Chapter 3

Matrices and Eigenvalue Problems

3.1 Introduction

In modern mathematics, matrix theory occupies an important place and has applications in almost al branches of engineering and physical sciences. Matrices of order $m \times n$ form a vector space and the define linear transformations which map vector spaces consisting of vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ into anothe vector space consisting of vectors in $\mathbb{R}^m$ or $\mathbb{C}^m$ under a given set of rules of vector addition and scalar multiplication. A matrix does not denote a number and no value can be assigned to it. The usua rules of arithmetic operations do not hold for matrices. The rules defining the operations on matrices are usually called its algebra. In this chapter, we shall discuss the matrix algebra and its use in solving linear system of algebraic equations $\mathbf{Ax} = \mathbf{b}$ and solving the eigenvalue problem $\mathbf{Ax} = \lambda \mathbf{x}$.

3.2 Matrices

An $m \times n$ matrix is an arrangement of mn objects (not necessarily distinct) in m *rows* and n *columns* in the form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}. \tag{3.1}$$

We say that the matrix is of *order* $m \times n$ (m by n). The objects $a_{11}, a_{12}, \dots, a_{mn}$ are called the *elements* of the matrix. Each element of the matrix can be a real or a complex number or a function of one or more variables or any other object. The element a_{ij} which is common to the ith row and the jth column is called its *general element*. The matrices are usually denoted by boldface uppercase letters $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, ... etc. When the order of the matrix is understood, we can simply write $\mathbf{A} = (a_{ij})$. If all the elements of a matrix are real, it is called a *real matrix*, whereas if one or more elements of a matrix are complex, it is called a *complex matrix*. We define the following particular types of matrices.

Row vector A matrix of order $1 \times n$, that is, it has one row and n columns is called a *row* vector or a row matrix of order n and is written as

$$[a_{11} \;\; a_{12} \;\; \dots \;\; a_{1n}], \quad \text{or} \quad [a_1 \;\; a_2 \;\; \dots \;\; a_n]$$

in which a_{1j} (or a_j) is the jth element.

Column vector A matrix of order $m \times 1$, that is, it has m rows and one column is called a *column vector* or a *column matrix* of order m and is written as

$$\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{bmatrix}, \text{ or } \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

in which b_{j1} (or b_j) is the jth element.

The number of elements in a row/column vector is called its *order*. The vectors are usually denoted by boldface lower case letters **a**, **b**, **c**, . . . etc. If a vector has n elements and all its elements are real numbers, then it is called an *ordered n-tuple* in $\mathbb{R}^n$, whereas if one or more elements are complex numbers, then it is called an ordered n-tuple in $\mathbb{C}^n$.

Rectangular matrix A matrix **A** of order $m \times n$, $m \neq n$ is called a *rectangular matrix*.

Square matrices A matrix **A** of order $m \times n$ in which $m = n$, that is number of rows is equal to the number of columns is called a square matrix of order n. The elements a_{ii}, that is the elements a_{11}, a_{22}, . . . , a_{nn} are called the *diagonal elements* and the line on which these elements lie is called the *principal diagonal* or the *main diagonal* of the matrix. The elements a_{ij}, when $i \neq j$ are called the *off-diagonal elements*. The sum of the diagonal elements of a square matrix is called the *trace* of the matrix.

Null matrix A matrix **A** of order $m \times n$ in which all the elements are zero is called a *null matrix* or a *zero matrix* and is denoted by **0**.

Equal matrices Two matrices $\mathbf{A} = (a_{ij})_{m \times n}$ and $\mathbf{B} = (b_{ij})_{p \times q}$ are said to be equal, when

(i) they are of the same order, that is $m = p$, $n = q$ and

(ii) their corresponding elements are equal, that is $a_{ij} = b_{ij}$ for all i, j.

Diagonal matrix A square matrix **A** in which all the off-diagonal elements a_{ij}, $i \neq j$ are zero is called a diagonal matrix. For example

$$\mathbf{A} = \begin{bmatrix} a_{11} & & & \mathbf{0} \\ & a_{22} & & \\ & & \ddots & \\ \mathbf{0} & & & a_{nn} \end{bmatrix} \text{ is a } \textit{diagonal matrix} \text{ of order } n.$$

A diagonal matrix is denoted by **D**. It is also written as diag $[a_{11} \quad a_{22} \quad \cdots \quad a_{nn}]$.

If all the elements of a diagonal matrix of order n are equal, that is $a_{ii} = \alpha$ for all i, then the matrix

$$\mathbf{A} = \begin{bmatrix} \alpha & & & \mathbf{0} \\ & \alpha & & \\ & & \ddots & \\ \mathbf{0} & & & \alpha \end{bmatrix} \text{ is called a } \textit{scalar matrix} \text{ of order } n.$$

If all the elements of a diagonal matrix of order n are 1, then the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$ is called an *unit matrix* or an *identity matrix* of order n.

An identity matrix is denoted by $\mathbf{I}$.

Submatrix A matrix obtained by omitting some rows and/or columns from a given matrix $\mathbf{A}$ is called a *submatrix* of $\mathbf{A}$. As a convention, the given matrix $\mathbf{A}$ is also taken as a submatrix of $\mathbf{A}$.

3.2.1 Matrix Algebra

The basic operations allowed on matrices are

(i) multiplication of a matrix by a scalar,

(ii) addition/subtraction of two matrices,

(iii) multiplication of two matrices.

Note that there is no concept of dividing a matrix by a matrix. Therefore, the operation $\mathbf{A}/\mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are matrices is not defined.

Multiplication of a matrix by a scalar

Let α be a scalar (real or complex) and $\mathbf{A} = (a_{ij})$ be a given matrix of order $m \times n$. Then

$$\mathbf{B} = \alpha \mathbf{A} = \alpha(a_{ij}) = (\alpha a_{ij}) \quad \text{for all } i \text{ and } j. \tag{3.2}$$

The order of the new matrix $\mathbf{B}$ is same as that of the matrix $\mathbf{A}$.

Addition/subtraction of two matrices

Let $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ be two matrices of the same order. Then

$$\mathbf{C} = (c_{ij}) = \mathbf{A} + \mathbf{B} = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \quad \text{for all } i \text{ and } j \tag{3.3a}$$

and

$$\mathbf{D} = (d_{ij}) = \mathbf{A} - \mathbf{B} = (a_{ij}) - (b_{ij}) = (a_{ij} - b_{ij}), \quad \text{for all } i \text{ and } j. \tag{3.3b}$$

The order of the new matrix $\mathbf{C}$ or $\mathbf{D}$ is the same as that of the matrices $\mathbf{A}$ and $\mathbf{B}$. Matrices of the same order are said to be *conformable* for addition/subtraction.

If $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$ are p matrices which are conformable for addition and $\alpha_1, \alpha_2, \ldots, \alpha_p$ are any scalars, then

$$\mathbf{C} = \alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2 + \ldots + \alpha_p \mathbf{A}_p \tag{3.4}$$

is called a linear combination of the matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$. The order of the matrix $\mathbf{C}$ is same as that of $\mathbf{A}_i$, $i = 1, 2, \ldots, p$.

Properties of the matrix addition and scalar multiplication

Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ be the matrices which are conformable for addition and α, β be scalars. Then

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. (commutative law)

2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ (associative law).
3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$ ($\mathbf{0}$ is the null matrix of the same order as $\mathbf{A}$).
4. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.
5. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$.
6. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$.
7. $\alpha(\beta\mathbf{A}) = \alpha\beta\mathbf{A}$.
8. $1 \times \mathbf{A} = \mathbf{A}$ and $0 \times \mathbf{A} = \mathbf{0}$.

Multiplication of two matrices

The product **AB** of two matrices **A** and **B** is defined only when the number of columns in **A** is equal to the number of rows in **B**. Such matrices are said to be *conformable* for multiplication. Let $\mathbf{A} = (a_{ij})$ be an $m \times n$ matrix and $\mathbf{B} = (b_{ij})$ be an $n \times p$ matrix. Then the product matrix

$$\mathbf{C} = (c_{ij}) = \mathbf{AB} = \underset{m \times n}{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}} \underset{n \times p}{\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2p} \\ \vdots & & & & & \\ b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix}}$$

is a matrix of order $m \times p$. The general element of the product matrix **C** is given by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}. \tag{3.5}$$

In the product **AB**, **B** is said to be pre-multiplied by **A** or **A** is said to be post-multiplied by **B**.

If **A** is a row matrix of order $1 \times n$ and **B** is a column matrix of order $n \times 1$, then **AB** is a matrix of order 1×1, that is a single element, and **BA** is a matrix of order $n \times n$.

Remark 1

(a) It is possible that for two given matrices **A** and **B**, the product matrix **AB** is defined but the product matrix **BA** may not be defined. For example, if **A** is a 2×3 matrix and **B** is a 3×4 matrix, then the product matrix **AB** is defined and is a matrix of order 2×4, whereas the product matrix **BA** is not defined.

(b) If both the product matrices **AB** and **BA** are defined, then both the matrices **AB** and **BA** are square matrices. In general $\mathbf{AB} \neq \mathbf{BA}$. Thus, the matrix product is not commutative. If $\mathbf{AB} = \mathbf{BA}$, then the matrices **A** and **B** are said to *commute* with each other.

(c) If $\mathbf{AB} = \mathbf{0}$, then it does not always imply that either $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. For example, let

$$\mathbf{A} = \begin{bmatrix} x & 0 \\ y & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}$$

then

$$\mathbf{AB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{BA} = \begin{bmatrix} 0 & 0 \\ ax + by & 0 \end{bmatrix} \neq \mathbf{AB}.$$

(d) If $\mathbf{AB} = \mathbf{AC}$, it does not always imply that $\mathbf{B} = \mathbf{C}$.

(e) Define $\mathbf{A}^k = \mathbf{A} \times \mathbf{A} \ldots \times \mathbf{A}$ (k times). Then, a matrix $\mathbf{A}$ such that $\mathbf{A}^k = \mathbf{0}$ for some positive integer k is said to be *nilpotent*. The smallest value of k for which $\mathbf{A}^k = \mathbf{0}$ is called the *index of nilpotency* of the matrix $\mathbf{A}$.

(f) If $\mathbf{A}^2 = \mathbf{A}$, then $\mathbf{A}$ is called an *idempotent matrix.*

Properties of matrix multiplication

1. If $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are matrices of orders $m \times n$, $n \times p$ and $p \times q$ respectively, then

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \qquad \text{(associative law)}$$

is a matrix of order $m \times q$.

2. If $\mathbf{A}$ is a matrix of order $m \times n$ and $\mathbf{B}$, $\mathbf{C}$ are matrices of order $n \times p$, then

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \qquad \text{(left distributive law).}$$

3. If $\mathbf{A}$, $\mathbf{B}$ are matrices of order $m \times n$ and $\mathbf{C}$ is a matrix of order $n \times p$, then

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \qquad \text{(right distributive law).}$$

4. If $\mathbf{A}$ is a matrix of order $m \times n$ and $\mathbf{B}$ is a matrix of order $n \times p$, then

$$\alpha(\mathbf{AB}) = \mathbf{A}(\alpha\mathbf{B}) = (\alpha\mathbf{A})\mathbf{B}$$

for any scalar α.

3.2.2 Some Special Matrices

We now define some special matrices.

Transpose of a matrix The matrix obtained by interchanging the corresponding rows and columns of a given matrix $\mathbf{A}$ is called the *transpose matrix* of $\mathbf{A}$ and is denoted by $\mathbf{A}^T$, that is, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \text{ then } \mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & & a_{mn} \end{bmatrix}.$$

If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^T$ is an $n \times m$ matrix. Also, both the product matrices $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$ are defined, and

$$\mathbf{A}^T\mathbf{A} = (n \times m)(m \times n) \text{ is an } n \times n \text{ square matrix}$$

and

$$\mathbf{A}\mathbf{A}^T = (m \times n)(n \times m) \text{ is an } m \times m \text{ square matrix.}$$

A column vector $\mathbf{b}$ can also be written as $[b_1 \quad b_2 \ldots b_n]^T$.

The following results can be easily verified

1. The transpose of a row matrix is a column matrix and the transpose of a column matrix is a row matrix.
2. $(\mathbf{A}^T)^T = \mathbf{A}$.

3. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$, when the matrices **A** and **B** are conformable for addition.
4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$, when the matrices **A** and **B** are conformable for multiplication.

   If the product $\mathbf{A}_1 \mathbf{A}_2 \ldots \mathbf{A}_p$ is defined, then

$$[\mathbf{A}_1\mathbf{A}_2 \ldots \mathbf{A}_p]^T = \mathbf{A}_p^T \mathbf{A}_{p-1}^T \ldots \mathbf{A}_1^T.$$

Remark 2

The product of a row vector $\mathbf{a}_i = (a_{i1} \;\; a_{i2} \;\; \ldots \;\; a_{in})$ of order $1 \times n$ and a column vector $\mathbf{b}_j = (b_{1j} \;\; b_{2j} \;\; \ldots \;\; b_{nj})^T$ of order $n \times 1$ is called the *dot product* or the *inner product* of the vectors $\mathbf{a}_i$ and $\mathbf{b}_j$, that is

$$c_{ij} = \mathbf{a}_i \cdot \mathbf{b}_j = \sum_{k=1}^{n} a_{ik} b_{kj}$$

which is a scalar. In terms of the inner products, the product matrix **C** in Eq. (3.5) can be written as

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_1 \cdot \mathbf{b}_p \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \ldots & \mathbf{a}_2 \cdot \mathbf{b}_p \\ \ldots & \ldots & \ldots & \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \ldots & \mathbf{b}_m \cdot \mathbf{b}_p \end{bmatrix}. \tag{3.6}$$

Symmetric and skew-symmetric matrices A real square matrix $\mathbf{A} = (a_{ij})$ is said to be

symmetric, if $a_{ij} = a_{ji}$ for all i and j, that is $\mathbf{A} = \mathbf{A}^T$

skew-symmetric, if $a_{ij} = -a_{ji}$ for all i and j, that is $\mathbf{A} = -\mathbf{A}^T$.

Remark 3

(a) In a skew-symmetric matrix $\mathbf{A} = (a_{ij})$, all its diagonal elements are zero.

(b) The matrix which is both symmetric and skew-symmetric must be a null matrix.

(c) For any real square matrix **A**, the matrix $\mathbf{A} + \mathbf{A}^T$ is always symmetric and the matrix $\mathbf{A} - \mathbf{A}^T$ is always skew-symmetric. Therefore, a real square matrix **A** can be written as the sum of a symmetric matrix and a skew-symmetric matrix. That is

$$\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^T).$$

Triangular matrices A square matrix $\mathbf{A} = (a_{ij})$ is called a *lower triangular matrix* if $a_{ij} = 0$, whenever $i < j$, that is all elements above the principal diagonal are zero and an *upper triangular matrix* if $a_{ij} = 0$, whenever $i > j$, that is all the elements below the principal diagonal are zero.

Conjugate matrix Let $\mathbf{A} = (a_{ij})$ be a complex matrix. Let $\overline{a}_{ij}$ denote the complex conjugate of a_{ij}. Then, the matrix $\overline{\mathbf{A}} = (\overline{a}_{ij})$ is called the *conjugate matrix* of *A*.

Hermitian and skew-Hermitian matrices A complex matrix **A** is called an *Hermitian matrix* if $\overline{\mathbf{A}} = \mathbf{A}^T$ or $\mathbf{A} = (\overline{\mathbf{A}})^T$ and a skew-Hermitian matrix if $\overline{\mathbf{A}} = -\mathbf{A}^T$ or $\mathbf{A} = -(\overline{\mathbf{A}})^T$. Sometimes, a Hermitian matrix is denoted by $\mathbf{A}^H$ or $\mathbf{A}^*$.

Remark 4

(a) If **A** is a real matrix, then an Hermitian matrix is same as a symmetric matrix and a skew-Hermitian matrix is same as a skew-symmetric matrix.

(b) In an Hermitian matrix, all the diagonal elements are real (let $a_{jj} = x_j + i\,y_j$; then $a_{jj} = \bar{a}_{jj}$ gives $x_j + i\,y_j = x_j - i\,y_j$ or $y_j = 0$ for all j).

(c) In a skew-Hermitian matrix, all the diagonal elements are either 0 or pure imaginary (let $a_{jj} = x_j + i\,y_j$; then $a_{jj} = -\bar{a}_{jj}$ gives $x_j + i\,y_j = -(x_j - i\,y_j)$ or $x_j = 0$ for all j).

(d) For any complex square matrix **A**, the matrix $\mathbf{A} + \overline{\mathbf{A}}^T$ is always an Hermitian matrix and the matrix $\mathbf{A} - \overline{\mathbf{A}}^T$ is always a skew-Hermitian matrix. Therefore, a complex square matrix **A** can be written as the sum of an Hermitian matrix and a skew-Hermitian matrix, that is

$$\mathbf{A} = \frac{1}{2}(\mathbf{A} + \overline{\mathbf{A}}^T) + \frac{1}{2}(\mathbf{A} - \overline{\mathbf{A}}^T).$$

Example 3.1 Let **A** and **B** be two symmetric matrices of the same order. Show that the matrix **AB** is symmetric if and only if **AB** = **BA**, that is the matrices **A** and **B** commute.

