    1/4 & 1/2
\end{vmatrix} - \frac{\lambda}{8}
\]

\[
= \lambda \left\{ \left( \lambda - \frac{1}{2} \right) \left( \lambda - \frac{3}{4} \right) - \frac{1}{8} \right\}
\]

\[
= \lambda \left\{ \left( \lambda - \frac{1}{2} \right) \left( \lambda - 3 \right) - \frac{1}{8} \right\}
\]

74
\[ \lambda \left( \lambda^2 - \frac{5\lambda}{4} + \frac{1}{4} \right) \]

\[ = \lambda(\lambda - 1) \left( \lambda - \frac{1}{4} \right). \]

(b) Hence the characteristic polynomial has no repeated roots and we can use Theorem 6.2.2 to find a non–singular matrix \( P \) such that

\[ P^{-1}AP = \text{diag}(1, 0, \frac{1}{4}). \]

We take \( P = [X_1|X_2|X_3] \), where \( X_1, X_2, X_3 \) are eigenvectors corresponding to the respective eigenvalues 1, 0, \( \frac{1}{4} \).

Finding \( X_1 \): We have to solve \((A - I_3)X = 0\). We have

\[ A - I_3 = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 1/4 & -3/4 & 1/2 \\ 1/4 & 1/4 & -1/2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \]

Hence the eigenspace consists of vectors \( X = [x, y, z]^t \) satisfying \( x = z \) and \( y = z \), with \( z \) arbitrary. Hence

\[ X = \begin{bmatrix} z \\ z \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \]

and we can take \( X_1 = [1, 1, 1]^t \).

Finding \( X_2 \): We solve \( AX = 0 \). We have

\[ A = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \]

Hence the eigenspace consists of vectors \( X = [x, y, z]^t \) satisfying \( x = -y \) and \( z = 0 \), with \( y \) arbitrary. Hence

\[ X = \begin{bmatrix} -y \\ y \\ 0 \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \]

and we can take \( X_2 = [-1, 1, 0]^t \).

Finding \( X_3 \): We solve \((A - \frac{1}{4}I_3)X = 0\). We have

\[ A - \frac{1}{4}I_3 = \begin{bmatrix} 1/4 & 1/2 & 0 \\ 1/4 & 0 & 1/2 \\ 1/4 & 1/4 & 1/4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \]
Hence the eigenspace consists of vectors $X = [x, y, z]^t$ satisfying $x = -2z$ and $y = z$, with $z$ arbitrary. Hence

$$X = \begin{bmatrix} -2z \\ z \\ 0 \end{bmatrix} = z \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

and we can take $X_3 = [-2, 1, 1]^t$.

Hence we can take $P = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

(c) $A = P \text{diag}(1, 0, \frac{1}{4}) P^{-1}$ so $A^n = P \text{diag}(1, 0, \frac{1}{4^n}) P^{-1}$.

Hence

$$A^n = \frac{1}{3} \begin{bmatrix} 1 & 0 & -\frac{2}{4^n} \\ 1 & 0 & \frac{1}{4^n} \\ 1 & 0 & \frac{1}{4^n} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & -3 \\ -1 & -1 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \frac{1}{3 \cdot 4^n} \begin{bmatrix} 2 & 2 & -4 \\ -1 & -1 & 2 \\ -1 & -1 & 2 \end{bmatrix}.$$
We have to find bases for each of the eigenspaces $N(A−9I_3)$ and $N(A−3I_3)$.

First we solve $(A−3I_3)X = 0$. We have

$$A−3I_3 = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$  

Hence the eigenspace consists of vectors $X = [x, y, z]^t$ satisfying $x = −y+z$, with $y$ and $z$ arbitrary. Hence

$$X = \begin{bmatrix} -y + z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

so $X_1 = [−1, 1, 0]^t$ and $X_2 = [1, 0, 1]^t$ form a basis for the eigenspace corresponding to the eigenvalue 3.

Next we solve $(A−9I_3)X = 0$. We have

$$A−9I_3 = \begin{bmatrix} -4 & 2 & -2 \\ 2 & -4 & -2 \\ -2 & -2 & -4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$  

Hence the eigenspace consists of vectors $X = [x, y, z]^t$ satisfying $x = −z$ and $y = −z$, with $z$ arbitrary. Hence

$$X = \begin{bmatrix} -z \\ -z \\ z \end{bmatrix} = z \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix},$$

and we can take $X_3 = [−1, −1, 1]^t$ as a basis for the eigenspace corresponding to the eigenvalue 9.
Then Theorem 6.2.3 assures us that $P = [X_1 | X_2 | X_3]$ is non-singular and

$$P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$
Figure 1: (a): $x^2 - 8x + 8y + 8 = 0$; (b): $y^2 - 12x + 2y + 25 = 0$

Section 7.3

1. (i) $x^2 - 8x + 8y + 8 = (x - 4)^2 + 8(y - 1)$. So the equation $x^2 - 8x + 8y + 8 = 0$ becomes

$$x_1^2 + 8y_1 = 0 \quad (1)$$

if we make a translation of axes $x - 4 = x_1$, $y - 1 = y_1$.

However equation (1) can be written as a standard form

$$y_1 = -\frac{1}{8}x_1^2,$$

which represents a parabola with vertex at $(4, 1)$. (See Figure 1(a).)