Solution Since the matrices **A** and **B** are symmetric, we have

$$\mathbf{A}^T = \mathbf{A} \quad \text{and} \quad \mathbf{B}^T = \mathbf{B}.$$

Let **AB** be symmetric. Then

$$(\mathbf{AB})^T = \mathbf{AB}, \quad \text{or} \quad \mathbf{B}^T\mathbf{A}^T = \mathbf{AB}, \quad \text{or} \quad \mathbf{BA} = \mathbf{AB}.$$

Now, let **AB** = **BA**. Taking transpose on both sides, we get

$$(\mathbf{AB})^T = (\mathbf{B\,A})^T = \mathbf{A}^T\mathbf{B}^T = \mathbf{AB}.$$

Hence, the result.

3.2.3 Determinants

With every square matrix **A** of order n, we associate a determinant of order n which is denoted by *det* (**A**) or | **A** |. The determinant has a value and this value is real if the matrix **A** is real and may be real or complex, if the matrix is complex. A determinant of order n is defined as

$$det\,(\mathbf{A}) = |\,\mathbf{A}\,| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{3.7}$$

We now discuss methods to find the value of a determinant. A determinant of order 2 has two rows and two columns. Its value is given by

$$|\,\mathbf{A}\,| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

We evaluate higher order determinants using minors and cofactors.

Minors and cofactors Let a_{ij} be the general element of a determinant. If we delete the ith row and the jth column from the determinant, we obtain a new determinant of order $(n - 1)$ which is called the *minor* of the element a_{ij}. We denote this minor by M_{ij}. The cofactor of the element a_{ij} is defined as

$$A_{ij} = (-1)^{i+j} M_{ij}. \tag{3.8}$$

We can expand a determinant of order n through the elements of any row or any column. The value of the determinant is the sum of the products of the elements of the ith row (or jth column) and thei[r] corresponding cofactors. Thus, we have

$$|\mathbf{A}| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} = \sum_{j=1}^{n} a_{ij} A_{ij} \tag{3.9a}$$

when we expand through the elements of the ith row, or

$$|\mathbf{A}| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij} \tag{3.9b}$$

when we expand through the elements of the jth column. Generally, we expand a determinant through that row or column which has a number of zeros. We can use one or more of the following properties of the determinants to simplify the evaluation of determinants.

Properties of determinants

1. If all the elements of a row (or column) are zero, then the value of the determinant is zero.
2. The value of a determinant remains unchanged if its corresponding rows and columns are interchanged, that is

$$|\mathbf{A}| = |\mathbf{A}^T|.$$

3. If any two rows (or columns) are interchanged, then the value of the determinant is multiplied by (– 1).
4. If the corresponding elements of two rows (or columns) are same, that is two rows (or columns) are identical, then the value of the determinant is zero.
5. If the corresponding elements of two rows (or columns) are proportional to each other, then the value of the determinant is zero.
6. The value of the determinant of a diagonal or a lower triangular or an upper triangular matrix is the product of its diagonal elements.
7. If each element of a row (or column) is multiplied by a scalar α, then the value of the determinant is multiplied by the scalar α. Therefore, if β is a factor of each element of a row (or column), then this factor β can be taken out of the determinant.

   Note that when we multiply a matrix by a scalar α, then every element of the matrix is multiplied by α. Therefore, $|\alpha\mathbf{A}| = \alpha^n |\mathbf{A}|$ where $\mathbf{A}$ is a matrix of order n.
8. If each element of any row (or column) can be written as the sum of two (or more) terms, then the determinant can be written as sum of two (or more) determinants.
9. If a non-zero constant multiple of the elements of some row (or column) is added to the corresponding elements of some other row (or column), then the value of the determinant remains unchanged.

   Remark 5

   When the elements of the jth row are multiplied by a non-zero constant k and added to the corresponding elements of the ith row, we denote this operation as $R_i \leftarrow R_i + kR_j$, where R_i is the ith row of $|\mathbf{A}|$. The elements of the jth row remain unchanged whereas the elements of the ith row get changed. This operation is called an *elementary row operation*. Similarly, the operation $C_i \leftarrow C_i + kC_j$, where C_i is the ith column of $|\mathbf{A}|$, is called the *elementary column operation*. Therefore, under elementary row (or column) operations, the value of a determinant is unchanged.

10. The sum of the products of elements of any row (or column) with their corresponding cofactors gives the value of the determinant. However, the sum of the products of the elements of any row (or column) with the corresponding cofactors of any other row (or column) is zero. Thus, we have

$$\sum_{k=1}^{n} a_{ik} A_{jk} = \begin{cases} |\mathbf{A}|, & i = j \\ 0 & i \neq j \end{cases} \quad \text{(expansion through } i\text{th row)} \tag{3.10a}$$

or

$$\sum_{k=1}^{n} a_{ki} A_{kj} = \begin{cases} |\mathbf{A}|, & i = j \\ 0 & i \neq j \end{cases} \quad \text{(expansion through } j\text{th column).} \tag{3.10b}$$

11. $|\mathbf{A} + \mathbf{B}| \neq |\mathbf{A}| + |\mathbf{B}|$, in general.

Product of two determinants

If $\mathbf{A}$ and $\mathbf{B}$ are two square matrices of the same order, then

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|.$$

Since $|\mathbf{A}| = |\mathbf{A}^T|$, we can multiply two determinants in any one of the following ways

(i) row by row, (ii) column by column,

(iii) row by column, (iv) column by row.

The value of the determinant is same in each case.

Rank of a matrix

The *rank* of a matrix $\mathbf{A}$, denoted by r or $r(\mathbf{A})$ is the order of the largest non-zero minor of $|\mathbf{A}|$. Therefore, the rank of a matrix is the largest value of r, for which there exists at least one $r \times r$ submatrix of $\mathbf{A}$ whose determinant is not zero. Thus, for an $m \times n$ matrix $r \leq \min(m, n)$. For a square matrix $\mathbf{A}$ of order n, the rank $r = n$ if $|\mathbf{A}| \neq 0$, otherwise $r < n$. The rank of a null matrix is zero and if the rank of a matrix is 0, then it must be a null matrix.

Example 3.2 Find the value of the determinant

$$|\mathbf{A}| = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 2 & 4 \\ -1 & 3 & 2 \end{vmatrix}$$

(i) using elementary row operations, (ii) using elementary column operations.

Solution

(i) Applying the operations $R_2 \leftarrow R_2 - (3/2)R_1$ and $R_3 \leftarrow R_3 + (1/2)R_1$, we obtain

$$|\mathbf{A}| = \begin{vmatrix} 2 & 1 & -1 \\ 0 & 1/2 & 11/2 \\ 0 & 7/2 & 3/2 \end{vmatrix}$$

Applying the operation $R_3 \leftarrow R_3 - 7R_2$, we obtain

$$|\mathbf{A}| = \begin{vmatrix} 2 & 1 & -1 \\ 0 & 1/2 & 11/2 \\ 0 & 0 & -37 \end{vmatrix} = 2(1/2)(-37) = -37$$

since the value of the determinant of an upper triangular matrix is the product of diagonal elements.

(ii) Applying the operations $C_2 \leftarrow C_2 - (1/2)C_1$ and $C_3 \leftarrow C_3 + (1/2)C_1$, we obtain

$$|\mathbf{A}| = \begin{vmatrix} 2 & 0 & 0 \\ 3 & 1/2 & 11/2 \\ -1 & 7/2 & 3/2 \end{vmatrix}$$

Applying the operation $C_3 \leftarrow C_3 - 11C_2$, we obtain

$$|\mathbf{A}| = \begin{vmatrix} 2 & 0 & 0 \\ 3 & 1/2 & 0 \\ -1 & 7/2 & -37 \end{vmatrix} = 2(1/2)(-37) = -37$$

since the value of the determinant of a lower triangular matrix is the product of diagonal elements.

Example 3.3 Show that

$$|\mathbf{A}| = \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = (\alpha - \beta)(\beta - \gamma)(\gamma - \alpha).$$

Solution Applying the operations $C_1 \leftarrow C_1 - C_3$ and $C_2 \leftarrow C_2 - C_3$, we get

$$|\mathbf{A}| = \begin{vmatrix} 0 & 0 & 1 \\ \alpha - \gamma & \beta - \gamma & \gamma \\ \alpha^2 - \gamma^2 & \beta^2 - \gamma^2 & \gamma^2 \end{vmatrix} = (\alpha - \gamma)(\beta - \gamma) \begin{vmatrix} 0 & 0 & 1 \\ 1 & 1 & \gamma \\ \alpha + \gamma & \beta + \gamma & \gamma^2 \end{vmatrix}$$

$$= (\alpha - \gamma)(\beta - \gamma) \begin{vmatrix} 1 & 1 \\ \alpha + \gamma & \beta + \gamma \end{vmatrix} = (\alpha - \gamma)(\beta - \gamma)(\beta + \gamma - \alpha - \gamma)$$

$$= (\alpha - \gamma)(\beta - \gamma)(\beta - \alpha) = (\alpha - \beta)(\beta - \gamma)(\gamma - \alpha).$$

Example 3.4 Given the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 2 & 1 \\ 1 & -5 & 2 \\ 1 & 7 & 1 \end{bmatrix}$$

verify that

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|.$$

Solution We have

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 1 & -5 & 2 \\ 1 & 7 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 13 & 8 \\ 23 & 25 & 20 \\ 22 & 117 & 0 \end{bmatrix}$$

$$\text{Therefore,} \quad |\mathbf{AB}| = \begin{vmatrix} 8 & 13 & 8 \\ 23 & 25 & 20 \\ 22 & 117 & 0 \end{vmatrix} = 8(0 - 2340) - 23(0 - 936) + 22(260 - 200)$$

$$= -18720 + 21528 + 1320 = 4128.$$

We can find the value of the product $|\mathbf{A}||\mathbf{B}|$ directly as

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} \begin{vmatrix} 3 & 2 & 1 \\ 1 & -5 & 2 \\ 1 & 7 & 1 \end{vmatrix} = \begin{vmatrix} 10 & -3 & 18 \\ 28 & -9 & 45 \\ 14 & 65 & -40 \end{vmatrix} \quad \text{(multiplying determinants row by row)}$$

$$= 10(360 - 2925) - 28(120 - 1170) + 14(-135 + 162)$$

$$= -25650 + 29400 + 378 = 4128.$$

We can also find $|\mathbf{A}|$ and $|\mathbf{B}|$ and then multiply.

Example 3.5 Without evaluating the determinant, show that

$$D = \begin{vmatrix} \cos(A-P) & \cos(A-Q) & \cos(A-R) \\ \cos(B-P) & \cos(B-Q) & \cos(B-R) \\ \cos(C-P) & \cos(C-Q) & \cos(C-R) \end{vmatrix} = 0.$$

Solution Expanding all the terms, we have

$$D = \begin{vmatrix} \cos A \cos P + \sin A \sin P & \cos A \cos Q + \sin A \sin Q & \cos A \cos R + \sin A \sin R \\ \cos B \cos P + \sin B \sin P & \cos B \cos Q + \sin B \sin Q & \cos B \cos R + \sin B \sin R \\ \cos C \cos P + \sin C \sin P & \cos C \cos Q + \sin C \sin Q & \cos C \cos R + \sin C \sin R \end{vmatrix}$$

$$= \begin{vmatrix} \cos A & \sin A & 0 \\ \cos B & \sin B & 0 \\ \cos C & \sin C & 0 \end{vmatrix} \begin{vmatrix} \cos P & \sin P & 0 \\ \cos Q & \sin Q & 0 \\ \cos R & \sin R & 0 \end{vmatrix} = 0 \times 0 = 0.$$

Example 3.6 Find all values of μ for which rank of the matrix

$$\mathbf{A} = \begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & \mu & -1 & 0 \\ 0 & 0 & \mu & -1 \\ -6 & 11 & -6 & 1 \end{bmatrix}$$

is equal to 3.

Solution Since the matrix $\mathbf{A}$ is of order 4, $r(\mathbf{A}) \leq 4$. Now, $r(\mathbf{A}) = 3$, if $|\mathbf{A}| = 0$ and there is at least one submatrix of order 3 whose determinant is not zero. Expanding the determinant through the elements of first row, we get

$$|\mathbf{A}| = \mu \begin{vmatrix} \mu & -1 & 0 \\ 0 & \mu & -1 \\ 11 & -6 & 1 \end{vmatrix} + \begin{vmatrix} 0 & -1 & 0 \\ 0 & \mu & -1 \\ -6 & -6 & 1 \end{vmatrix} = \mu[\mu(\mu-6)+11]-6$$

$$= \mu^3 - 6\mu^2 + 11\mu - 6 = (\mu-1)(\mu-2)(\mu-3).$$

Setting $|\mathbf{A}| = 0$, we obtain $\mu = 1, 2, 3$. For $\mu = 1, 2, 3$, the determinant of the leading third order submatrix

$$|\mathbf{A}_1| = \begin{vmatrix} \mu & -1 & 0 \\ 0 & \mu & -1 \\ 0 & 0 & \mu \end{vmatrix} = \mu^3 \neq 0.$$

Hence $r(\mathbf{A}) = 3$, when $\mu = 1$ or 2 or 3. For other values of μ, $r(\mathbf{A}) = 4$.

3.2.4 Inverse of a Square Matrix

Let $\mathbf{A} = (a_{ij})$ be a square matrix of order n. Then, $\mathbf{A}$ is called a

(i) *singular matrix* if $|\mathbf{A}| = 0$,
(ii) *non-singular matrix* if $|\mathbf{A}| \neq 0$.

In other words, a square matrix of order n is singular if its rank $r(\mathbf{A}) < n$ and non-singular if its rank $r(\mathbf{A}) = n$. A square non-singular matrix $\mathbf{A}$ of order n is said to be *invertible*, if there exists a non-singular square matrix $\mathbf{B}$ of order n such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I} \tag{3.11}$$

where $\mathbf{I}$ is an identity matrix of order n. The matrix $\mathbf{B}$ is called the *inverse matrix* of $\mathbf{A}$ and we write $\mathbf{B} = \mathbf{A}^{-1}$ or $\mathbf{A} = \mathbf{B}^{-1}$. Hence, we say that $\mathbf{A}^{-1}$ is the inverse of the matrix $\mathbf{A}$, if

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}. \tag{3.12}$$

The inverse $\mathbf{A}^{-1}$ of the matrix $\mathbf{A}$ is given by

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} adj(\mathbf{A}) \tag{3.13}$$

where $adj(\mathbf{A})$ = adjoint matrix of $\mathbf{A}$
= transpose of the matrix of cofactors of $\mathbf{A}$.

Remark 6

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

This result can be easily proved. We have

$$(\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{I}.$$

Premultiplying both sides first by $\mathbf{A}^{-1}$ and then by $\mathbf{B}^{-1}$ we obtain

$$\mathbf{B}^{-1}\mathbf{A}^{-1}(\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B}(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad \text{or} \quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

In general, we have $(\mathbf{A}_1\mathbf{A}_2 \ldots \mathbf{A}_p)^{-1} = \mathbf{A}_p^{-1}\mathbf{A}_{p-1}^{-1} \ldots \mathbf{A}_1^{-1}$.

(b) If $\mathbf{A}$ and $\mathbf{B}$ are non-singular matrices, then $\mathbf{AB}$ is also a non-singular matrix.

(c) If $\mathbf{AB} = \mathbf{0}$ and $\mathbf{A}$ is a non-singular matrix, then $\mathbf{B}$ must be a null matrix, since $\mathbf{AB} = \mathbf{0}$ can be premultiplied by $\mathbf{A}^{-1}$. If $\mathbf{B}$ is a non-singular matrix, then $\mathbf{A}$ must be a null matrix, since $\mathbf{AB} = \mathbf{0}$ can be post multiplied by $\mathbf{B}^{-1}$.

(d) If $\mathbf{AB} = \mathbf{AC}$ and $\mathbf{A}$ is a non-singular matrix, then $\mathbf{B} = \mathbf{C}$ (see Remark 1(d)).

(e) $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$, in general.

Properties of inverse matrices

1. If $\mathbf{A}^{-1}$ exists, then it is unique.
2. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
3. $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$, (From $(\mathbf{AA}^{-1})^T = \mathbf{I}^T = \mathbf{I}$, we get $(\mathbf{A}^{-1})^T\mathbf{A}^T = \mathbf{I}$. Hence, the result).
4. Let $\mathbf{D} = \text{diag}\,(d_{11}, d_{22}, \ldots, d_{nn})$, $d_{ii} \neq 0$. Then $\mathbf{D}^{-1} = \text{diag}\,(1/d_{11}, 1/d_{22}, \ldots, 1/d_{nn})$.
5. The inverse of a non-singular upper or lower triangular matrix is respectively an upper or a lower triangular matrix.
6. The inverse of a non-singular symmetric matrix is a symmetric matrix.
7. $(\mathbf{A}^{-1})^n = \mathbf{A}^{-n}$ for any positive integer n.