(ii) $y^2 - 12x + 2y + 25 = (y + 1)^2 - 12(x - 2)$. Hence $y^2 - 12x + 2y + 25 = 0$ becomes

$$y_1^2 - 12x_1 = 0 \quad (2)$$

if we make a translation of axes $x - 2 = x_1$, $y + 1 = y_1$.

However equation (2) can be written as a standard form

$$y_1^2 = 12x_1,$$

which represents a parabola with vertex at $(2, -1)$. (See Figure 1(b).)

2. $4xy - 3y^2 = X^tAX$, where $A = \begin{bmatrix} 0 & 2 \\ 2 & -3 \end{bmatrix}$ and $X = \begin{bmatrix} x \\ y \end{bmatrix}$. The eigenvalues of $A$ are the roots of $\lambda^2 + 3\lambda - 4 = 0$, namely $\lambda_1 = -4$ and $\lambda_2 = 1$.  

The eigenvectors corresponding to an eigenvalue $\lambda$ are the non-zero vectors $[x, y]^t$ satisfying

$$
\begin{bmatrix}
0 - \lambda & 2 \\
2 & -3 - \lambda
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0
\end{bmatrix}.
$$

$\lambda_1 = -4$ gives equations

\begin{align*}
4x + 2y &= 0 \\
2x + y &= 0
\end{align*}

which has the solution $y = -2x$. Hence

$$
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
x \\
-2x
\end{bmatrix}
=
x
\begin{bmatrix}
1 \\
-2
\end{bmatrix}.
$$

A corresponding unit eigenvector is $[1/\sqrt{5}, -2/\sqrt{5}]^t$.

$\lambda_2 = 1$ gives equations

\begin{align*}
-x + 2y &= 0 \\
2x - 4y &= 0
\end{align*}

which has the solution $x = 2y$. Hence

$$
\begin{bmatrix}
x \\
y
\end{bmatrix}
=
\begin{bmatrix}
2y \\
y
\end{bmatrix}
=y
\begin{bmatrix}
2 \\
1
\end{bmatrix}.
$$

A corresponding unit eigenvector is $[2/\sqrt{5}, 1/\sqrt{5}]^t$.

Hence if

$$
P = \begin{bmatrix}
\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\
\frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}}
\end{bmatrix},
$$

then $P$ is an orthogonal matrix. Also as $\det P = 1$, $P$ is a proper orthogonal matrix and the equation

$$
\begin{bmatrix}
x \\
y
\end{bmatrix}
= P
\begin{bmatrix}
x_1 \\
y_1
\end{bmatrix}
$$

represents a rotation to new $x_1, y_1$ axes whose positive directions are given by the respective columns of $P$. Also

$$
P^t A P = \begin{bmatrix}
-4 & 0 \\
0 & 1
\end{bmatrix}.
$$
Then $X^t AX = -4x_1^2 + y_1^2$ and the original equation $4xy - 3y^2 = 8$ becomes

$$-4x_1^2 + y_1^2 = 8,$$

or the standard form

$$\frac{-x_1^2}{2} + \frac{y_1^2}{8} = 1,$$

which represents an hyperbola.

The asymptotes assist in drawing the curve. They are given by the equations

$$\frac{-x_1^2}{2} + \frac{y_1^2}{8} = 0, \quad \text{or} \quad y_1 = \pm 2x_1.$$

Now

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = P^t \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

so

$$x_1 = \frac{x - 2y}{\sqrt{5}}, \quad y_1 = \frac{2x + y}{\sqrt{5}}.$$

Hence the asymptotes are

$$\frac{2x + y}{\sqrt{5}} = \pm 2 \left( \frac{x - 2y}{\sqrt{5}} \right),$$

which reduces to $y = 0$ and $y = 4x/3$. (See Figure 2(a).)

3. $8x^2 - 4xy + 5y^2 = X^t AX$, where $A = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$ and $X = \begin{bmatrix} x \\ y \end{bmatrix}$. The eigenvalues of $A$ are the roots of $\lambda^2 - 13\lambda + 36 = 0$, namely $\lambda_1 = 4$ and $\lambda_2 = 9$. Corresponding unit eigenvectors turn out to be $[1/\sqrt{5}, 2/\sqrt{5}]^t$ and $[-2/\sqrt{5}, 1/\sqrt{5}]^t$. Hence if

$$P = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix},$$

then $P$ is an orthogonal matrix. Also as $\det P = 1$, $P$ is a proper orthogonal matrix and the equation

$$\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

represents a rotation to new $x_1$, $y_1$ axes whose positive directions are given by the respective columns of $P$. Also

$$P^t AP = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}.$$
Then \( X^tAX = 4x_1^2 + 9y_1^2 \) and the original equation \( 8x^2 - 4xy + 5y^2 = 36 \) becomes \( 4x_1^2 + 9y_1^2 = 36 \), or the standard form

\[
\frac{x_1^2}{9} + \frac{y_1^2}{4} = 1,
\]

which represents an ellipse as in Figure 2(b).

The axes of symmetry turn out to be \( y = 2x \) and \( x = -2y \).

4. We give the sketch only for parts (i), (iii) and (iv). We give the working for (ii) only. See Figures 3(a) and 4(a) and 4(b), respectively.

(ii) We have to investigate the equation

\[
5x^2 - 4xy + 8y^2 + 4\sqrt{5}x - 16\sqrt{5}y + 4 = 0. \tag{3}
\]

Here \( 5x^2 - 4xy + 8y^2 = X^tAX \), where \( A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix} \) and \( X = \begin{bmatrix} x \\ y \end{bmatrix} \).

The eigenvalues of \( A \) are the roots of \( \lambda^2 - 13\lambda + 36 = 0 \), namely \( \lambda_1 = 9 \) and \( \lambda_2 = 4 \). Corresponding unit eigenvectors turn out to be \([1/\sqrt{5}, -2/\sqrt{5}]^t\) and \([2/\sqrt{5}, 1/\sqrt{5}]^t\). Hence if

\[
P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ \sqrt{5} & \sqrt{5} \end{bmatrix},
\]

then \( P \) is an orthogonal matrix. Also as \( \det P = 1 \), \( P \) is a proper orthogonal matrix and the equation

\[
\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}
\]
Figure 3: (a): $4x^2 - 9y^2 - 24x - 36y - 36 = 0$; (b): $5x^2 - 4xy + 8y^2 + \sqrt{5}x - 16\sqrt{5}y + 4 = 0$

Figure 4: (a): $4x^2 + y^2 - 4xy - 10y - 19 = 0$; (b): $77x^2 + 78xy - 27y^2 + 70x - 30y + 29 = 0$
represents a rotation to new $x_1$, $y_1$ axes whose positive directions are given by the respective columns of $P$. Also

$$P^tAP = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}.$$  

Moreover

$$5x^2 - 4xy + 8y^2 = 9x_1^2 + 4y_1^2.$$  

To get the coefficients of $x_1$ and $y_1$ in the transformed form of equation (3), we have to use the rotation equations

$$x = \frac{1}{\sqrt{5}}(x_1 + 2y_1), \quad y = \frac{1}{\sqrt{5}}(-2x_1 + y_1).$$  

Then equation (3) transforms to

$$9x_1^2 + 4y_1^2 + 36x_1 - 8y_1 + 4 = 0,$$

or, on completing the square,

$$9(x_1 + 2)^2 + 4(y_1 - 1)^2 = 36,$$

or in standard form

$$\frac{x_2^2}{4} + \frac{y_2^2}{9} = 1,$$

where $x_2 = x_1 + 2$ and $y_2 = y_1 - 1$. Thus we have an ellipse, centre $(x_2, y_2) = (0, 0)$, or $(x_1, y_1) = (-2, 1)$, or $(x, y) = (0, \sqrt{5})$.

The axes of symmetry are given by $x_2 = 0$ and $y_2 = 0$, or $x_1 + 2 = 0$ and $y_1 - 1 = 0$, or

$$\frac{1}{\sqrt{5}}(x - 2y) + 2 = 0 \quad \text{and} \quad \frac{1}{\sqrt{5}}(2x + y) - 1 = 0,$$

which reduce to $x - 2y + 2\sqrt{5} = 0$ and $2x + y - \sqrt{5} = 0$. See Figure 3(b).

5. (i) Consider the equation

$$2x^2 + y^2 + 3xy - 5x - 4y + 3 = 0.$$  

$$\Delta = \begin{vmatrix} 2 & 3/2 & -5/2 \\ 3/2 & 1 & -2 \\ -5/2 & -2 & 3 \end{vmatrix} = \begin{vmatrix} 4 & 3 & -5 \\ 3 & 2 & -4 \\ -5 & -4 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 \\ 3 & 2 & -4 \\ -2 & -2 & 2 \end{vmatrix} = 0.$$  

84
Let \( x = x_1 + \alpha \), \( y = y_1 + \beta \) and substitute in equation (4) to get
\[
2(x_1 + \alpha)^2 + (y_1 + \beta)^2 + 3(x_1 + \alpha)(y_1 + \beta) - 5(x_1 + \alpha) - 4(y_1 + \beta) + 3 = 0 \quad (5).
\]
Then equating the coefficients of \( x_1 \) and \( y_1 \) to 0 gives
\[
\begin{align*}
4\alpha + 3\beta - 5 &= 0 \\
3\alpha + 2\beta - 4 &= 0,
\end{align*}
\]
which has the unique solution \( \alpha = 2, \beta = -1 \). Then equation (5) simplifies to
\[
2x_1^2 + y_1^2 + 3x_1y_1 = 0 = (2x_1 + y_1)(x_1 + y_1).
\]
So relative to the \( x_1, y_1 \) coordinates, equation (4) describes two lines: \( 2x_1 + y_1 = 0 \) and \( x_1 + y_1 = 0 \). In terms of the original \( x, y \) coordinates, these lines become \( 2(x - 2) + (y + 1) = 0 \) and \( (x - 2) + (y + 1) = 0 \), i.e. \( 2x + y - 3 = 0 \) and \( x + y - 1 = 0 \), which intersect in the point
\[
(x, y) = (\alpha, \beta) = (2, -1).
\]

(ii) Consider the equation
\[
9x^2 + y^2 - 6xy + 6x - 2y + 1 = 0. \quad (6)
\]
Here
\[
\Delta = \begin{vmatrix} 9 & -3 & 3 \\ 3 & 1 & -1 \\ 3 & -1 & 1 \end{vmatrix} = 0,
\]
as column 3 = − column 2.