Example 3.7 Show that the matrix $\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$ satisfies the matrix equation $\mathbf{A}^3 - 6\mathbf{A}^2 + 11\mathbf{A} - \mathbf{I} = \mathbf{0}$ where $\mathbf{I}$ is an identity matrix of order 3. Hence, find the matrix (i) $\mathbf{A}^{-1}$ and (ii) $\mathbf{A}^{-2}$.

Solution We have

$$\mathbf{A}^2 = \begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & -1 & -5 \\ 15 & 1 & -5 \\ 5 & 4 & 9 \end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{A}^2\,\mathbf{A} = \begin{bmatrix} 4 & -1 & -5 \\ 15 & 1 & -5 \\ 5 & 4 & 9 \end{bmatrix}\begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -6 & -19 \\ 35 & -4 & -30 \\ 30 & 13 & 22 \end{bmatrix}.$$

Substituting in $\mathbf{B} = \mathbf{A}^3 - 6\mathbf{A}^2 + 11\mathbf{A} - \mathbf{I}$, we get

$$\mathbf{B} = \begin{bmatrix} 3 & -6 & -19 \\ 35 & -4 & -30 \\ 30 & 13 & 22 \end{bmatrix} - \begin{bmatrix} 24 & -6 & -30 \\ 90 & 6 & -30 \\ 30 & 24 & 54 \end{bmatrix} + \begin{bmatrix} 22 & 0 & -11 \\ 55 & 11 & 0 \\ 0 & 11 & 33 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}.$$

(i) Premultiplying $\mathbf{A}^3 - 6\mathbf{A}^2 + 11\mathbf{A} - \mathbf{I} = \mathbf{0}$ by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}\mathbf{A}^3 - 6\mathbf{A}^{-1}\mathbf{A}^2 + 11\mathbf{A}^{-1}\mathbf{A} - \mathbf{A}^{-1} = \mathbf{0}$$

or $\qquad \mathbf{A}^{-1} = \mathbf{A}^2 - 6\mathbf{A} + 11\,\mathbf{I}$

$$= \begin{bmatrix} 4 & -1 & -5 \\ 15 & 1 & -5 \\ 5 & 4 & 9 \end{bmatrix} - \begin{bmatrix} 12 & 0 & -6 \\ 30 & 6 & 0 \\ 0 & 6 & 18 \end{bmatrix} + \begin{bmatrix} 11 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 11 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}.$$

(ii) $$\mathbf{A}^{-2} = (\mathbf{A}^{-1})^2 = \begin{bmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}\begin{bmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 29 & -11 & 10 \\ -160 & 61 & -55 \\ 55 & -21 & 19 \end{bmatrix}.$$

3.2.5 Solution of $n \times n$ Linear System of Equations

Consider the system of n equations in n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n. \end{aligned} \tag{3.14}$$

In matrix form, we can write the system of equations (3.14) as

$$\mathbf{Ax} = \mathbf{b} \tag{3.15}$$

where $$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and **A**, **b**, **x** are respectively called the *coefficient matrix*, the right hand side column vector and th solution vector. If $\mathbf{b} \neq \mathbf{0}$, that is, at least one of the elements $b_1, b_2, \ldots, b_n$ is not zero, then the systen of equations is called *non-homogeneous*. If $\mathbf{b} = \mathbf{0}$, then the system of equations is called *homogeneous*. Th system of equations is called *consistent* if it has at least one solution and *inconsistent* if it has n solution.

Non-homogeneous system of equations

The non-homogeneous system of equations $\mathbf{Ax} = \mathbf{b}$ can be solved by the following methods.

Matrix method

Let **A** be non-singular. Premultiplying $\mathbf{Ax} = \mathbf{b}$ by $\mathbf{A}^{-1}$, we obtain

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{3.16}$$

The system of equations is consistent and has a unique solution. If $\mathbf{b} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$ (trivial solution) is the only solution.

Cramer's rule

Let $\mathbf{A}$ be non-singular. The Cramer's rule for the solution of $\mathbf{Ax} = \mathbf{b}$ is given by

$$x_i = \frac{|\mathbf{A}_i|}{|\mathbf{A}|}, \quad i = 1, 2, \ldots, n \tag{3.17}$$

where $|\mathbf{A}_i|$ is the determinant of the matrix $\mathbf{A}_i$ obtained by replacing the ith column of $\mathbf{A}$ by the right hand side column vector $\mathbf{b}$.

We discuss the following cases.

Case 1 When $|\mathbf{A}| \neq 0$, the system of equations is consistent and the unique solution is obtained by using Eq. (3.17).

Case 2 When $|\mathbf{A}| = 0$ and one or more of $|\mathbf{A}_i|$, $i = 1, 2, \ldots, n$ are not zero, then the system of equations has no solution, that is the system is inconsistent.

Case 3 When $|\mathbf{A}| = 0$ and all $|\mathbf{A}_i| = 0$, $i = 1, 2, \ldots, n$, then the system of equations is consistent and has infinite number of solutions. The system of equations has at least a one-parameter family of solutions.

Homogeneous system of equations

Consider the homogeneous system of equations

$$\mathbf{Ax} = \mathbf{0}. \tag{3.18}$$

Trivial solution $\mathbf{x} = \mathbf{0}$ is always a solution of this system.

If $\mathbf{A}$ is non-singular, then again $\mathbf{x} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$ is the solution.

Therefore, a homogeneous system of equations is always consistent. We conclude that non-trivial solutions for $\mathbf{Ax} = \mathbf{0}$ exist if and only if $\mathbf{A}$ is singular. In this case, the homogeneous system of equations has infinite number of solutions.

Example 3.8 Show that the system of equations

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$$

has a unique solution. Solve this system using (i) matrix method, (ii) Cramer's rule.

Solution We find that

$$|\mathbf{A}| = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & -3 \\ 1 & 1 & 1 \end{vmatrix} = 1(1+3) - 2(-1-1) + 1(3-1) = 10 \neq 0.$$

Therefore, the coefficient matrix $\mathbf{A}$ is non-singular and the given system of equations has a unique solution. Let $\mathbf{x} = [x, y, z]^T$.

(i) We obtain

$$\mathbf{A}^{-1} = \frac{1}{10}\begin{bmatrix} 4 & 2 & 2 \\ -5 & 0 & 5 \\ 1 & -2 & 3 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$$

Therefore, $$\mathbf{x} = \mathbf{A}^{-1}\,\mathbf{b} = \frac{1}{10}\begin{bmatrix} 4 & 2 & 2 \\ -5 & 0 & 5 \\ 1 & -2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

Hence, $x = 2$, $y = -1$ and $z = 1$.

(ii) We have

$$|\mathbf{A}_1| = \begin{vmatrix} 4 & -1 & 1 \\ 0 & 1 & -3 \\ 2 & 1 & 1 \end{vmatrix} = 4(1+3) - 0 + 2(3-1) = 20 \cdot$$

$$|\mathbf{A}_2| = \begin{vmatrix} 1 & 4 & 1 \\ 2 & 0 & -3 \\ 1 & 2 & 1 \end{vmatrix} = 1(0+6) - 2(4-2) + 1(-12-0) = -10 \cdot$$

$$|\mathbf{A}_3| = \begin{vmatrix} 1 & -1 & 4 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{vmatrix} = 1(2-0) - 2(-2-4) + 1(0-4) = 10 \cdot$$

Therefore, $x = \dfrac{|\mathbf{A}_1|}{|\mathbf{A}|} = 2, \quad y = \dfrac{|\mathbf{A}_2|}{|\mathbf{A}|} = -1, \quad z = \dfrac{|\mathbf{A}_3|}{|\mathbf{A}|} = 1.$

Example 3.9 Show that the system of equations

$$\begin{bmatrix} 1 & -1 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

has infinite number of solutions. Hence, find the solutions.

Solutions We find that

$$|\mathbf{A}| = \begin{vmatrix} 1 & -1 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 4 \end{vmatrix} = 0, \quad |\mathbf{A}_1| = \begin{vmatrix} 3 & -1 & 3 \\ 2 & 3 & 1 \\ 5 & 2 & 4 \end{vmatrix} = 0,$$

$$|\mathbf{A}_2| = \begin{vmatrix} 1 & 3 & 3 \\ 2 & 2 & 1 \\ 3 & 5 & 4 \end{vmatrix} = 0, \quad |\mathbf{A}_3| = \begin{vmatrix} 1 & -1 & 3 \\ 2 & 3 & 2 \\ 3 & 2 & 5 \end{vmatrix} = 0.$$

Therefore, the system of equations has infinite number of solutions. Using the first two equations

$$x_1 - x_2 = 3 - 3x_3$$

$$2x_1 + 3x_2 = 2 - x_3$$

and solving, we obtain $x_1 = (11 - 10\, x_3)/5$ and $x_2 = (5x_3 - 4)/5$ where x_3 is arbitrary. This solution satisfies the third equation.

Example 3.10 Show that the system of equations

$$\begin{bmatrix} 4 & 9 & 3 \\ 2 & 3 & 1 \\ 2 & 6 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 7 \end{bmatrix}$$

is inconsistent.

Solution We find that

$$|\mathbf{A}| = \begin{vmatrix} 4 & 9 & 3 \\ 2 & 3 & 1 \\ 2 & 6 & 2 \end{vmatrix} = 0, \quad |\mathbf{A}_1| = \begin{vmatrix} 6 & 9 & 3 \\ 2 & 3 & 1 \\ 7 & 6 & 2 \end{vmatrix} = 0,$$

$$|\mathbf{A}_2| = \begin{vmatrix} 4 & 6 & 3 \\ 2 & 2 & 1 \\ 2 & 7 & 2 \end{vmatrix} = 6.$$

Since $|\mathbf{A}| = 0$ and $|\mathbf{A}_2| \neq 0$, the system of equations is inconsistent.

Example 3.11 Solve the homogeneous system of equations

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & -2 \\ 4 & 7 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Solution We find that $|\mathbf{A}| = 0$. Hence, the given system has infinite number of solutions. Solving the first two equations

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3z \\ 2z \end{bmatrix}$$

we obtain $x = 13z$, $y = -8z$ where z is arbitrary. This solution satisfies the third equation.

Exercise 3.1

1. Given the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix}$, verify that

   (i) $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$, (ii) $|\mathbf{A} + \mathbf{B}| \neq |\mathbf{A}| + |\mathbf{B}|$.

2. If $\mathbf{A}^T = [1, -5, 7]$, $\mathbf{B} = [3, 1, 2]$, verify that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

3. Show that the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ satisfies the matrix equation $\mathbf{A}^2 - 4\mathbf{A} - 5\mathbf{I} = \mathbf{0}$. Hence, find $\mathbf{A}^{-1}$.

4. Show that the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \\ 2 & -1 & 3 \end{bmatrix}$ satisfies the matrix equation $\mathbf{A}^3 - 6\mathbf{A}^2 + 5\mathbf{A} + 11\mathbf{I} = \mathbf{0}$.

   Hence, find $\mathbf{A}^{-1}$.

5. For the matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$, verify that

   (i) $[adj\ (\mathbf{A})]^T = adj\ (\mathbf{A}^T)$, (ii) $[adj\ (\mathbf{A})]^{-1} = adj\ (\mathbf{A}^{-1})$.

6. For the matrix $\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 3 & 1 \\ 1 & 1 & 5 \end{bmatrix}$, verify that

   (i) $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$, (ii) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

7. For the matrices $\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 4 \\ 2 & 0 & 9 \end{bmatrix}$, verify that

   (i) $adj\ (\mathbf{AB}) = adj\ (\mathbf{A})\ adj\ (\mathbf{B})$, (ii) $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$.

8. For any non-singular matrix $\mathbf{A} = (a_{ij})$ of order n, show that

   (i) $|\ adj\ (\mathbf{A})\ | = |\ \mathbf{A}\ |^{n-1}$, (ii) $adj\ (adj\ (\mathbf{A})) = |\ \mathbf{A}\ |^{n-2}\mathbf{A}$.

9. For any non-singular matrix $\mathbf{A}$, show that $|\ \mathbf{A}^{-1}\ | = 1/|\ \mathbf{A}\ |$.

10. For any symmetric matrix $\mathbf{A}$, show that $\mathbf{BAB}^T$ is symmetric, where $\mathbf{B}$ is any matrix for which the product matrix $\mathbf{BAB}^T$ is defined.

11. If $\mathbf{A}$ is a symmetric matrix, prove that $(\mathbf{BA}^{-1})^T(\mathbf{A}^{-1}\mathbf{B}^T)^{-1} = \mathbf{I}$ where $\mathbf{B}$ is any matrix for which the product matrices are defined.

12. If $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices, then prove that

    (i) $\mathbf{A} + \mathbf{B}$ is symmetric, (ii) $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$ are both symmetric,

    (iii) $\mathbf{AB} - \mathbf{BA}$ is skew-symmetric.

13. If $\mathbf{A}$ and $\mathbf{B}$ are non-singular, commutative and symmetric matrices, then prove that

    (i) $\mathbf{AB}^{-1}$, (ii) $\mathbf{A}^{-1}\mathbf{B}$, (iii) $\mathbf{A}^{-1}\mathbf{B}^{-1}$

    are symmetric.

14. Let $\mathbf{A}$ be a non-singular matrix. Show that

    (i) if $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^n = \mathbf{0}$, then $\mathbf{A}^{-1} = \mathbf{A}^n$,

    (ii) if $\mathbf{I} - \mathbf{A} + \mathbf{A}^2 - \ldots + (-1)^n\mathbf{A}^n = \mathbf{0}$, then $\mathbf{A}^{-1} = (-1)^{n-1}\mathbf{A}^n$.

15. Let $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{A}$ be non-singular square matrices of order n and $\mathbf{PAQ} = \mathbf{I}$, then show that $\mathbf{A}^{-1} = \mathbf{QP}$.

16. If $\mathbf{I} - \mathbf{A}$ is a non-singular matrix, then show that

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots$$

    assuming that the series on the right hand side converges.

17. For any three non-singular matrices **A**, **B**, **C**, each of order n, show that $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Establish the following identities:

18. $\begin{vmatrix} 1 & w & w^2 \\ w & w^2 & 1 \\ w^2 & 1 & w \end{vmatrix} = 0$, where w is a cube root of unity.

19. $\begin{vmatrix} 1 & bc & a(b+c) \\ 1 & ac & b(a+c) \\ 1 & ab & c(a+b) \end{vmatrix} = 0.$

20. $\begin{vmatrix} a-b & b-c & c-a \\ b-c & c-a & a-b \\ c-a & a-b & b-c \end{vmatrix} = 0.$

21. $\begin{vmatrix} bc & a^2 & a^2 \\ b^2 & ac & b^2 \\ c^2 & c^2 & ab \end{vmatrix} = \begin{vmatrix} bc & ab & ac \\ ab & ac & bc \\ ac & bc & ab \end{vmatrix}.$

22. $\begin{vmatrix} -a^2 & ab & ac \\ ab & -b^2 & bc \\ ac & bc & -c^2 \end{vmatrix} = 4a^2b^2c^2.$

23. $\begin{vmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (a+c)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{vmatrix} = 2abc(a+b+c)^3.$

24. $\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ b+c & c+a & a+b \end{vmatrix} = (a-b)(b-c)(c-a)(a+b+c).$

25. $\begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}^2 = (x^3+y^3+z^3-3xyz)^2.$

26. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 \end{vmatrix} = (\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta).$

27. $\begin{vmatrix} \sin(a+\alpha) & \sin(b+\alpha) & \sin(c+\alpha) \\ \sin(a+\beta) & \sin(b+\beta) & \sin(c+\beta) \\ \sin(a+\gamma) & \sin(b+\gamma) & \sin(c+\gamma) \end{vmatrix} = 0$ for all a, b, c, α, β and γ.

Solve the following system of equations:

28. $x - y + z = 2,\ 2x + 3y - z = 5,\ x + y - z = 0.$

29. $x + 2y + 3z = 6,\ 2x + 4y + z = 7,\ 3x + 2y + 9z = 14.$

30. $-x + y + 2z = 2,\ 3x - y + z = 3,\ -x + 3y + 4z = 6.$

31. $2x - z = 1,\ 5x + y = 7,\ y + 3z = 5.$

32. Determine the values of k for which the system of equations

$$x - ky + z = 0, \quad kx + 3y - kz = 0, \quad 3x + y - z = 0$$

has (i) only trivial solution, (ii) non-trivial solution.

33. Find the value of θ for which the system of equations

$$2(\sin\theta)x + y - 2z = 0,\quad 3x + 2(\cos 2\theta)y + 3z = 0,\quad 5x + 3y - z = 0$$

has a non-trivial solution.

34. If the system of equations $x + ay + az = 0$, $bx + y + bz = 0$, $cx + cy + z = 0$, where a, b, c are non-zero and non-unity, has a non-trivial solution, then show that

$$\frac{a}{1-a} + \frac{b}{1-b} + \frac{c}{1-c} = -1.$$

35. Find the values of λ and μ for which the system of equations

$$x + 2y + z = 6,\quad x + 4y + 3z = 10,\quad x + 4y + \lambda z = \mu$$

has a (i) unique solution, (ii) infinite number of solutions, (iii) no solution.