Let \( x = x_1 + \alpha \), \( y = y_1 + \beta \) and substitute in equation (6) to get
\[
9(x_1 + \alpha)^2 + (y_1 + \beta)^2 - 6(x_1 + \alpha)(y_1 + \beta) + 6(x_1 + \alpha) - 2(y_1 + \beta) + 1 = 0.
\]
Then equating the coefficients of \( x_1 \) and \( y_1 \) to 0 gives
\[
\begin{align*}
18\alpha - 6\beta + 6 &= 0 \\
-6\alpha + 2\beta - 2 &= 0,
\end{align*}
\]
or equivalently \(-3\alpha + \beta - 1 = 0\). Take \( \alpha = 0 \) and \( \beta = 1 \). Then equation (6) simplifies to
\[
9x_1^2 + y_1^2 - 6x_1y_1 = 0 = (3x_1 - y_1)^2. \quad (7)
\]
In terms of $x, y$ coordinates, equation (7) becomes

$$(3x - (y - 1))^2 = 0,$$ or $3x - y + 1 = 0.$

(iii) Consider the equation

$$x^2 + 4xy + 4y^2 - x - 2y - 2 = 0. \quad (8)$$

Arguing as in the previous examples, we find that any translation

$$x = x_1 + \alpha, \quad y = y_1 + \beta$$

where $2\alpha + 4\beta - 1 = 0$ has the property that the coefficients of $x_1$ and $y_1$ will be zero in the transformed version of equation (8). Take $\beta = 0$ and $\alpha = 1/2$. Then (8) reduces to

$$x_1^2 + 4x_1y_1 + 4y_1^2 - \frac{9}{4} = 0,$$

or $(x_1 + 2y_1)^2 = 3/2$. Hence $x_1 + 2y_1 = \pm3/2$, with corresponding equations

$$x + 2y = 2 \quad \text{and} \quad x + 2y = -1.$$
Section 8.8

1. The given line has equations

\[ x = 3 + t(13 - 3) = 3 + 10t, \]
\[ y = -2 + t(3 + 2) = -2 + 5t, \]
\[ z = 7 + t(-8 - 7) = 7 - 15t. \]

The line meets the plane \( y = 0 \) in the point \((x, 0, z)\), where \( 0 = -2 + 5t \), or \( t = 2/5 \). The corresponding values for \( x \) and \( z \) are 7 and 1, respectively.

2. \( E = \frac{1}{2}(B + C) \), \( F = (1 - t)A + tE \), where

\[ t = \frac{AF}{AE} = \frac{AF}{AF + FE} = \frac{AF/FE}{(AF/FE) + 1} = \frac{2}{3}. \]

Hence

\[ F = \frac{1}{3}A + \frac{2}{3}\left(\frac{1}{2}(B + C)\right) \]
\[ = \frac{1}{3}A + \frac{1}{3}(B + C) \]
\[ = \frac{1}{3}(A + B + C). \]

3. Let \( A = (2, 1, 4) \), \( B = (1, -1, 2) \), \( C = (3, 3, 6) \). Then we prove \( \overrightarrow{AC} = t \overrightarrow{AB} \) for some real \( t \). We have

\[ \overrightarrow{AC} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \overrightarrow{AB} = \begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix}. \]

Hence \( \overrightarrow{AC} = -1 \overrightarrow{AB} \) and consequently \( C \) is on the line \( AB \). In fact \( A \) is between \( C \) and \( B \), with \( AC = AB \).

4. The points \( P \) on the line \( AB \) which satisfy \( AP = \frac{2}{5}PB \) are given by

\( P = A + t \overrightarrow{AB} \), where \(|t/(1 - t)| = 2/5\). Hence \( t/(1 - t) = \pm 2/5 \).

The equation \( t/(1 - t) = 2/5 \) gives \( t = 2/7 \) and hence

\[ P = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} + \frac{2}{7} \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 16/7 \\ 29/7 \\ 3/7 \end{bmatrix}. \]
Hence $P = (16/7, 29/7, 3/7)$.

The equation $t/(1-t) = -2/5$ gives $t = -2/3$ and hence

$$
\mathbf{P} = \begin{bmatrix}
\frac{2}{3} \\
\frac{1}{5} \\
-1
\end{bmatrix} - \frac{2}{3} \begin{bmatrix}
\frac{1}{4} \\
\frac{1}{3} \\
\frac{4}{3}
\end{bmatrix} = \begin{bmatrix}
\frac{4}{3} \\
\frac{1}{3} \\
-13/3
\end{bmatrix}.
$$

Hence $P = (4/3, 1/3, -13/3)$.

5. An equation for $\mathcal{M}$ is $\mathbf{P} = \mathbf{A} + t \mathbf{BC}$, which reduces to

$$
\begin{align*}
x &= 1 + 6t \\
y &= 2 - 3t \\
z &= 3 + 7t.
\end{align*}
$$

An equation for $\mathcal{N}$ is $\mathbf{Q} = \mathbf{E} + s \mathbf{EF}$, which reduces to

$$
\begin{align*}
x &= 1 + 9s \\
y &= -1 \\
z &= 8 + 3s.
\end{align*}
$$

To find if and where $\mathcal{M}$ and $\mathcal{N}$ intersect, we set $\mathbf{P} = \mathbf{Q}$ and attempt to solve for $s$ and $t$. We find the unique solution $t = 1$, $s = 2/3$, proving that the lines meet in the point $(x, y, z) = (1+6, 2-3, 3+7) = (7, -1, 10)$.

6. Let $A = (3, 5, 6)$, $B = (-2, 7, 9)$, $C = (2, 1, 7)$. Then

(i)

$$
\cos \angle ABC = (\overrightarrow{BA} \cdot \overrightarrow{BC})/(\overrightarrow{BA} \cdot \overrightarrow{BC}),
$$

where $\overrightarrow{BA} = [-1, -2, -3]^t$ and $\overrightarrow{BC} = [4, -6, -2]^t$. Hence

$$
\cos \angle ABC = \frac{-4 + 12 + 6}{\sqrt{14} \sqrt{56}} = \frac{14}{\sqrt{14} \sqrt{56}} = \frac{1}{2}.
$$

Hence $\angle ABC = \pi/3$ radians or $60^\circ$.
(ii) \[ \cos \angle BAC = (\overrightarrow{AB} \cdot \overrightarrow{AC})/(\overrightarrow{AB} \cdot \overrightarrow{AC}), \]

where \( \overrightarrow{AB} = [1, 2, 3]^t \) and \( \overrightarrow{AC} = [5, -4, 1]^t \). Hence \[ \cos \angle BAC = \frac{5 - 8 + 3}{\sqrt{14}\sqrt{42}} = 0. \]

Hence \( \angle ABC = \pi/2 \) radians or 90°.

(iii) \[ \cos \angle ACB = (\overrightarrow{CA} \cdot \overrightarrow{CB})/(\overrightarrow{CA} \cdot \overrightarrow{CB}), \]

where \( \overrightarrow{CA} = [-5, 4, -1]^t \) and \( \overrightarrow{CB} = [-4, 6, 2]^t \). Hence \[ \cos \angle ACB = \frac{20 + 24 - 2}{\sqrt{42}\sqrt{56}} = \frac{42}{\sqrt{42}\sqrt{56}} = \frac{\sqrt{3}}{2}. \]

Hence \( \angle ACB = \pi/6 \) radians or 30°.

7. By Theorem 8.5.2, the closest point \( P \) on the line \( AB \) to the origin \( O \) is given by \( P = A + t \overrightarrow{AB} \), where \[ t = \frac{\overrightarrow{AO} \cdot \overrightarrow{AB}}{\overrightarrow{AB}^2} = \frac{-A \cdot \overrightarrow{AB}}{\overrightarrow{AB}^2}. \]

Now \[ A \cdot \overrightarrow{AB} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = -2. \]

Hence \( t = 2/11 \) and \[ P = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} + \frac{2}{11} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -16/11 \\ 13/11 \\ 35/11 \end{bmatrix} \]

and \( P = (-16/11, 13/11, 35/11) \).

Consequently the shortest distance \( OP \) is given by \[ \sqrt{\left(\frac{-16}{11}\right)^2 + \left(\frac{13}{11}\right)^2 + \left(\frac{35}{11}\right)^2} = \sqrt{\left(\frac{1650}{11}\right)} = \sqrt{\frac{15 \times 11 \times 10}{11}} = \frac{\sqrt{150}}{\sqrt{11}}. \]
Alternatively, we can calculate the distance $OP^2$, where $P$ is an arbitrary point on the line $AB$ and then minimize $OP^2$:

$P = A + t \overrightarrow{AB} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 + 3t \\ 1 + t \\ 3 + t \end{bmatrix}.$

Hence

$$OP^2 = (-2 + 3t)^2 + (1 + t)^2 + (3 + t)^2 = 11t^2 - 4t + 14 = 11 \left( t^2 - \frac{4}{11} t + \frac{14}{11} \right) = 11 \left( \left\{ t - \frac{2}{11} \right\}^2 + \frac{14}{11} - \frac{4}{121} \right) = 11 \left( \left\{ t - \frac{2}{11} \right\}^2 + \frac{150}{121} \right).$$

Consequently

$$OP^2 \geq 11 \times \frac{150}{121}$$

for all $t$; moreover

$$OP^2 = 11 \times \frac{150}{121}$$

when $t = 2/11$.

8. We first find parametric equations for $\mathcal{N}$ by solving the equations

\[
\begin{align*}
x + y - 2z &= 1 \\
x + 3y - z &= 4.
\end{align*}
\]

The augmented matrix is

$$\begin{bmatrix} 1 & 1 & -2 & 1 \\ 1 & 3 & -1 & 4 \end{bmatrix},$$

which reduces to

$$\begin{bmatrix} 1 & 0 & -5/2 & -1/2 \\ 0 & 1 & 1/2 & 3/2 \end{bmatrix},$$

Hence $x = -\frac{1}{2} + \frac{5}{2} z$, $y = \frac{3}{2} - \frac{z}{2}$, with $z$ arbitrary. Taking $z = 0$ gives a point $A = (-\frac{1}{2}, \frac{3}{2}, 0)$, while $z = 1$ gives a point $B = (2, 1, 1)$. 