Find the rank of the matrix **A**, where **A** is given by

36. $\begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 7 \end{bmatrix}$.

37. $\begin{bmatrix} 1 & 3 & -4 \\ -1 & -3 & 4 \\ 2 & 6 & -8 \end{bmatrix}$.

38. $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \\ 5 & 4 & -5 \end{bmatrix}$.

39. $\begin{bmatrix} 1 & 1 & 1 \\ p & q & r \\ p^3 & q^3 & r^3 \end{bmatrix}$.

40. (a) $\begin{bmatrix} 2 & 1 & 5 & -1 \\ -1 & 2 & 5 & 3 \\ 3 & 2 & 9 & -1 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & c_1 & -b_1 & a_2 \\ -c_1 & 0 & a_1 & b_2 \\ b_1 & -a_1 & 0 & c_2 \\ -a_2 & -b_2 & -c_2 & 0 \end{bmatrix}$, $a_i, b_i, c_i \neq 0$, $i = 1, 2$.

41. Prove that if **A** is an Hermitian matrix, then i**A** is a skew-Hermitian matrix and if **A** is a skew-Hermitian matrix, then i**A** is an Hermitian matrix.

42. Prove that if **A** is a real matrix and $\mathbf{A}^n \to \mathbf{0}$ as $n \to \infty$, then $\mathbf{I} + \mathbf{A}$ is invertible.

43. Let **A**, **B** be $n \times n$ real matrices. Then, show that

    (i) Trace $(\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha$ Trace $(\mathbf{A}) + \beta$ Trace $(\mathbf{B})$ for any scalars α and β,

    (ii) Trace $(\mathbf{AB})$ = Trace $(\mathbf{BA})$, (iii) $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$ is never true.

44. If **B**, **C** are $n \times n$ matrices, $\mathbf{A} = \mathbf{B} + \mathbf{C}$, $\mathbf{BC} = \mathbf{CB}$ and $\mathbf{C}^2 = \mathbf{0}$, then show that $\mathbf{A}^{p+1} = \mathbf{B}^p[\mathbf{B} + (p+1)\mathbf{C}]$ for any positive integer p.

45. Let $\mathbf{A} = (a_{ij})$ be a square matrix of order n, such that $a_{ij} = d$, $i \neq j$ and $a_{ij} = c$, $i = j$. Then show that $|\mathbf{A}| = (c-d)^{n-1}[c + (n-1)d]$.

Identify the following matrices as symmetric, skew-symmetric, Hermitian, skew-Hermitian or none of these:

46. $\begin{bmatrix} 1 & 2 & 3 \\ -2 & 5 & 4 \\ -3 & -4 & 6 \end{bmatrix}$.

47. $\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$.

48. $\begin{bmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{bmatrix}$.

49. $\begin{bmatrix} 1 & 2+4i & 1-i \\ 2-4i & -5 & 3-5i \\ 1+i & 3+5i & 6 \end{bmatrix}$.

50. $\begin{bmatrix} 1 & 2+4i & 1-i \\ -2+4i & -5 & 3-5i \\ -1-i & -3-5i & 6 \end{bmatrix}$.

51. $\begin{bmatrix} 0 & 2+4i & 1-i \\ -2+4i & 0 & 3-5i \\ -1-i & -3-5i & 0 \end{bmatrix}$.

52. $\begin{bmatrix} 0 & i & i \\ i & 0 & i \\ i & i & 0 \end{bmatrix}$.

53. $\begin{bmatrix} 0 & -i & 1+i \\ -i & -2i & 0 \\ -1+i & 0 & i \end{bmatrix}$.

54. $\begin{bmatrix} 1 & -1 & i \\ -1 & 0 & 1-i \\ -i & 1+i & 2 \end{bmatrix}$.

55. $\begin{bmatrix} 1 & 2i & -i \\ -2i & i & 1 \\ i & 1 & 2 \end{bmatrix}$.

3.3 Vector Spaces

Let V be a non-empty set of certain objects, which may be vectors, matrices, functions or some other objects. Each object is an element of V and is called a vector. The elements of V are denoted by **a**, **b**, **c**, **u**, **v**, etc. Assume that the two algebraic operations

(i) vector addition and (ii) scalar multiplication

are defined on elements of V.

If the vector addition is defined as the usual addition of vectors, then

$$\mathbf{a} + \mathbf{b} = (a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

If the scalar multiplication is defined as the usual scalar multiplication of a vector by the scaler α, then

$$\alpha\mathbf{a} = \alpha(a_1, a_2, \ldots, a_n) = (\alpha a_1, \alpha a_2, \ldots, \alpha a_n).$$

The set V defines a vector space if for any elements **a**, **b**, **c** in V and any scalars α, β the following properties (axioms) are satisfied.

Properties (axioms) with respect to vector addition

1. $\mathbf{a} + \mathbf{b}$ is in V.
2. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$. (commutative law)
3. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$. (associative law)
4. $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$. (existence of a unique zero element in V)
5. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$. (existence of additive inverse or negative vector in V)

Properties (axioms) with respect to scalar multiplication

6. $\alpha\mathbf{a}$ is in V.
7. $(\alpha + \beta)\,\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$. (left distributive law)
8. $(\alpha\beta)\,\mathbf{a} = \alpha(\beta\mathbf{a})$.
9. $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$. (right distributive law)
10. $1\mathbf{a} = \mathbf{a}$. (existence of multiplicative inverse)

The properties defined in **1** and **6** are called the *closure* properties. When these two properties are satisfied, we say that the vector space is closed under the vector addition and scalar multiplication. The vector addition and scalar multiplication defined above need not always be the usual addition and multiplication operators. Thus, *the vector space depends not only on the set V of vectors, but also on the definition of vector addition and scalar multiplication on V.*

If the elements of V are real, then it is called a *real vector space* when the scalars α, β are real numbers, whereas V is called a *complex vector space*, if the elements of V are complex and the scalars α, β may be real or complex numbers or if the elements of V are real and the scalars α, β are complex numbers.

Remark 7

(a) If even one of the above properties is not satisfied, then V is not a vector space. We usually check the closure properties first before checking the other properties.

(b) The concepts of length, dot product, vector product etc. are not part of the properties to be satisfied.

(c) The set of real numbers and complex numbers are called *fields* of scalars. We shall consider vector spaces only on the fields of scalars. In an advanced course on linear algebra, vector spaces over arbitrary fields are considered.

(d) The vector space $V = \{\mathbf{0}\}$ is called a trivial vector space.

The following are some examples of vector spaces under the usual operations of vector addition and scalar multiplication.

1. The set V of real or complex numbers.
2. The set of real valued continuous functions f on any closed interval $[a, b]$. The **0** vector defined in property **4** is the zero function.
3. The set of polynomials P_n of degree less than or equal to n.
4. The set V of n-tuples in $\mathbb{R}^n$ or $\mathbb{C}^n$.
5. The set V of all $m \times n$ matrices. The element **0** defined in property **4** is the null matrix of order $m \times n$.

The following are some examples which are not vector spaces. Assume that usual operations of vector addition and scalar multiplication are being used.

1. The set V of all polynomials of degree n. Let P_n and Q_n be two polynomials of degree n in V. Then, $\alpha P_n + \beta Q_n$ need not be a polynomial of degree n and thus may not be in V. For example, if $P_n = x^n + a_1x^{n-1} + \ldots + a_n$ and $Q_n = -x^n + b_1x^{n-1} + \ldots + b_n$, then $P_n + Q_n$ is a polynomial of degree $(n - 1)$.
2. The set V of all real-valued functions of one variable x, defined and continuous on the closed interval $[a, b]$ such that the value of the function at b is some non-zero constant p. For example, let $f(x)$ and $g(x)$ be two elements in V. Now, $f(b) = g(b) = p$. Since $f(b) + g(b) = 2p$, $f(x) + g(x)$ is not in V. Note that if $p = 0$, then V forms a vector space.

Example 3.12 Let V be the set of all polynomials, with real coefficients, of degree n, where addition is defined by $\mathbf{a} + \mathbf{b} = \mathbf{ab}$ and under usual scalar multiplication. Show that V is not a vector space.

Solution Let P_n and Q_n be two elements in V. Now, $P_n + Q_n = (P_n)(Q_n)$ is a polynomial of degree $2n$, which is not in V. Therefore, V does not define a vector space.

Example 3.13 Let V be the set of all ordered pairs (x, y), where x, y are real numbers. Let $\mathbf{a} = (x_1, y_1)$ and $\mathbf{b} = (x_2, y_2)$ be two elements in V. Define the addition as

$$\mathbf{a} + \mathbf{b} = (x_1, y_1) + (x_2, y_2) = (2x_1 - 3x_2, y_1 - y_2)$$

and the scalar multiplication as

$$\alpha(x_1, y_1) = (\alpha x_1/3, \alpha y_1/3).$$

Show that V is not a vector space. Which of the properties are not satisfied?

Solution We illustrate the properties that are not satisfied.

(i) $(x_2, y_2) + (x_1, y_1) = (2x_2 - 3x_1, y_2 - y_1) \neq (x_1, y_1) + (x_2, y_2)$.

Therefore, property **2** (commutative law) does not hold.

(ii)
$$\begin{aligned}((x_1, y_1) + (x_2, y_2)) + (x_3, y_3) &= (2x_1 - 3x_2, y_1 - y_2) + (x_3, y_3)\\ &= (4x_1 - 6x_2 - 3x_3, y_1 - y_2 - y_3)\\ (x_1, y_1) + ((x_2, y_2) + (x_3, y_3)) &= (x_1, y_1) + (2x_2 - 3x_3, y_2 - y_3)\\ &= (2x_1 - 6x_2 + 9x_3, y_1 - y_2 + y_3).\end{aligned}$$

Therefore, property **3** (associative law) is not satisfied.

(iii) $1(x_1, y_1) = (x_1/3, y_1/3) \neq (x_1, y_1)$.

Therefore, property **10** (existence of multiplicative inverse) is not satisfied.

Hence, V is not a vector space.

Example 3.14 Let V be the set of all ordered pairs (x, y), where x, y are real numbers. Let $\mathbf{a} = (x_1, y_1)$ and $\mathbf{b} = (x_2, y_2)$ be two elements in V. Define the addition as

$$\mathbf{a} + \mathbf{b} = (x_1, y_1) + (x_2, y_2) = (x_1 x_2, y_1 y_2)$$

and the scalar multiplication as

$$\alpha(x_1, y_1) = (\alpha x_1, \alpha y_1).$$

Show that V is not a vector space. Which of the properties are not satisfied?

Solution Note that $(1, 1)$ is an element of V. From the given definition of vector addition, we find that

$$(x_1, y_1) + (1, 1) = (x_1, y_1).$$

and this is true only for the element $(1, 1)$. Therefore, the element $(1, 1)$ plays the role of $\mathbf{0}$ element as defined in property **4**. Now, there is no element in V for which $(\mathbf{a}) + (-\mathbf{a}) = \mathbf{0} = (1, 1)$, since

$$(x_1, y_1) + (-x_1, -y_1) = (-x_1^2, -y_1^2) \neq (1, 1).$$

Therefore, property **5** is not satisfied.

Now, let $\alpha = 1, \beta = 2$ be any two scalars. We have

$$(\alpha + \beta)(x_1, y_1) = 3(x_1, y_1) = (3x_1, 3y_1)$$

and $$\alpha(x_1, y_1) + \beta(x_1, y_1) = 1(x_1, y_1) + 2(x_1, y_1) = (x_1, y_1) + (2x_1, 2y_1) = (2x_1^2, 2y_1^2)$$

Therefore, $(\alpha + \beta)(x_1, y_1) \neq \alpha(x_1, y_1) + \beta(x_1, y_1)$ and property **7** is not satisfied.

Similarly, it can be shown that property **9** is not satisfied. Hence, V is not a vector space.

3.3.1 Subspaces

Let V be an arbitrary vector space defined under a given vector addition and scalar multiplication. A non-empty subset W of V, such that W is also a vector space under the same two operations of vector addition and scalar multiplication, is called a *subspace* of V. Thus, W is also closed under the two given algebraic operations on V. As a convention, the vector space V is also taken as a subspace of V.

Remark 8

To show that W is a subspace of a vector space V, it is not necessary to verify all the 10 properties as given in section 3.3. If it is shown that W is closed under the given definition of vector addition and scalar multiplication, then the properties **2, 3, 7, 8, 9** and **10** are automatically satisfied because these properties are valid for all elements in V and hence are also valid for all elements in W. Thus, we need to verify the remaining properties, that is, the existence of the zero element and the additive inverse in W.

Consider the following examples:

1. Let V be the set of n-tuples $(x_1\ x_2 \ldots x_n)$ in $\mathbb{R}^n$ with usual addition and scalar multiplication. Then

   (i) W consisting of n-tuples $(x_1\ \ x_2\ \ \ldots\ \ x_n)$ with $x_1 = 0$ is a subspace of V.

   (ii) W consisting of n-tuples $(x_1\ \ x_2\ \ \ldots\ \ x_n)$ with $x_1 \geq 0$ is not a subspace of V, since W is not closed under scalar multiplication ($\alpha\mathbf{x}$, when α is a negative real number, is not in W).

   (iii) W consisting of n-tuples $(x_1\ \ x_2\ \ \ldots\ \ x_n)$ with $x_2 = x_1 + 1$ is not a subspace of V, since W is not closed under addition.
   (Let $\mathbf{x} = (x_1\ \ x_2\ \ \ldots\ \ x_n)$ with $x_2 = x_1 + 1$ and $\mathbf{y} = (y_1\ \ y_2\ \ \ldots\ \ y_n)$ with $y_2 = y_1 + 1$ be two elements in W. Then
   $$\mathbf{x} + \mathbf{y} = (x_1 + y_1\ .x_2 + y_2\ \ \ldots\ \ x_n + y_n)$$
   is not in W as $x_2 + y_2 = x_1 + y_1 + 2 \neq x_1 + y_1 + 1$).

2. Let V be the set of all real polynomials P of degree $\leq m$ with usual addition and scalar multiplication. Then

   (i) W consisting of all real polynomials of degree $\leq m$ with $P(0) = 0$ is a subspace of V.

   (ii) W consisting of all real polynomials of degree $\leq m$ with $P(0) = 1$ is not a subspace of V, since W is not closed under addition (If P and $Q \in W$, then $P + Q \notin W$).

   (iii) W consisting of all polynomials of degree $\leq m$ with real positive coefficients is not a subspace of V since W is not closed under scalar multiplication (If P is an element of W, then $-P \notin W$).

3. Let V be the set of all $n \times n$ real square matrices with usual matrix addition and scalar multiplication. Then

(i) W consisting of all symmetric/skew-symmetric matrices of order n is a subspace of V.

(ii) W consisting of all upper/lower triangular matrices of order n is a subspace of V.

(iii) W consisting of all $n \times n$ matrices having real positive elements is not a subspace of V since W is not closed under scalar multiplication (if $\mathbf{A}$ is an element of W, then $-\mathbf{A} \notin W$).

4. Let V be the set of all $n \times n$ complex matrices with usual matrix addition and scalar multiplication. Then

(i) W consisting of all Hermitian matrices of order n forms a vector space when scalars are real numbers and does not form a vector space when scalars are complex numbers (W is not closed under scalar multiplication).

Let $$\mathbf{A} = \begin{pmatrix} a & x + iy \\ x - iy & b \end{pmatrix} \in W.$$

Let $\alpha = i$. We get $\alpha \mathbf{A} = i\mathbf{A} = \begin{pmatrix} ai & xi - y \\ xi + y & bi \end{pmatrix} \notin W.$

(ii) W consisting of all skew-Hermitian matrices of order n forms a vector space when scalars are real numbers and does not form a vector space when scalars are complex numbers.

Let $$\mathbf{A} = \begin{pmatrix} i & x + iy \\ -x + iy & 2i \end{pmatrix} \in W.$$

Let $\alpha = i$. We get $i\mathbf{A} - \begin{pmatrix} -1 & ix - y \\ -ix - y & -2 \end{pmatrix} \notin W.$

Example 3.15 Let F and G be subspaces of a vector space V such that $F \cap G = \{\mathbf{0}\}$. The *sum* of F and G is written as $F + G$ and is defined by

$$F + G = \{\, \mathbf{f} + \mathbf{g} : \mathbf{f} \in F, \mathbf{g} \in G\}.$$

Show that $F + G$ is a subspace of V assuming the usual definition of vector addition and scalar multiplication.

Solution Let $W = F + G$ and $\mathbf{f} \in F$, $\mathbf{g} \in G$. Since $\mathbf{0} \in F$ and $\mathbf{0} \in G$, we have $\mathbf{0} + \mathbf{0} = \mathbf{0} \in W$. Let $\mathbf{f}_1 + \mathbf{g}_1$ and $\mathbf{f}_2 + \mathbf{g}_2$ belong to W where $\mathbf{f}_1, \mathbf{f}_2 \in F$ and $\mathbf{g}_1, \mathbf{g}_2 \in G$. Then

$$(\mathbf{f}_1 + \mathbf{g}_1) + (\mathbf{f}_2 + \mathbf{g}_2) = (\mathbf{f}_1 + \mathbf{f}_2) + (\mathbf{g}_1 + \mathbf{g}_2) \in F + G = W.$$

Also for any scalar α, $$\alpha(\mathbf{f} + \mathbf{g}) = \alpha \mathbf{f} + \alpha \mathbf{g} \in F + G = W.$$

Therefore, $W = F + G$ is a subspace of V.

We now present an important result on subspaces.

Theorem 3.1 Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ be any r elements of a vector space V under usual vector addition and scalar multiplication. Then, the set of all linear combinations of these elements, that is the set of all elements of the form

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_r \mathbf{v}_r \tag{3.19}$$

is a subspace of V, where $\alpha_1, \alpha_2, \ldots, \alpha_r$ are scalars.

Proof Let W be the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. Let

$$\mathbf{w}_1 = \sum_{i=1}^{r} a_i \mathbf{v}_i \quad \text{and} \quad \mathbf{w}_2 = \sum_{i=1}^{r} b_i \mathbf{v}_i$$

be any two linear combinations (any two elements of W). Then,

$$\mathbf{w}_1 + \mathbf{w}_2 = (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2 + \ldots + (a_r + b_r)\mathbf{v}_r \in W$$

$$\alpha \mathbf{w}_1 = (\alpha a_1)\mathbf{v}_1 + (\alpha a_2)\mathbf{v}_2 + \ldots + (\alpha a_r)\mathbf{v}_r \in W$$

$$\alpha \mathbf{w}_2 = (\alpha b_1)\mathbf{v}_1 + (\alpha b_2)\mathbf{v}_2 + \ldots + (\alpha b_r)\mathbf{v}_r \in W$$

and $\qquad \alpha(\mathbf{w}_1 + \mathbf{w}_2) = \alpha \mathbf{w}_1 + \alpha \mathbf{w}_2.$

Taking $\alpha = 0$, we find that $0\mathbf{w}_1 = \mathbf{0} \in W$. This implies that $\mathbf{w}_1 + \mathbf{0} = \mathbf{0} + \mathbf{w}_1 = \mathbf{w}_1$

Taking $\alpha = -1$, we find that $(-1)\mathbf{w}_1 = (-\mathbf{w}_1) \in W$. This implies that $\mathbf{w}_1 + (-\mathbf{w}_1) = \mathbf{0}$

Therefore, W is a subspace of V.

The elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are in the subspace W as

$$\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_r,\ \mathbf{v}_2 = 0\mathbf{v}_1 + 1\mathbf{v}_2 + \ldots + 0\mathbf{v}_r, \ldots$$

We say that the subspace W is *spanned* by the r elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. Also, any subspace that contains the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ must contain every linear combination of these elements.

Spanning set Let S be a subset of a vector space V and suppose that every element in V can be obtained as a linear combination of the elements taken from S. Then S is said to be the *spanning set* for V. We also say that S spans V.

Example 3.16 Let V be the vector space of all 2×2 real matrices. Show that the sets

(i) $$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

(ii) $$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$$

span V.

Solution Let $\mathbf{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an arbitrary element of V.

(i) We write

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since every element of V can be written as a linear combination of the elements of S, the set S spans the vector space V.

(ii) We need to determine the scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ so that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Equating the corresponding elements, we obtain the system of equations

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = a, \quad \alpha_2 + \alpha_3 + \alpha_4 = b,$$
$$\alpha_3 + \alpha_4 = c, \qquad \alpha_4 = d.$$

The solution of this system of equations is

$$\alpha_4 = d,\ \alpha_3 = c - d,\ \alpha_2 = b - c,\ \alpha_1 = a - b.$$

Therefore, we can write

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a-b) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + (b-c) \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + (c-d) \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since every element of V can be written as a linear combination of the elements of S, the set S spans the vector space V.

Example 3.17 Let V be the vector space of all polynomials of degree ≤ 3. Determine whether or not the set

$$S = \{t^3, t^2 + t, t^3 + t + 1\}$$

spans V?

Solution Let $P(t) = \alpha t^3 + \beta t^2 + \gamma t + \delta$ be an arbitrary element in V. We need to find whether or not there exist scalars a_1, a_2, a_3 such that

$$\alpha t^3 + \beta t^2 + \gamma t + \delta = a_1 t^3 + a_2 (t^2 + t) + a_3 (t^3 + t + 1)$$

or

$$\alpha t^3 + \beta t^2 + \gamma t + \delta = (a_1 + a_3)\, t^3 + a_2 t^2 + (a_2 + a_3)\, t + a_3.$$

Comparing the coefficients of various powers of t, we get

$$a_1 + a_3 = \alpha,\ a_2 = \beta,\ a_2 + a_3 = \gamma,\ a_3 = \delta.$$

The solution of the first three equations is given by

$$a_1 = \alpha + \beta - \gamma,\ a_2 = \beta,\ a_3 = \gamma - \beta.$$

Substituting in the last equation, we obtain $\gamma - \beta = \delta$, which may not be true for all elements in V. For example, the polynomial $t^3 + 2t^2 + t + 3$ does not satisfy this condition and therefore, it cannot be written as a linear combination of the elements of S. Therefore, S does not span the vector space V.

3.3.2 Linear Independence of Vectors

Let V be a vector space. A finite set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$ of the elements of V is said to be *linearly dependent* if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all zero, such that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0}. \tag{3.20}$$

If Eq. (3.20) is satisfied only for $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, then the set of vectors is said to be *linearly independent*.

The above definition of linear dependence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ can be written alternately as follows.

Theorem 3.2 The set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly dependent if and only if at least one element of the set is a linear combination of the remaining elements.

Proof Let the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be linearly dependent. Then, there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, not all zero such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_{i-1}\mathbf{v}_{i-1} + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}.$$

Let $\alpha_i \neq 0$. Then, we can write

$$\mathbf{v}_i = -\left(\frac{\alpha_1}{\alpha_i}\right)\mathbf{v}_1 - \left(\frac{\alpha_2}{\alpha_i}\right)\mathbf{v}_2 - \ldots - \left(\frac{\alpha_{i-1}}{\alpha_i}\right)\mathbf{v}_{i-1} - \left(\frac{\alpha_{i+1}}{\alpha_i}\right)\mathbf{v}_{i+1} - \ldots - \left(\frac{\alpha_n}{\alpha_i}\right)\mathbf{v}_n$$

$$= \alpha_1^*\mathbf{v}_1 + \alpha_2^*\mathbf{v}_2 + \ldots + \alpha_{i-1}^*\mathbf{v}_{i-1} + \alpha_{i+1}^*\mathbf{v}_{i+1} + \ldots + \alpha_n^*\mathbf{v}_n$$

where $\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*$ are some scalars. Hence, the vector $\mathbf{v}_i$ is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n$.

Now let $\mathbf{v}_i$ be a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n$. Therefore, we have

$$\mathbf{v}_i = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_{i-1}\mathbf{v}_{i-1} + a_{i+1}\mathbf{v}_{i+1} + \ldots + a_n\mathbf{v}_n$$

where a_i's are scalars. Then

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_{i-1}\mathbf{v}_{i-1} + (-1)\,\mathbf{v}_i + a_{i+1}\mathbf{v}_{i+1} + \ldots + a_n\mathbf{v}_n = \mathbf{0}.$$

Since the coefficient of $\mathbf{v}_i$ is not zero, the elements are linearly dependent.

Remark 9

Eq. (3.20) gives a homogeneous system of algebraic equations. Non-trivial solutions exist if *det* (coefficient matrix) = 0, that is the vectors are linearly dependent in this case. If the *det* (coefficient matrix) $\neq 0$, then by Cramer's rule, $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ and the vectors are linearly independent.

Example 3.18 Let $\mathbf{v}_1 = (1, -1, 0)$, $\mathbf{v}_2 = (0, 1, -1)$ and $\mathbf{v}_3 = (0, 0, 1)$ be elements of $\mathbb{R}^3$. Show that the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

Solution We consider the vector equation

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}.$$

Substituting for $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, we obtain

$$\alpha_1(1, -1, 0) + \alpha_2(0, 1, -1) + \alpha_3(0, 0, 1) = \mathbf{0}$$

or

$$(\alpha_1, -\alpha_1 + \alpha_2, -\alpha_2 + \alpha_3) = \mathbf{0}.$$

Comparing, we obtain $\alpha_1 = 0$, $-\alpha_1 + \alpha_2 = 0$ and $-\alpha_2 + \alpha_3 = 0$. The solution of these equations is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore, the given set of vectors is linearly independent.

Alternative

$$det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \begin{vmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{vmatrix} = 1 \neq 0.$$

Therefore, the given vectors are linearly independent.

Example 3.19 Let $\mathbf{v}_1 = (1, -1, 0)$, $\mathbf{v}_2 = (0, 1, -1)$, $\mathbf{v}_3 = (0, 2, 1)$ and $\mathbf{v}_4 = (1, 0, 3)$ be elements of $\mathbb{R}^3$. Show that the set of vectors $\{\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \mathbf{v}_4\}$ is linearly dependent.

Solution The given set of elements will be linearly dependent if there exists scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, not all zero, such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 = \mathbf{0}. \tag{3.21}$$

Substituting for $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ and comparing, we obtain

$$\alpha_1 + \alpha_4 = 0, -\alpha_1 + \alpha_2 + 2\alpha_3 = 0, -\alpha_2 + \alpha_3 + 3\alpha_4 = 0.$$

The solution of this system of equations is

$$\alpha_1 = -\alpha_4,\ \alpha_2 = 5\alpha_4/3,\ \alpha_3 = -4\alpha_4/3,\ \alpha_4 \text{ arbitrary}.$$

Substituting in Eq. (3.21) and cancelling α_4, we obtain

$$-\mathbf{v}_1 + \frac{5}{3}\mathbf{v}_2 - \frac{4}{3}\mathbf{v}_3 + \mathbf{v}_4 = \mathbf{0}.$$

Hence, there exist scalars not all zero, such that Eq. (3.21) is satisfied. Therefore, the set of vectors is linearly dependent.

3.3.3 Dimension and Basis

Let V be a vector space. If for some positive integer n, there exists a set S of n linearly independent elements of V and if every set of $n + 1$ or more elements in V is linearly dependent, then V is said to have *dimension* n. Then, we write $dim(V) = n$. Thus, the maximum number of linearly independent elements of V is the dimension of V. The set S of n linearly independent vectors is called the *basis* of V. Note that a vector space whose only element is zero has dimension zero.

Theorem 3.3 Let V be a vector space of dimension n. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be the linearly independent elements of V. Then, every other element of V can be written as a linear combination of these elements. Further, this representation is unique.

Proof Let $\mathbf{v}$ be an element of V. Then, the set $\{\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent as it has $n + 1$ elements. Therefore, there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_n$, not all zero, such that

$$\alpha_0\mathbf{v} + \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}. \tag{3.22}$$

Now, $\alpha_0 \neq 0$. Because, if $\alpha_0 = 0$, we get $\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}$ and since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, we get $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. This implies that the set of $n + 1$ elements $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly independent, which is not possible as the dimension of V is n.

Therefore, we obtain from Eq. (3.22)

$$\mathbf{v} = \sum_{i=1}^{n} (-\alpha_i/\alpha_0)\mathbf{v}_i. \tag{3.23}$$

Hence, $\mathbf{v}$ is a linear combination of n linearly independent vectors of V.

Now, let there be two representations of $\mathbf{v}$ given by

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n \quad \text{and} \quad \mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_n\mathbf{v}_n$$

where $b_i \neq a_i$ for at least one i. Subtracting these two equations, we get

$$\mathbf{0} = (a_1 - b_1)\mathbf{v}_1 + (a_2 - b_2)\mathbf{v}_2 + \ldots + (a_n - b_n)\mathbf{v}_n.$$

Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, we get

$$a_i - b_i = 0 \quad \text{or} \quad a_i = b_i,\ i = 1, 2, \ldots, n.$$

Therefore, both the representations of $\mathbf{v}$ are same and the representation of $\mathbf{v}$ given by Eq. (3.23) is unique.

Remark 10

(a) A set of $(n + 1)$ vectors in $\mathbb{R}^n$ is linearly dependent.

(b) A set of vectors containing $\mathbf{0}$ as one of its elements is linearly dependent as $\mathbf{0}$ is the linear combination of any set of vectors.

Theorem 3.4 Let V be an n-dimensional vector space. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, $k < n$ are linearly independent elements of V, then there exist elements $\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n$ such that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of V.

Proof There exists an element $\mathbf{v}_{k+1}$ such that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ are linearly independent. Otherwise, every element of V can be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and therefore V has dimension $k < n$. This argument can be continued. If $n > k + 1$, we keep adding elements $\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n$ such that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of V.

Since all the elements of a vector space V of dimension n can be represented as linear combinations of the n elements in the basis of V, the basis of V spans V. However, there can be many basis for the same vector space. For example, consider the vector space $\mathbb{R}^3$. Each of the following set of vectors

(i) $[1, -1, 0], [0, 1, -1], [0, 0, 1]$

(ii) $[1, -1, 0], [0, 0, 1], [1, 2, 3]$

(iii) $[1, 0, 0], [0, 1, 0], [0, 0, 1]$

are linearly independent and therefore forms a basis in $\mathbb{R}^3$. Some of the standard basis are the following.

1. If V consists of n-tuples in $\mathbb{R}^n$, then

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 0, 1)$$

is called a standard basis in $\mathbb{R}^n$.

2. If V consists of all $m \times n$ matrices, then

$$\mathbf{E}_{rs} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}, r = 1, 2, \ldots, m \text{ and } s = 1, 2, \ldots, n$$

where 1 is located in the (r, s) location, that is in the rth row and the sth column, is called its standard basis.

For example, if V consists of all 2×3 matrices, then any matrix $\begin{bmatrix} a & b & c \\ x & y & z \end{bmatrix}$ in V can be written as

$$\begin{bmatrix} a & b & c \\ x & y & z \end{bmatrix} = a\mathbf{E}_{11} + b\mathbf{E}_{12} + c\mathbf{E}_{13} + x\mathbf{E}_{21} + y\mathbf{E}_{22} + z\mathbf{E}_{23}$$

where $\mathbf{E}_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{E}_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ etc.

3. If V consists of all polynomials $P(t)$ of degree $\leq n$, then $\{1, t, t^2, \ldots, t^n\}$ is taken as its standard basis.

Example 3.20 Determine whether the following set of vectors $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ forms a basis in $\mathbb{R}^3$, where

(i) $\mathbf{u} = (2, 2, 0)$, $\mathbf{v} = (3, 0, 2)$, $\mathbf{w} = (2, -2, 2)$

(ii) $\mathbf{u} = (0, 1, -1)$, $\mathbf{v} = (-1, 0, -1)$, $\mathbf{w} = (3, 1, 3)$.

Solution If the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ forms a basis in $\mathbb{R}^3$, then $\mathbf{u}, \mathbf{v}, \mathbf{w}$ must be linearly independent. Let $\alpha_1, \alpha_2, \alpha_3$ be scalars. Then, the only solution of the equation

$$\alpha_1\mathbf{u} + \alpha_2\mathbf{v} + \alpha_3\mathbf{w} = \mathbf{0} \tag{3.24}$$

must be $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

(i) Using Eq. (3.24), we obtain the system of equations

$$2\alpha_1 + 3\alpha_2 + 2\alpha_3 = 0,\ 2\alpha_1 - 2\alpha_3 = 0 \quad \text{and} \quad 2\alpha_2 + 2\alpha_3 = 0.$$

The solution of this system of equations is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore, $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are linearly independent and they form a basis in $\mathbb{R}^3$.

(ii) Using Eq. (3.24), we obtain the system of equations

$$-\alpha_2 + 3\alpha_3 = 0,\ \alpha_1 + \alpha_3 = 0,\ -\alpha_1 - \alpha_2 + 3\alpha_3 = 0.$$

The solution of this system of equations is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore, $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are linearly independent and they form a basis in $\mathbb{R}^3$.

Example 3.21 Find the dimension of the subspace of $\mathbb{R}^4$ spanned by the set $\{(1\ \ 0\ \ 0\ \ 0), (0\ \ 1\ \ 0\ \ 0), (1\ \ 2\ \ 0\ \ 1), (0\ \ 0\ \ 0\ \ 1)\}$. Hence find its basis.

Solution The dimension of the set is ≤ 4. If it is 4, then the only solution of the vector equation

$$\alpha_1(1\ \ 0\ \ 0\ \ 0) + \alpha_2(0\ \ 1\ \ 0\ \ 0) + \alpha_3(1\ \ 2\ \ 0\ \ 1) + \alpha_4(0\ \ 0\ \ 0\ \ 1) = \mathbf{0} \tag{3.25a}$$

should be $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. Comparing, we obtain the system of equations

$$\alpha_1 + \alpha_3 = 0,\ \alpha_2 + 2\alpha_3 = 0,\ \alpha_3 + \alpha_4 = 0.$$

The solution of this system of equations is given by

$$\alpha_1 = \alpha_4,\ \alpha_2 = 2\alpha_4,\ \alpha_3 = -\alpha_4, \text{ where } \alpha_4 \text{ is arbitrary.}$$

Hence, the vector equation (3.25a) is satisfied for non-zero values of $\alpha_1, \alpha_2, \alpha_3$ and α_4. Therefore, the dimension of the set is less than 4.