90
Hence if \( C = (1, 0, 1) \), then the closest point on \( N \) to \( C \) is given by \( P = A + t \overrightarrow{AB} \), where \( t = (\overrightarrow{AC} \cdot \overrightarrow{AB})/AB^2 \).

Now 
\[
\overrightarrow{AC} = \begin{bmatrix} 3/2 \\ -3/2 \\ 1 \end{bmatrix} \quad \text{and} \quad \overrightarrow{AB} = \begin{bmatrix} 5/2 \\ -1/2 \\ 1 \end{bmatrix}, 
\]
so 
\[
t = \frac{\frac{3}{2} \times \frac{5}{2} + \frac{-3}{2} \times \frac{-1}{2} + 1 \times 1}{\left(\frac{3}{2}\right)^2 + \left(\frac{-1}{2}\right)^2 + 1^2} = \frac{11}{15}.
\]
Hence 
\[
P = \begin{bmatrix} -1/2 \\ 3/2 \\ 0 \end{bmatrix} + \frac{11}{15} \begin{bmatrix} 5/2 \\ -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 17/15 \\ 11/15 \end{bmatrix},
\]
so \( P = (4/3, 17/15, 11/15) \).

Also the shortest distance \( PC \) is given by 
\[
PC = \sqrt{\left(1 - \frac{4}{3}\right)^2 + \left(0 - \frac{17}{15}\right)^2 + \left(1 - \frac{11}{15}\right)^2} = \sqrt{\frac{330}{15}}.
\]

9. The intersection of the planes \( x + y - 2z = 4 \) and \( 3x - 2y + z = 1 \) is the line given by the equations 
\[
x = \frac{9}{5} + \frac{3}{5}z, \quad y = \frac{11}{5} + \frac{7}{5}z,
\]
where \( z \) is arbitrary. Hence the line \( \mathcal{L} \) has a direction vector \([3/5, 7/5, 1]^t\) or the simpler \([3, 7, 5]^t\). Then any plane of the form \( 3x + 7y + 5z = d \) will be perpendicular to \( \mathcal{L} \). The required plane has to pass through the point \((6, 0, 2)\), so this determines \( d \):
\[
3 \times 6 + 7 \times 0 + 5 \times 2 = d = 28.
\]

10. The length of the projection of the segment \( AB \) onto the line \( CD \) is given by the formula 
\[
\frac{|\overrightarrow{CD} \cdot \overrightarrow{AB}|}{CD}.
\]
Here \( \overrightarrow{CD} = [-8, 4, -1]^t \) and \( \overrightarrow{AB} = [4, -4, 3]^t \), so
\[
\frac{|\overrightarrow{CD} \cdot \overrightarrow{AB}|}{CD} = \frac{|(-8) \times 4 + 4 \times (-4) + (-1) \times 3|}{\sqrt{(-8)^2 + 4^2 + (-1)^2}} = \frac{|-51|}{\sqrt{81}} = \frac{51}{9} = \frac{17}{3}.
\]
11. A direction vector for \( \mathbf{L} \) is given by \( \overrightarrow{BC} = [-5, -2, 3]^t \). Hence the plane through \( A \) perpendicular to \( \mathbf{L} \) is given by

\[
-5x - 2y + 3z = (-5) \times 3 + (-2) \times (-1) + 3 \times 2 = -7.
\]

The position vector \( \mathbf{P} \) of an arbitrary point \( P \) on \( \mathbf{L} \) is given by \( \mathbf{P} = \mathbf{B} + t \overrightarrow{BC} \), or

\[
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} + t \begin{bmatrix} -5 \\ -2 \\ 3 \end{bmatrix},
\]

or equivalently \( x = 2 - 5t, y = 1 - 2t, z = 4 + 3t \).

To find the intersection of line \( \mathbf{L} \) and the given plane, we substitute the expressions for \( x, y, z \) found in terms of \( t \) into the plane equation and solve the resulting linear equation for \( t \):

\[
-5(2 - 5t) - 2(1 - 2t) + 3(4 + 3t) = -7,
\]

which gives \( t = -7/38 \). Hence \( P = \left( \frac{111}{38}, \frac{52}{38}, \frac{131}{38} \right) \) and

\[
AP = \sqrt{\left(3 - \frac{111}{38}\right)^2 + \left(-1 - \frac{52}{38}\right)^2 + \left(2 - \frac{131}{38}\right)^2}
\]

\[
= \frac{\sqrt{11134}}{38} = \frac{\sqrt{293 \times 38}}{38} = \frac{\sqrt{293}}{\sqrt{38}}.
\]

12. Let \( P \) be a point inside the triangle \( ABC \). Then the line through \( P \) and parallel to \( AC \) will meet the segments \( AB \) and \( BC \) in \( D \) and \( E \), respectively. Then

\[
\mathbf{P} = (1-r)\mathbf{D} + r\mathbf{E}, \quad 0 < r < 1;
\]

\[
\mathbf{D} = (1-s)\mathbf{B} + s\mathbf{A}, \quad 0 < s < 1;
\]

\[
\mathbf{E} = (1-t)\mathbf{B} + t\mathbf{C}, \quad 0 < t < 1.
\]

Hence

\[
\mathbf{P} = (1-r)\{(1-s)\mathbf{B} + s\mathbf{A}\} + r\{(1-t)\mathbf{B} + t\mathbf{C}\}
\]

\[
= (1-r)s\mathbf{A} + \{(1-r)(1-s) + r(1-t)\} \mathbf{B} + rt\mathbf{C}
\]

\[
= \alpha\mathbf{A} + \beta\mathbf{B} + \mathbf{C},
\]

92
where
\[ \alpha = (1 - r)s, \quad \beta = (1 - r)(1 - s) + r(1 - t), \quad \gamma = rt. \]

Then \(0 < \alpha < 1\), \(0 < \beta < 1 - (1 - r) + r = 1\). Also
\[ \alpha + \beta + \gamma = (1 - r)s + (1 - r)(1 - s) + r(1 - t) + rt = 1. \]

13. The line \(AB\) is given by \(P = A + t[3, 4, 5]^t\), or
\[ x = 6 + 3t, \quad y = -1 + 4t, \quad z = 11 + 5t. \]

Then \(B\) is found by substituting these expressions in the plane equation
\[ 3x + 4y + 5z = 10. \]

We find \(t = -59/50\) and consequently
\[ B = \left(6 - \frac{177}{50}, -1 - \frac{236}{50}, 11 - \frac{295}{50}\right) = \left(\frac{123}{50}, -\frac{286}{50}, \frac{255}{50}\right). \]

Then
\[ AB = \|\overrightarrow{AB}\| = ||t\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}|| = |t|\sqrt{3^2 + 4^2 + 5^2} = \frac{59}{50} \times \sqrt{50} = \frac{59}{\sqrt{50}}. \]

14. Let \(A = (-3, 0, 2)\), \(B = (6, 1, 4)\), \(C = (-5, 1, 0)\). Then the area of triangle \(ABC\) is \(\frac{1}{2}||\overrightarrow{AB} \times \overrightarrow{AC}||\). Now
\[ \overrightarrow{AB} \times \overrightarrow{AC} = \begin{bmatrix} 9 \\ 1 \\ 2 \end{bmatrix} \times \begin{bmatrix} -2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -4 \\ 14 \\ 11 \end{bmatrix}. \]

Hence \(||\overrightarrow{AB} \times \overrightarrow{AC}|| = \sqrt{333}.\)

15. Let \(A_1 = (2, 1, 4)\), \(A_2 = (1, -1, 2)\), \(A_3 = (4, -1, 1)\). Then the point \(P = (x, y, z)\) lies on the plane \(A_1A_2A_3\) if and only if
\[ A_1P \cdot (A_1A_2 \times A_1A_3) = 0, \]

93
16. Non–parallel lines $\mathcal{L}$ and $\mathcal{M}$ in three dimensional space are given by equations
\[ P = \mathbf{A} + s\mathbf{X}, \quad Q = \mathbf{B} + t\mathbf{Y}. \]

(i) Suppose $\overrightarrow{PQ}$ is orthogonal to both $\mathbf{X}$ and $\mathbf{Y}$. Now
\[ \overrightarrow{PQ} = Q - P = (\mathbf{B} + t\mathbf{Y}) - (\mathbf{A} + s\mathbf{X}) = \overrightarrow{AB} + t\mathbf{Y} - s\mathbf{X}. \]

Hence
\[ (\overrightarrow{AB} + t\mathbf{Y} + s\mathbf{X}) \cdot \mathbf{X} = 0 \]
\[ (\overrightarrow{AB} + t\mathbf{Y} + s\mathbf{X}) \cdot \mathbf{Y} = 0. \]

More explicitly
\[ t(\mathbf{Y} \cdot \mathbf{X}) - s(\mathbf{X} \cdot \mathbf{X}) = -\overrightarrow{AB} \cdot \mathbf{X} \]
\[ t(\mathbf{Y} \cdot \mathbf{Y}) - s(\mathbf{X} \cdot \mathbf{Y}) = -\overrightarrow{AB} \cdot \mathbf{Y}. \]

However the coefficient determinant of this system of linear equations in $t$ and $s$ is equal to
\[ \begin{vmatrix} \mathbf{Y} \cdot \mathbf{X} & -\mathbf{X} \cdot \mathbf{X} \\ \mathbf{Y} \cdot \mathbf{Y} & -\mathbf{X} \cdot \mathbf{Y} \end{vmatrix} = -(\mathbf{Y} \cdot \mathbf{X})^2 + (\mathbf{X} \cdot \mathbf{X})(\mathbf{Y} \cdot \mathbf{Y}) \]
\[ = ||\mathbf{X} \times \mathbf{Y}||^2 \neq 0, \]
as $\mathbf{X} \neq 0$, $\mathbf{Y} \neq 0$ and $\mathbf{X}$ and $\mathbf{Y}$ are not proportional ($\mathcal{L}$ and $\mathcal{M}$ are not parallel).

(ii) $P$ and $Q$ can be viewed as the projections of $C$ and $D$ onto the line $PQ$, where $C$ and $D$ are arbitrary points on the lines $\mathcal{L}$ and $\mathcal{M}$, respectively. Hence by equation (8.14) of Theorem 8.5.3, we have
\[ PQ \leq CD. \]

Finally we derive a useful formula for $PQ$. Again by Theorem 8.5.3
\[ PQ = \frac{|\overrightarrow{AB} \cdot \overrightarrow{PQ}|}{PQ} = |\overrightarrow{AB} \cdot \hat{n}|, \]

94
where \( \hat{n} = \frac{1}{||PQ||} PQ \) is a unit vector which is orthogonal to \( X \) and \( Y \).