Now, consider any three elements of the set, say (1 0 0 0), (0 1 0 0) and (1 2 0 1). Consider the vector equation

$$\alpha_1(1 \quad 0 \quad 0 \quad 0) + \alpha_2(0 \quad 1 \quad 0 \quad 0) + \alpha_3(1 \quad 2 \quad 0 \quad 1) = \mathbf{0}. \tag{3.25b}$$

Comparing, we obtain the system of equations

$$\alpha_1 + \alpha_3 = 0,\ \alpha_2 + 2\alpha_3 = 0 \quad \text{and} \quad \alpha_3 = 0.$$

The solution of this system of equations is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Hence, these three elements are linearly independent. Therefore, the dimension of the given set is 3 and the basis is the set of vectors {(1 0 0 0), (0 1 0 0), (1 2 0 1)}. We find that the fourth vector can be written as

$$(0 \quad 0 \quad 0 \quad 1) = -(1 \quad 0 \quad 0 \quad 0) - 2\,(0 \quad 1 \quad 0 \quad 0) + 1(1 \quad 2 \quad 0 \quad 1).$$

Example 3.22 Let $\mathbf{u} = \{(a, b, c, d)$, such that $a + c + d = 0,\ b + d = 0\}$ be a subspace of $\mathbb{R}^4$. Find the dimension and the basis of the subspace.

Solution $\mathbf{u}$ satisfies the closure properties. From the given equations, we have

$$a + c + d = 0 \quad \text{and} \quad b + d = 0 \quad \text{or} \quad a = -c - d \quad \text{and} \quad b = -d.$$

We have two free parameters, say, c and d. Therefore, the dimension of the given subspace is 2. Choosing $c = 0, d = 1$ and $c = 1, d = 0$, we may write a basis as {(– 1 –1 0 1), (– 1 0 1 0)}.

3.3.4 Lnear Transformations

Let A and B be two arbitrary sets. A rule that assigns to elements of A exactly one element of B is called a *function* or a *mapping* or a *transformation*. Thus, a transformation maps the elements of A into the elements of B. The set A is called the *domain* of the transformation. We use capital letters T, S etc. to denote a transformation. If T is a transformation from A into B, we write

$$T : A \to B. \tag{3.26}$$

For each element $\mathbf{a} \in A$, we get a unique element $\mathbf{b} \in B$. We write $\mathbf{b} = T(\mathbf{a})$ or $\mathbf{b} = T\mathbf{a}$ and $\mathbf{b}$ is called the image of $\mathbf{a}$ under the mapping T. The collection of all such images in B is called the *range* or the image set of the transformation T.

In this section, we shall discuss mappings from a vector space into a vector space. let V and W be two vector spaces, both real or complex, over the same field F of scalars. Let T be a mapping from V into W. The mapping T is said to be a *linear transformation* or a *linear mapping*, if it satisfies the following two properties:

(i) For every scalar α and every element $\mathbf{v}$ in V

$$T(\alpha\mathbf{v}) = \alpha T(\mathbf{v}). \tag{3.27}$$

(ii) For any two elements $\mathbf{v}_1, \mathbf{v}_2$ in V

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2). \tag{3.28}$$

Since V is a vector space, the product $\alpha\mathbf{v}$ and the sum $\mathbf{v}_1 + \mathbf{v}_2$ are defined and are elements in V. Then, T defines a mapping from V into W. Since $T(\mathbf{v}_1)$ and $T(\mathbf{v}_2)$ are in W, the product $\alpha T(\mathbf{v})$ and the sum $T(\mathbf{v}_1) + T(\mathbf{v}_2)$ are in W. The conditions given in Eqs. (3.27) and (3.28) are equivalent to

$$T(\alpha\,\mathbf{v}_1 + \beta\,\mathbf{v}_2) = T(\alpha\,\mathbf{v}_1) + T(\alpha\,\mathbf{v}_2) = \alpha T(\mathbf{v}_1) + \beta T(\mathbf{v}_2)$$

for $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and any scalars α, β.

Let V be a vector space of dimension n and let the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be its basis. Then, any element $\mathbf{v}$ in V can be written as a linear combination of the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Therefore,

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars, not all zero. If T is a linear transformation defined in V, then

$$\begin{aligned} T(\mathbf{v}) &= T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n) \\ &= T(\alpha_1\mathbf{v}_1) + T(\alpha_2\mathbf{v}_2) + \ldots + T(\alpha_n\mathbf{v}_n) \\ &= \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \ldots + \alpha_n T(\mathbf{v}_n). \end{aligned}$$

Thus, a linear transformation is completely determind by its action on the basis vectors of a vector space.

Letting $\alpha = 0$ in Eq. (3.27), we find that for every element $\mathbf{v}$ in V

$$T(0\,\mathbf{v}) = T(\mathbf{0}) = 0T(\mathbf{v}) = \mathbf{0}.$$

Therefore, the zero element in V is mapped into zero element in W by the linear transformation T.

The collection of all elements $\mathbf{w} = T(\mathbf{v})$ is called the *range* of T and is written as *ran* (T). The set of all elements of V that are mapped into the zero element by the linear transformation T is called the *kernel* or the *null-space* of T and is denoted by *ker* (T). Therefore, we have

$$ker(T) = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{0}\} \quad \text{and} \quad ran(T) = \{T(\mathbf{v}) \mid \mathbf{v} \in V\}.$$

Thus, the null space of T is a subspace of V and the range of T is a subspace of W.

The dimension of *ran* (T) is called the rank (T) and the dimension of *ker* (T) is called the *nullity* of T. We have the following result.

Theorem 3.5 If T has rank r and the dimension of V is n, then the nullity of T is $n - r$, that is,

$$\text{rank}\,(T) + \text{nullity} = n = dim\,(V).$$

We shall discuss the linear transformation only in the context of matrices.

Let $\mathbf{A}$ be an $m \times n$ real (or complex) matrix. Let the rows of $\mathbf{A}$ represent the elements in $\mathbb{R}^n$ (or $\mathbb{C}^n$) and the columns of $\mathbf{A}$ represent the elements in $\mathbb{R}^m$ (or $\mathbb{C}^m$). If $\mathbf{x}$ is in $\mathbb{R}^n$, then $\mathbf{Ax}$ is in $\mathbb{R}^m$. Thus, an $m \times n$ matrix maps the elements in $\mathbb{R}^n$ into the elements in $\mathbb{R}^m$. We write

$$T = \mathbf{A} : \mathbb{R}^n \to \mathbb{R}^m, \quad \text{and} \quad T\mathbf{x} = \mathbf{Ax}.$$

We now prove that the mapping $\mathbf{A}$ is a linear transformation. Let $\mathbf{v}_1, \mathbf{v}_2$ be two elements in $\mathbb{R}^n$ and α, β be scalars. Then

$$T(\alpha\,\mathbf{v}_1) = \mathbf{A}(\alpha\,\mathbf{v}_1) = \alpha\mathbf{A}\mathbf{v}_1$$

$$T(\beta\,\mathbf{v}_2) = \mathbf{A}(\beta\,\mathbf{v}_2) = \beta\,\mathbf{A}\mathbf{v}_2$$

and

$$T(\alpha\,\mathbf{v}_1 + \beta\,\mathbf{v}_2) = \mathbf{A}(\alpha\,\mathbf{v}_1 + \beta\,\mathbf{v}_2) = \alpha\mathbf{A}\mathbf{v}_1 + \beta\mathbf{A}\mathbf{v}_2.$$

The range of T is a linear subspace of $\mathbb{R}^m$ and the kernel of T is a linear subspace of $\mathbb{R}^n$.

Sum and product of linear transformations

Let T_1 and T_2 be two linear transformations from V into W. We define the sum $T_1 + T_2$ to be the transformation S such that

$$S\mathbf{v} = T_1\mathbf{v} + T_2\mathbf{v}, \; \mathbf{v} \in V.$$

It can be easily verified that $T_1 + T_2$ is a linear transformation and $T_1 + T_2 = T_2 + T_1$.

Now, let U, V, W be three vector spaces, all real or all complex, on the same field of scalars. Let T_1 and T_2 be linear transformations such that

$$T_1 : U \to V \quad \text{and} \quad T_2 : V \to W.$$

The product T_2T_1 is defined to be the transformation S from U into W such that

$$\mathbf{w} = S\mathbf{u} = T_2(T_1\mathbf{u}),\ \mathbf{u} \in U.$$

The transformation T_2T_1 is also called a *composite* transformation (Fig. 3.1). The transformation T_2T_1 means applying first the transformation T_1 and then applying the transformation T_2. It can be easily verified that T_2T_1 is a linear transformation.

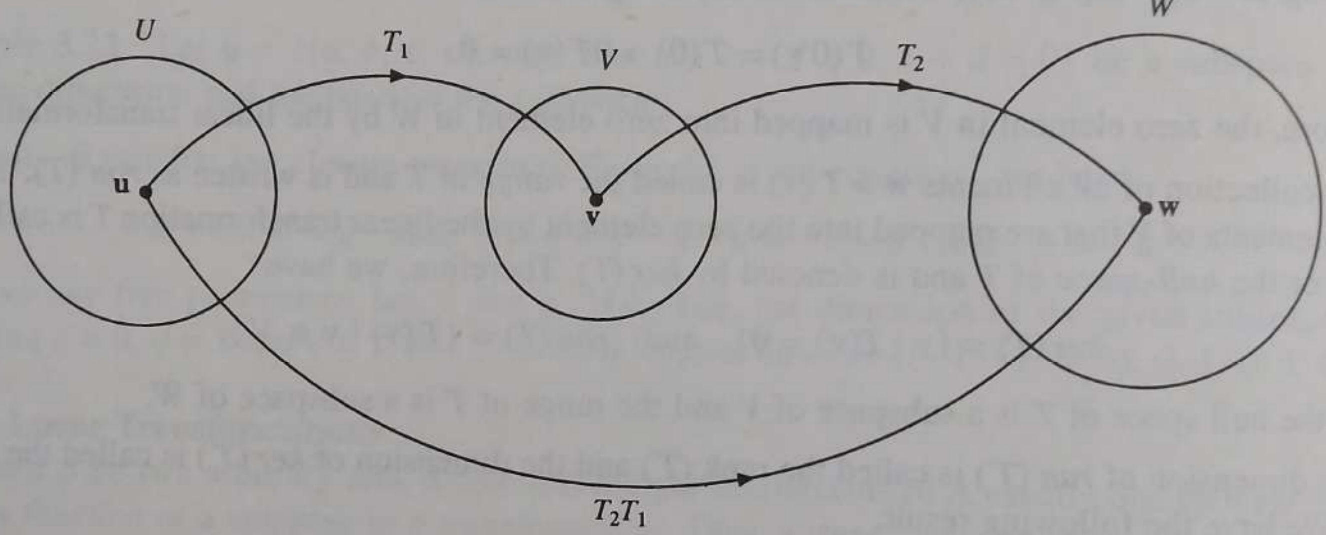

Fig. 3.1. Composite transformation.

If $T_1 : V \to V$ and $T_2 : V \to V$ are linear transformations, then both T_2T_1 and T_1T_2 are defined and map V into V. In general, $T_2T_1 \neq T_1T_2$. For example, let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ matrices and $\mathbf{x}$ be any element in $\mathbb{R}^n$. Let T_1 and T_2 be the transformations

$$T_1(\mathbf{x}) = \mathbf{Ax} \quad \text{and} \quad T_2(\mathbf{x}) = \mathbf{Bx}$$

from $\mathbb{R}^n$ into $\mathbb{R}^n$. Then

$$T_2(T_1(\mathbf{x})) = \mathbf{BAx} \quad \text{and} \quad T_1(T_2(\mathbf{x})) = \mathbf{ABx}.$$

Therefore, $T_2T_1 \neq T_1T_2$ unless the matrices $\mathbf{A}$ and $\mathbf{B}$ commute.

Example 3.23 Let T be a linear transformation from $\mathbb{R}^3$ into $\mathbb{R}^2$ defined by the relations

$$T\mathbf{x} = \mathbf{Ax},\ \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Find $T\mathbf{x}$ when $\mathbf{x}$ is given by $[3\ \ 4\ \ 5]^T$.

Solution We have

$$T\mathbf{x} = \mathbf{Ax} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 26 \\ 62 \end{bmatrix}.$$

Example 3.24 Let T be a linear transformation defined by

$$T\left[\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\right]=\begin{pmatrix}1\\2\\3\end{pmatrix}, T\left[\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}\right]=\begin{pmatrix}1\\-2\\3\end{pmatrix}, T\left[\begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}\right]=\begin{pmatrix}1\\-2\\-3\end{pmatrix}, T\left[\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right]=\begin{pmatrix}-1\\2\\3\end{pmatrix}.$$

Find $T\left[\begin{pmatrix}4 & 5\\3 & 8\end{pmatrix}\right]$.

Solution The matrices $\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$ are linearly independent and hence form a basis in the space of 2×2 matrices. We write for any scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, not all zero

$$\begin{pmatrix}4 & 5\\3 & 8\end{pmatrix}=\alpha_1\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}+\alpha_2\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}+\alpha_3\begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}+\alpha_4\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$$

$$=\begin{bmatrix}\alpha_1 & \alpha_1+\alpha_2\\ \alpha_1+\alpha_2+\alpha_3 & \alpha_1+\alpha_2+\alpha_3+\alpha_4\end{bmatrix}.$$

Comparing the elements and solving the resulting system of equations, we get $\alpha_1 = 4$, $\alpha_2 = 1$, $\alpha_3 = -2$, $\alpha_4 = 5$. Since T is a linear transformation, we get

$$T\left[\begin{pmatrix}4 & 5\\3 & 8\end{pmatrix}\right]=\alpha_1 T\left[\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\right]+\alpha_2 T\left[\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}\right]+\alpha_3 T\left[\begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}\right]+\alpha_4 T\left[\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right]$$

$$=4\begin{pmatrix}1\\2\\3\end{pmatrix}+1\begin{pmatrix}1\\-2\\3\end{pmatrix}-2\begin{pmatrix}1\\-2\\-3\end{pmatrix}+5\begin{pmatrix}-1\\2\\3\end{pmatrix}=\begin{pmatrix}-2\\20\\36\end{pmatrix}.$$

Example 3.25 Let T be a linear transformation defined by

$$T\mathbf{x} = \mathbf{A}\mathbf{x}, \quad \text{where} \quad \mathbf{A}=\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}, \mathbf{x}=(x_1, x_2)^T.$$

Find all points, if any, that are mapped into the point (3, 2).

Solution Let $(y_1, y_2)^T$ be the point that is mapped into (3, 2). Therefore, we have

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix}=\begin{bmatrix}3\\2\end{bmatrix}.$$

Multiplying and comparing we obtain the system of equations $y_1 + 2y_2 = 3$, $3y_1 + 4y_2 = 2$. The solution of this system of equations is $y_1 = -4$, $y_2 = 7/2$.

Example 3.26 For the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = (1, 3)^T$, $\mathbf{x}_2 = (4, 6)^T$, are in $\mathbb{R}^2$, find the matrix of linear transformation T: $\mathbb{R}^2 \to \mathbb{R}^3$, such that

$$T\mathbf{x}_1 = (-2 \quad 2 \quad -7)^T \quad \text{and} \quad T\mathbf{x}_2 = (-2 \quad -4 \quad -10)^T.$$

Solution The transformation T maps column vectors in $\mathbb{R}^2$ into column vectors in $\mathbb{R}^3$. Therefore, T must be a matrix $\mathbf{A}$ of order 3×2. Let

$$\mathbf{A} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}.$$

Therefore, we have

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ -7 \end{bmatrix} \text{ and } \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ -10 \end{bmatrix}.$$

Multiplying and comparing the corresponding elements, we get

$$\begin{aligned} a_1 + 3b_1 &= -2, & 4a_1 + 6b_1 &= -2, \\ a_2 + 3b_2 &= 2, & 4a_2 + 6b_2 &= -4, \\ a_3 + 3b_3 &= -7, & 4a_3 + 6b_3 &= -10. \end{aligned}$$

Solving these equations, we obtain

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -4 & 2 \\ 2 & -3 \end{bmatrix}.$$

Example 3.27 Let T be a linear transformation from $\mathbb{R}^3$ into $\mathbb{R}^2$, where $T\mathbf{x} = \mathbf{A}\mathbf{x}$, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$, $\mathbf{x} = (x \quad y \quad z)^T$. Find $ker(T)$, $ran(T)$ and their dimensions.

Solution To find $ker(T)$, we need to determine all $\mathbf{v} = (v_1 \ v_2 \ v_3)^T$ such that $T\mathbf{v} = \mathbf{0}$. Now, $T\mathbf{v} = \mathbf{A}\mathbf{v} = \mathbf{0}$ gives the equations

$$v_1 + v_2 = 0, \ -v_1 + v_3 = 0$$

whose solution is $v_1 = -v_2 = v_3$. Therefore $\mathbf{v} = v_1[1 \ -1 \ 1]^T$.

Therefore, dimension of $ker(T)$ is 1.

Now, $ran(T)$ is defined as $\{T(\mathbf{v}) \mid \mathbf{v} \in V\}$. We have

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 \\ -v_1 + v_3 \end{bmatrix}$$

$$= v_1\begin{pmatrix}1\\-1\end{pmatrix} + v_2\begin{pmatrix}1\\0\end{pmatrix} + v_3\begin{pmatrix}0\\1\end{pmatrix}.$$

Since $\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix}$, the dimension of *ran* (T) is 2.

Example 3.28 Let T be a linear transformation $T\mathbf{x} = \mathbf{A}\mathbf{x}$ from $\mathbb{R}^2$ into $\mathbb{R}^3$, where

$$\mathbf{A} = \begin{bmatrix}2 & 1\\1 & -1\\3 & 2\end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix}x\\y\end{pmatrix}$$

Find $ker\,(T)$, $ran\,(T)$ and their dimensions.

Solution To find $ker\,(T)$, we need to determine all $\mathbf{v} = (v_1 \;\; v_2)^T$ such that $T\mathbf{v} = \mathbf{0}$. Now, $T\mathbf{v} = A\mathbf{v} = \mathbf{0}$ gives the equations

$$2v_1 + v_2 = 0, \;\; v_1 - v_2 = 0 \quad \text{and} \quad 3v_1 + 2v_2 = 0$$

whose solution is $v_1 = v_2 = 0$. Therefore $\mathbf{v} = (0 \;\; 0)^T$ and the dimension of $ker\,(T)$ is zero.

Now,
$$ran(T) = T(\mathbf{v}) = \begin{bmatrix}2 & 1\\1 & -1\\3 & 2\end{bmatrix}\begin{bmatrix}v_1\\v_2\end{bmatrix} = v_1\begin{bmatrix}2\\1\\3\end{bmatrix} + v_2\begin{bmatrix}1\\-1\\2\end{bmatrix}.$$

Since $(2 \;\; 1 \;\; 3)^T$, $(1 \;\; -1 \;\; 2)^T$ are linearly independent, the dimension of $ran\,(T)$ is 2.

Example 3.29 Find the matrix of a linear transformation T from $\mathbb{R}^3$ into $\mathbb{R}^3$ such that

$$T\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}6\\2\\4\end{pmatrix}, \;\; T\begin{pmatrix}1\\-1\\1\end{pmatrix} = \begin{pmatrix}2\\-4\\2\end{pmatrix}, \;\; T\begin{pmatrix}1\\-2\\3\end{pmatrix} = \begin{pmatrix}6\\6\\5\end{pmatrix}.$$

Solution The transformation T maps elements in $\mathbb{R}^3$ into $\mathbb{R}^3$. Therefore, the transformation is a matrix of order 3×3. Let this matrix be written as

$$T = \mathbf{A} = \begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\\ \gamma_1 & \gamma_2 & \gamma_3\end{bmatrix}.$$

We determine the elements of the matrix $\mathbf{A}$ such that

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\\ \gamma_1 & \gamma_2 & \gamma_3\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}6\\2\\4\end{bmatrix}, \begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\\ \gamma_1 & \gamma_2 & \gamma_3\end{bmatrix}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}2\\-4\\2\end{bmatrix}, \begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3\\ \beta_1 & \beta_2 & \beta_3\\ \gamma_1 & \gamma_2 & \gamma_3\end{bmatrix}\begin{bmatrix}1\\-2\\3\end{bmatrix} = \begin{bmatrix}6\\6\\5\end{bmatrix}.$$

Equating the elements and solving the resulting equations, we obtain

$$\mathbf{A} = \begin{bmatrix}1 & 2 & 3\\ -15/2 & 3 & 13/2\\ 1 & 1 & 2\end{bmatrix}.$$

Example 3.30 Let T be a transformation from $\mathbb{R}^3$ into $\mathbb{R}^1$ defined by

$$T(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2.$$

Show that T is not a linear transformation.

Solution Let $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ be any two elements in $\mathbb{R}^3$. Then

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3).$$

We have

$$T(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2,\ T(\mathbf{y}) = y_1^2 + y_2^2 + y_3^2$$

$$T(\mathbf{x} + \mathbf{y}) = (x_1 + y_1)^2 + (x_2 + y_2)^2 + (x_3 + y_3)^2 \neq T(\mathbf{x}) + T(\mathbf{y}).$$

Therefore, T is not a linear transformation.

Matrix representation of a linear transformation

We observe from the earlier discussion that a matrix $\mathbf{A}$ of order $m \times n$ is a linear transformation whic maps the elements in $\mathbb{R}^n$ into the elements in $\mathbb{R}^m$. Now, let T be a linear transformation from finite dimensional vector space into another finite dimensional vector space over the same field F. W shall now show that with this linear transformation, we may associate a matrix $\mathbf{A}$.

Let V and W be respectively, n-dimensional and m-dimensional vector spaces over the same fiel F. Let T be a linear transformation such that T: $V \to W$. Let

$$X = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\},\ Y = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$$

be the ordered basis of V and W respectively. Let $\mathbf{v}$ be an arbitrary element in V and $\mathbf{w}$ be an arbitrar element in W. Then, there exist scalars, $\alpha_1, \alpha_2, \ldots \alpha_n$ and $\beta_1, \beta_2, \ldots, \beta_m$, not all zero, such that

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n \quad (3.29\ \text{i}$$

$$\mathbf{w} = \beta_1 \mathbf{w}_1 + \beta_2 \mathbf{w}_2 + \ldots + \beta_m \mathbf{w}_m \quad (3.29\ \text{ii}$$

and

$$\mathbf{w} = T\mathbf{v} = T(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n)$$

$$= \alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \ldots + \alpha_n T\mathbf{v}_n \quad (3.29\ \text{iii)}$$

Since every element $T\mathbf{v}_i$, $i = 1, 2, \ldots, n$ is in W, it can be written as a linear combination of the basis vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ in W. That is, there exist scalars a_{ij}, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$ not all zero, such that

$$T\mathbf{v}_i = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \ldots + a_{mi}\mathbf{w}_m$$

$$= [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]\,[a_{1i}, a_{2i}, \ldots, a_{mi}],\ i = 1, 2, \ldots, n \quad (3.29\ \text{iv)}$$

Hence, we can write

$$T[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n] = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (3.29\ \text{v)}$$

or $T\mathbf{X} = \mathbf{YA}$

where $\mathbf{A}$ is the $m \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{3.29 vi}$$

The $m \times n$ matrix $\mathbf{A}$ is called the matrix representation of T or the matrix of T with respect to the ordered basis $\mathbf{X}$ and $\mathbf{Y}$. It may be observed that $\mathbf{X}$ is a basis of the vector space V, on which T acts and $\mathbf{Y}$ is the basis of the vector space W that contains the range of T. Therefore, the matrix representation of T depends not only on T but also on the basis $\mathbf{X}$ and $\mathbf{Y}$. For a given linear transformation T, the elements a_{ij} of the matrix $\mathbf{A} = (a_{ij})$ are determined from (3.29 v), using the given basis vectors in $\mathbf{X}$ and $\mathbf{Y}$. From (3.29 iii), we have (using (3.29 iv))

$$\begin{aligned} \mathbf{w} &= \alpha_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \dots + a_{m1}\mathbf{w}_m) + \alpha_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \dots + a_{m2}\mathbf{w}_m) \\ &\quad + \dots + \alpha_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \dots + a_{mn}\mathbf{w}_m) \\ &= (\alpha_1 a_{11} + \alpha_2 a_{12} + \dots + \alpha_n a_{1n})\,\mathbf{w}_1 + (\alpha_1 a_{21} + \alpha_2 a_{22} + \dots + \alpha_n a_{2n})\mathbf{w}_2 \\ &\quad + \dots + (\alpha_1 a_{m1} + \alpha_2 a_{m2} + \dots + \alpha_n a_{mn})\,\mathbf{w}_m \\ &= \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \dots + \beta_m\mathbf{w}_m \end{aligned}$$

where $\beta_i = \alpha_1 a_{i1} + \alpha_2 a_{i2} + \dots + \alpha_n a_{in}$, $i = 1, 2, \dots, m$.

Hence,
$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

or
$$\boldsymbol{\beta} = \mathbf{A}\,\boldsymbol{\alpha} \tag{3.29 vii}$$

where the matrix $\mathbf{A}$ is as defined in (3.29 vi) and

$$\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_m]^{\mathrm{T}},\ \boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_n]^{\mathrm{T}}.$$

For a given ordered basis vectors $\mathbf{X}$ and $\mathbf{Y}$ of vector spaces V and W respectively, and a linear transformation $T: V \to W$, the matrix $\mathbf{A}$ obtained from (3.29 v) is unique. We prove this result as follows:

Let $\mathbf{A} = (a_{ij})$ and $B = (b_{ij})$ be two matries each of order $m \times n$ such that

$$TX = Y\mathbf{A} \quad \text{and} \quad TX = YB.$$

Therefore, we have

$$Y\mathbf{A} = YB$$

or
$$\sum_{i=1}^{m} \mathbf{w}_i a_{ij} = \sum_{i=1}^{m} \mathbf{w}_i b_{ij}, j = 1, 2, \dots, n.$$

Since $Y = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is a given basis, we obtain $a_{ij} = b_{ij}$ for all i and j and hence $\mathbf{A} \equiv B$.

Example 3.31 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation defined by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y + z \\ y - z \end{pmatrix}$$

Determine the matrix of the linear transformation T, with respect to the ordered basis

(i) $X = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ in $\mathbb{R}^3$ and $Y = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ in $\mathbb{R}^2$

(standard basis e_1, e_2, e_3 in $\mathbb{R}^3$ and e_1, e_2, in $\mathbb{R}^2$).

(ii) $X = \left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\}$ in $\mathbb{R}^3$ and $Y = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ in $\mathbb{R}^2$.

(iii) $X = \left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\}$ in $\mathbb{R}^3$ and $Y = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$ in $\mathbb{R}^2$.

Solution Let $V = \mathbb{R}^3$, $W = \mathbb{R}^2$. Let $X = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, $Y = \{\mathbf{w}_1, \mathbf{w}_2\}$.

(i) We have $\quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

We obtain

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0),$$

$$T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1),$$

$$T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(-1).$$

Using the notation given in (3.29 v), that is $TX = YA$, we write

$$T[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = [\mathbf{w}_1, \mathbf{w}_2]\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

or
$$T\left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Therefore, the matrix of the linear transformation T with respect to the given basis vectors is given by

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

(ii) We have
$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We obtain

$$T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(2) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0),$$

$$T\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(-1),$$

$$T\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1).$$

Using (3.29 v), that is $TX = Y\mathbf{A}$, we write

$$T\left[\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]\begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Therefore, the matrix of the linear transformation T with respect to the given basis vectors is given by

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}.$$

(iii) We have $\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

We obtain

$$T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}(1) + \begin{pmatrix} 1 \\ -1 \end{pmatrix}(1)$$

$$T\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}(0) + \begin{pmatrix} 1 \\ -1 \end{pmatrix}(1)$$

$$T\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}(1) + \begin{pmatrix} 1 \\ -1 \end{pmatrix}(0)$$

Using (3.29 v), that is $TX = Y\mathbf{A}$, we write

$$T\left[\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right] = \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right]\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Therefore, the matrix of the linear transformation T with respect to the given basis vectors is given by

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

Exercise 3.2

Discuss whether V defined in Problems **1** to **10** is a vector space or not. If V is not a vector space, state which of the properties are not satisfied.

1. Let V be the set of the real polynomials of degree $\leq m$ and having 2 as a root with the usual addition and scalar multiplication.
2. Let V be the set of all real polynomials of degree 4 or 6 with the usual addition and scalar multiplication.
3. Let V be the set of all real polynomials of degree ≥ 4 with the usual addition and scalar multiplication.
4. Let V be the set of all rational numbers with the usual addition and scalar multiplication.
5. Let V be the set of all positive real numbers with addition defined as $x + y = xy$ and usual scalar multiplication.
6. Let V be the set of all ordered pairs (x, y) in $\mathbb{R}^2$ with vector addition defined as $(x, y) + (u, v) = (x + u, y + v)$ and scalar multiplication defined as $\alpha(x, y) = (3\alpha x, y)$.

7. Let V be the set of all ordered triplets (x, y, z), $x, y, z \in \mathbb{R}$, with vector addition defined as

$$(x, y, z) + (u, v, w) = (3x + 4u, y - 2v, z + w)$$

and scalar multiplication defined as

$$\alpha(x, y, z) = (\alpha x, \alpha y, \alpha z/3).$$

8. Let V be the set of all positive real numbers with addition defined as $x + y = xy$ and scalar multiplication defined as $\alpha x = x^{\alpha}$.

9. Let V be the set of all real valued continuous functions f on $[a, b]$ such that (i) $\int_a^b f(x)\,dx = 0$ and (ii) $\int_a^b f(x)\,dx = 2$ with usual addition and scalar multiplication.

10. Let V be the set of all solutions of the
   (i) homogeneous linear differential equation $y'' - 3y' + 2y = 0$.
   (ii) non-homogeneous linear differential equation $y'' - 3y' + 2y = x$.
   under the usual addition and scalar multiplication.

Is W a subspace of V in Problems **11** to **15**? If not, state why?

11. Let V be the set of all 3×1 real matrices with usual matrix addition and scalar multiplication and W consisting of all 3×1 real matrices of the form

(i) $\begin{bmatrix} a \\ b \\ a+b \end{bmatrix}$, (ii) $\begin{bmatrix} a \\ a \\ a^2 \end{bmatrix}$, (iii) $\begin{bmatrix} a \\ b \\ 2 \end{bmatrix}$, (iv) $\begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$.

12. Let V be the set of all 3×3 real matrices with the usual matrix addition and scalar multiplication and W consisting of all 3×3 matrices $\mathbf{A}$ which
   (i) have positive elements, (ii) are non-singular,
   (iii) are symmetric, (iv) $\mathbf{A}^2 = \mathbf{A}$.

13. Let V be the set of all 2×2 complex matrices with the usual matrix addition and scalar multiplication and W consisting of all matrices of the form $\begin{bmatrix} z & x+iy \\ x-iy & u \end{bmatrix}$, where x, y, z, u are real numbers and (i) scalars are real numbers, (ii) scalars are complex numbers.

14. Let V consist of all real polynomials of degree ≤ 4 with the usual polynomial addition and scalar multiplication and W consisting of polynomials of degree ≤ 4 having
   (i) constant term 1, (ii) coefficient of t^2 as 0,
   (iii) coefficient of t^3 as 1, (iv) only real roots.

15. Let V be the vector space of all triplets of the form (x_1, x_2, x_3) in $\mathbb{R}^3$ with the usual addition and scalar multiplication and W be the set of triplets of the form (x_1, x_2, x_3) such that
   (i) $x_1 = 2x_2 = 3x_3$, (ii) $x_1 = x_2 = x_3 + 1$,
   (iii) $x_1 \ge 0$, x_2, x_3 arbitrary, (iv) $x_1^2 + x_2^2 + x_3^2 \le 4$.
   (v) x_3 is an integer.

16. Let $\mathbf{u} = (1, 2, -1)$, $\mathbf{v} = (2, 3, 4)$ and $\mathbf{w} = (1, 5, -3)$. Determine whether or not $\mathbf{x}$ is a linear combination of $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, where $\mathbf{x}$ is given by

(i) (4, 3, 10), (ii) (3, 2, 5), (iii) (– 2, 1, – 5).

17. Let $\mathbf{u} = (1, -2, 1, 3)$, $\mathbf{v} = (1, 2, -1, 1)$ and $\mathbf{w} = (2, 3, 1, -1)$. Determine whether or not $\mathbf{x}$ is a linear combination of $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, where $\mathbf{x}$ is given by

(i) (3, 0, 5, – 1), (ii) (2, – 7, 1, 11), (iii) (4, 3, 0, 3).

18. Let $P_1(t) = t^2 - 4t - 6$, $P_2(t) = 2t^2 - 7t - 8$, $P_3(t) = 2t - 3$. Write $P(t)$ as a linear combination of $P_1(t)$, $P_2(t)$, $P_3(t)$, when

(i) $P(t) = -t^2 + 1$, (ii) $P(t) = 2t^2 - 3t - 25$.

19. Let V be the set of all 3×1 real matrices. Show that the set

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \text{ spans } V.$$

20. Let V be the set of all 2×2 real matrices. Show that the set

$$S = \left\{ \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} \right\} \text{ spans } V.$$

21. Examine whether the following vectors in $\mathbb{R}^3/\mathbb{C}^3$ are linearly independent.

(i) (2, 2, 1), (1, – 1, 1), (1, 0, 1), (ii) (1, 2, 3), (3, 4, 5), (6, 7, 8),

(iii) (0, 0, 0), (1, 2, 3), (3, 4, 5), (iv) $(2, i, -1), (1, -3, i), (2i, -1, 5)$,

(v) (1, 3, 4), (1, 1, 0), (1, 4, 2), (1, – 2, 1).