Hence
\[
\hat{n} = t(X \times Y),
\]
where \( t = \pm 1/||X \times Y|| \). Hence
\[
PQ = \frac{|\overrightarrow{AB} \cdot (X \times Y)|}{||X \times Y||}.
\]

17. We use the formula of the previous question.

Line \( \mathcal{L} \) has the equation \( \mathbf{P} = \mathbf{A} + s\mathbf{X} \), where
\[
X = \overrightarrow{AC} = \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}.
\]

Line \( \mathcal{M} \) has the equation \( \mathbf{Q} = \mathbf{B} + t\mathbf{Y} \), where
\[
Y = \overrightarrow{BD} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.
\]

Hence \( X \times Y = [-6, 1, 5]' \) and \( ||X \times Y|| = \sqrt{62} \).
Hence the shortest distance between lines $AC$ and $BD$ is equal to

$$\frac{|\overrightarrow{AB} \cdot (X \times Y)|}{||X \times Y||} = \frac{\begin{vmatrix} 0 & -2 & -6 \\ -1 & 1 & 5 \end{vmatrix}}{\sqrt{62}} = \frac{3}{\sqrt{62}}.$$

18. Let $E$ be the foot of the perpendicular from $A_4$ to the plane $A_1A_2A_3$. Then

$$\text{vol } A_1A_2A_3A_4 = \frac{1}{3} (\text{area } A_1A_2A_3) \cdot A_4E.$$

Now

$$\text{area } A_1A_2A_3 = \frac{1}{2} || \overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3} ||.$$

Also $A_4E$ is the length of the projection of $A_1A_4$ onto the line $A_4E$. See figure below.)

Hence $A_4E = |\overrightarrow{A_1A_4} \cdot X|$, where $X$ is a unit direction vector for the line $A_4E$. We can take

$$X = \frac{\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3}}{|| \overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3} ||}.$$

Hence

$$\text{vol } A_1A_2A_3A_4 = \frac{1}{6} || \overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3} || |\overrightarrow{A_1A_4} \cdot (\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3})| \\ || \overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3} ||$$

$$= \frac{1}{6} |\overrightarrow{A_1A_4} \cdot (\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3})|.$$
\[
\frac{1}{6}(\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3}) \cdot \overrightarrow{A_4A_1}.
\]

19. We have \( \overrightarrow{CB} = [1, 4, -1]^t, \overrightarrow{CD} = [-3, 3, 0]^t, \overrightarrow{AD} = [3, 0, 3]^t \). Hence
\[
\overrightarrow{CB} \times \overrightarrow{CD} = 3i + 3j + 15k,
\]
so the vector \( i + j + 5k \) is perpendicular to the plane \( BCD \).

Now the plane \( BCD \) has equation \( x + y + 5z = 9 \), as \( B = (2, 2, 1) \) is on the plane.

Also the line through \( A \) normal to plane \( BCD \) has equation
\[
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} = (1 + t) \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}.
\]

Hence \( x = 1 + t, \ y = 1 + t, \ z = 5(1 + t) \).

[We remark that this line meets plane \( BCD \) in a point \( E \) which is given by a value of \( t \) found by solving \( (1 + t) + (1 + t) + 5(5 + 5t) = 9 \).]

So \( t = -2/3 \) and \( E = (1/3, 1/3, 5/3). \]

The distance from \( A \) to plane \( BCD \) is
\[
\frac{|1 \times 1 + 1 \times 1 + 5 \times 5 - 9|}{1^2 + 1^2 + 5^2} = \frac{18}{27} = 2\sqrt{3}.
\]

To find the distance between lines \( AD \) and \( BC \), we first note that

(a) The equation of \( AD \) is
\[
\mathbf{P} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} + t \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 + 3t \\ 1 \\ 5 + 3t \end{bmatrix};
\]

(b) The equation of \( BC \) is
\[
\mathbf{Q} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 + s \\ 2 + 4s \\ 1 - s \end{bmatrix}.
\]
Then $\overrightarrow{PQ} = [1 + s - 3t, 1 + 4s, -4 - s - 3t]^t$ and we find $s$ and $t$ by solving the equations $\overrightarrow{PQ} \cdot \overrightarrow{AD} = 0$ and $\overrightarrow{PQ} \cdot \overrightarrow{BC} = 0$, or
\begin{align*}
(1 + s - 3t)3 + (1 + 4s)0 + (-4 - s - 3t)3 &= 0 \\
(1 + s - 3t) + 4(1 + 4s) - (-4 - s - 3t) &= 0.
\end{align*}
Hence $t = -1/2 = s$.

Correspondingly, $P = (-1/2, 1, 7/2)$ and $Q = (3/2, 0, 3/2)$.

Thus we have found the closest points $P$ and $Q$ on the respective lines $AD$ and $BC$. Finally the shortest distance between the lines is
$$PQ = \| \overrightarrow{PQ} \| = 3.$$