22. Examine whether the following vectors in $\mathbb{R}^4$ are linearly independent.

(i) (4, 1, 2, – 6), (1, 1, 0, 3), (1, – 1, 0, 2), (– 2, 1, 0, 3),

(ii) (1, 2, 3, 1), (2, 1, – 1, 1), (4, 5, 5, 3), (5, 4, 1, 3),

(iii) (1, 2, 3, 4), (2, 0, 1, – 2), (3, 2, 4, 2),

(iv) (1, 1, 0, 1), (1, 1, 1, 1), (– 1, – 1, 1, 1), (1, 0, 0, 1),

(v) (1, 2, 3, – 1), (0, 1, – 1, 2), (1, 5, 1, 8), (– 1, 7, 8, 3).

23. If $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are linearly independent vectors in $\mathbb{R}^3$, then show that

(i) $\mathbf{x} + \mathbf{y}, \mathbf{y} + \mathbf{z}, \mathbf{z} + \mathbf{x}$; (ii) $\mathbf{x}, \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} + \mathbf{z}$

are also linearly independent in $\mathbb{R}^3$.

24. Write (– 4, 7, 9) as a linear combination of the elements of the set S : {(1, 2, 3), (– 1, 3, 4), (3, 1, 2)}. Show that S is not a spanning set in $\mathbb{R}^3$.

25. Write $t^2 + t + 1$ as a linear combination of the elements of the set $S : \{3t, t^2 - 1, t^2 + 2t + 2\}$. Show that S is the spanning set for all polynomials of degree 2 and can be taken as its basis.

26. Let V be the set of all vectors in $\mathbb{R}^4$ and S be a subset of V consisting of all vectors of the form

(i) $(x, y, -y, -x)$, (ii) (x, y, z, w) such that $x + y + z - w = 0$,

(iii) $(x, 0, z, w)$, (iv) (x, x, x, x).

Find the dimension and the basis of S.

27. For what values of k do the following set of vectors form a basis in $\mathbb{R}^3$?

(i) $\{(k, 1-k, k), (0, 3k-1, 2), (-k, 1, 0)\}$,

(ii) $\{(k, 1, 1), (0, 1, 1), (k, 0, k)\}$,

(iii) $\{(k, k, k), (0, k, k), (k, 0, k)\}$,

(iv) $\{(1, k, 5), (1, -3, 2), (2, -1, 1)\}$.

28. Find the dimension and the basis for the vector space V, when V is the set of all 2×2 (i) real matrices, (ii) symmetric matrices, (iii) skew-symmetric matrices, (iv) skew-Hermitian matrices, (v) real matrices $\mathbf{A} = (a_{ij})$ with $a_{11} + a_{22} = 0$, (vi) real matrices $\mathbf{A} = (a_{ij})$ with $a_{11} + a_{12} = 0$.

29. Find the dimension and the basis for the vector space V, when V is the set of all 3×3 (i) diagonal matrices, (ii) upper triangular matrices, (iii) lower triangular matrices.

30. Find the dimension of the vector space V, when V is the set of all $n \times n$ (i) real matrices, (ii) diagonal matrices, (iii) symmetric matrices, (iv) skew-symmetric matrices.

Examine whether the transformation T given in problems **31** to **35** is linear or not. If not linear, state why?

31. $T : \mathbb{R}^2 \to \mathbb{R}^1 ; T\begin{pmatrix} x \\ y \end{pmatrix} = x + y + a,\ a \neq 0$, a real constant.

32. $T : \mathbb{R}^3 \to \mathbb{R}^2 ; T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ x + z \end{pmatrix}$.

33. $T : \mathbb{R}^1 \to \mathbb{R}^2 ;\ T(x) = \begin{pmatrix} x^2 \\ 3x \end{pmatrix}$.

34. $T : \mathbb{R}^2 \to \mathbb{R}^1 ; T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} 0 & x \neq 0, y \neq 0 \\ 2y, & x = 0 \\ 3x, & y = 0. \end{cases}$

35. $T : \mathbb{R}^3 \to \mathbb{R}^1 ; T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = xy + x + z$.

Find $ker(T)$ and $ran(T)$ and their dimensions in problems **36** to **42**.

36. $T : \mathbb{R}^3 \to \mathbb{R}^3 ; T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y \\ z \\ x - y \end{pmatrix}$.

37. $T : \mathbb{R}^2 \to \mathbb{R}^3 ; T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ y - x \\ 3x + 4y \end{pmatrix}$.

38. $T : \mathbb{R}^4 \to \mathbb{R}^3 ; T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x + y + w \\ z \\ y + 2w \end{pmatrix}$.

39. $T : \mathbb{R}^2 \to \mathbb{R}^1 ; T\begin{pmatrix} x \\ y \end{pmatrix} = x + 3y$.

40. $T : \mathbb{R}^3 \to \mathbb{R}^1 ; T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + 3y$.

41. $T : \mathbb{R}^2 \to \mathbb{R}^2 ; T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ x - y \end{pmatrix}$.

42. $T : \mathbb{R}^3 \to \mathbb{R}^2 ; T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y \\ 3x + z \end{pmatrix}$.

43. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation difined by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ x-z \end{pmatrix}.$$

Find the matrix representation of T with respect to the ordered basis

(i) $$X = \left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} \text{ in } \mathbb{R}^3 \text{ and } Y = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\} \text{ in } \mathbb{R}^2.$$

(ii) $$X = \left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} \text{ in } \mathbb{R}^3 \text{ and } Y = \left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}\right\} \text{ in } \mathbb{R}^2.$$

44. Let V and W be two vector spaces in $\mathbb{R}^3$. Let $T: V \to W$ be a linear transformation difined by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ x+y \\ x+y+z \end{pmatrix}.$$

Find the matrix representation of T with respect to the ordered basis

(i) $$X = \left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} \text{ in } V \text{ and } Y = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\} \text{ in } W$$

(ii) $$X = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\} \text{ in } V \text{ and } Y = \left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} \text{ in } W.$$

45. Let V and W be two vector spaces in $\mathbb{R}^3$. Let $T : V \to W$ be a linear transformation defined by

$$T\begin{pmatrix} x \\ y \\ x \end{pmatrix} = \begin{pmatrix} x+z \\ x+y \\ x+y+z \end{pmatrix}.$$

Find the matrix representation of T with respect to the ordered basis

$$X = \left\{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\right\} \text{ in } V \text{ and } Y = \left\{\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right\} \text{ in } W$$

46. Let $T: \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation defined by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y+z \\ x+z \\ x+y+z \end{pmatrix}.$$

Find the matrix representation of T with respect to the ordered basis

$$X = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \text{ in } \mathbb{R}^3 \quad \text{and} \quad Y = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\} \text{ in } \mathbb{R}^4$$

47. Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation.

Let $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$ be the matrix representation of the linear transformation T with respect to the ordered basis vectors $\mathbf{v}_1 = [1, 2]^T$, $\mathbf{v}_2 = [3, 4]^T$ in $\mathbb{R}^2$ and $\mathbf{w}_1 = [-1, 1, 1]^T$, $\mathbf{w}_2 = [1, -1, 1]^T$, $\mathbf{w}_3 = [1, 1, -1]^T$ in $\mathbb{R}^3$. Then, determine the linear transformation T.

48. Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation. Let $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -3 & -4 \end{bmatrix}$ be the matrix representation of the linear transformation with respect to the ordered basis vectors $\mathbf{v}_1 = [1, -1, 1]^T$, $\mathbf{v}_2 = [2, 3, -1]^T$, $\mathbf{v}_3 = [1, 1, -1]^T$ in $\mathbb{R}^3$ and $\mathbf{w}_1 = [1, 1]^T$, $\mathbf{w}_2 = [2, 3]^T$ in $\mathbb{R}^2$. Then, determine the linear transformation T.

49. Let $T: P_1(t) \to P_2(t)$ be a linear tansformation. Let $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ -1 & 1 \end{bmatrix}$ be the matrix representation of the linear transformation with respect to the ordered basis $[1 + t, t]$ in $P_1(t)$ and $[1 - t, 2t, 2 + 3t - t^2]$ in $P_2(t)$. Then, determine the linear transformation T.

50. Let V be the set of all vectors of the form (x_1, x_2, x_3) in $\mathbb{R}^3$ satisfying (i) $x_1 - 3x_2 + 2x_3 = 0$; (ii) $3x_1 - 2x_2 + x_3 = 0$ and $4x_1 + 5x_2 = 0$. Find the dimension and basis for V.

3.4 Solution of General Linear System of Equations

In section 3.2.5, we have discussed the matrix method and the Cramer's rule for solving a system of n equations in n unknowns, $\mathbf{Ax} = \mathbf{b}$. We assumed that the coefficient matrix $\mathbf{A}$ is non-singular, that is $|\mathbf{A}| \neq 0$, or the rank of the matrix $\mathbf{A}$ is n. The matrix method requires evaluation of n^2 determinants each of order $(n - 1)$, to generate the cofactor matrix, and one determinant of order n, whereas the Cramer's rule requires evaluation of $(n + 1)$ determinants each of order n. Since the evaluation of high order determinants is very time consuming, these methods are not used for large values of n, say

$n > 4$. In this section, we discuss a method for solving a general system of m equations in n unknow given by

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.3}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

are respectively called the *coefficient matrix, right hand side column vector* and the *solution vect* The orders of the matrices **A**, **b**, **x** are respectively $m \times n$, $m \times 1$ and $n \times 1$. The matrix

$$(\mathbf{A} \mid \mathbf{b}) = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right] \tag{3.3}$$

is called the *augmented matrix* and has m rows and $(n + 1)$ columns. The augmented matrix describ completely the system of equations. The solution vector of the system of equations (3.30) is n-tuple $(x_1, x_2, \dots, x_n)$ that satisfies all the equations. There are three possibilities:

(i) the system has a unique solution,

(ii) the system has no solution,

(iii) the system has infinite number of solutions.

The system of equations is said to be *consistent*, if it has atleast one solution and *inconsistent*, if it h no solution. Using the concepts of ranks and vector spaces, we now obtain the necessary and sufficie conditions for the existence and uniqueness of the solution of the linear system of equations.

3.4.1 Existence and Uniqueness of the Solution

Let V_n be a vector space consisting of n-tuples in $\mathbb{R}^n$ (or $\mathbb{C}^n$). The row vectors $R_1, R_2, \dots, R_m$ of t $m \times n$ matrix **A** are n-tuples which belong to V_n. Let S be the subspace of V_n generated by the ro of **A**. Then, S is called the *row-space* of the matrix **A** and its dimension is the *row-rank* of **A** and denoted by $rr(\mathbf{A})$. Therefore,

$$\textit{row-rank} \text{ of } \mathbf{A} = rr(\mathbf{A}) = dim(S). \tag{3.3}$$

Similarly, we define the column-space of **A** and the column-rank of **A** denoted by $cr(\mathbf{A})$.

Since the row-space of $m \times n$ matrix **A** is generated by m row vectors of **A**, we have $\dim(S) \leq$ and since S is a subspace of V_n, we have $dim(S) \leq n$. Therefore, we have

$$rr(\mathbf{A}) \leq min(m, n) \text{ and similarly } cr(\mathbf{A}) \leq min(m, n). \tag{3.3}$$

Theorem 3.6 Let $\mathbf{A} = (a_{ij})$ be an $m \times n$ matrix. Then the row-rank and the column-rank of **A** are sam

Proof Let S be the row-space of **A**. The dimension of S is the number of linearly independent row of **A**. Let the dimension of S be r. Therefore, r rows of the matrix A are linearly independent and th remaining $m - r$ rows can be written as a linear combination of these r rows. Let $R_1, R_2, \ldots, R_r$ be the linearly independent rows of **A**. Then, we can write

$$\begin{aligned} R_{r+1} &= \alpha_{r+1,1} R_1 + \alpha_{r+1,2} R_2 + \ldots + \alpha_{r+1,r} R_r \\ R_{r+2} &= \alpha_{r+2,1} R_1 + \alpha_{r+2,2} R_2 + \ldots + \alpha_{r+2,r} R_r \\ &\ldots \\ R_m &= \alpha_{m,1} R_1 + \alpha_{m,2} R_2 + \ldots + \alpha_{m,r} R_r \end{aligned}$$

where $\alpha_{i,j}$ are scalars.

Therefore, the jth element of the row R_{r+1} is given by

$$a_{r+1,j} = \alpha_{r+1,1}\, a_{1j} + \alpha_{r+1,2}\, a_{2j} + \ldots + \alpha_{r+1,r}\, a_{rj}.$$

Similarly, the jth elements of the rows $R_{r+2}, \ldots, R_m$ can be written.

Hence, the jth column of the matrix **A** can be written as

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \ldots \\ a_{rj} \\ a_{r+1,j} \\ \ldots \\ a_{m,j} \end{bmatrix} = a_{1j} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ 0 \\ \alpha_{r+1,1} \\ \ldots \\ \alpha_{m,1} \end{bmatrix} + a_{2j} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \ldots \\ 0 \\ \alpha_{r+1,2} \\ \ldots \\ \alpha_{m,2} \end{bmatrix} + \ldots + a_{rj} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ 1 \\ \alpha_{r+1,r} \\ \ldots \\ \alpha_{m,r} \end{bmatrix}.$$

Thus, every column of the matrix **A** can be written as a linear combination of r linearly independent rows of **A**. Therefore, the dimension of the column-space cannot exceed r, which is the maximum number of linearly independent rows of **A**, that is

$$cr\,(\mathbf{A}) \le r = \text{row-rank of } \mathbf{A}.$$

Similarly, by reversing the roles of rows and columns in the above discussion, we obtain

$$rr\,(\mathbf{A}) \le r = \text{column-rank of } \mathbf{A}.$$

Combining the above results, we have

$$\text{rank}\,(\mathbf{A}) = rr\,(\mathbf{A}) = cr\,(\mathbf{A}) = r.$$

Now, we prove the important result which is known as the *fundamental theorem of linear algebra*.

Theorem 3.7 The non-homogeneous system of equations $\mathbf{Ax} = \mathbf{b}$, where **A** is an $m \times n$ matrix, has a solution if and only if the matrix **A** and the augmented matrix $(\mathbf{A} \mid \mathbf{b})$ have the same rank.

Proof We can write the given system of equations $\mathbf{Ax} = \mathbf{b}$ as

$$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n = \mathbf{b} \tag{3.34}$$

where C_i is the ith column of **A**. Thus, finding solution of the system $\mathbf{Ax} = \mathbf{b}$ is equivalent to finding scalars $x_1, x_2, \ldots, x_n$ which satisfy the equation (3.34).

Let rank (**A**) = rank (**A** | **b**) = r. Then, the column-rank of the matrix (**A** | **b**) is r and $r \le n$. Therefore, there are r linearly independent column vectors. Suppose these are the first r columns,. Then, the remaining columns $C_{r+1}, C_{r+2}, \ldots, C_{n+1}$ can be written as a linear combination of these r linearly independent column vectors. Thus, the $(n + 1)$th column of (**A** | **b**) is a linear combination of its first n columns, that is

$$\mathbf{b} = \alpha_1 C_1 + \alpha_2 C_2 + \ldots + \alpha_n C_n$$

which means that $\mathbf{A\alpha} = \mathbf{b}$, or $\mathbf{Ax} = \mathbf{b}$ has a solution.

Conversely, let $\mathbf{Ax} = \mathbf{b}$ have a solution, say $\mathbf{x} = \boldsymbol{\alpha}$. Then, we can write

$$\mathbf{b} = \alpha_1 C_1 + \alpha_2 C_2 + \ldots + \alpha_n C_n.$$

Thus, the column-spaces of **A** and (**A** | **b**) are the same and have the same dimension. Since, these dimensions are rank (**A**) and rank (**A** | **b**) respectively, we obtain rank (**A**) = rank (**A** | **b**).

Remark 11

A system of linear equations $\mathbf{Ax} = \mathbf{b}$ is consistent, if the vector **b** can be written as a linear combination of the columns $C_1, C_2, \ldots, C_n$ of **A**. If **b** is not a linear combination of the columns of **A**, that is, **b** is linearly independent of the columns $C_1, C_2, \ldots, C_n$, then no scalars can be determined that satisfy Eq. (3.34) and the system is inconsistent in this case.

In section 3.2.3, we defined the rank of an $m \times n$ matrix **A** in terms of the determinants of the submatrices of **A**. An $m \times n$ matrix has rank r if it has at least one square submatrix of order r which is non-singular and all square submatrices of order greater than r are singular. This approach is very time consuming when n is large. Now, we discuss an alternative procedure to obtain the rank of a matrix.

3.4.2 Elementary Row and Column Operations

The following three operations on a matrix **A** are called the *elementary row operations*:

(i) Interchange of any two rows (written as $R_i \sim R_j$).

(ii) Multiplication/division of any row by a non-zero scalar (written as αR_i).

(iii) Adding/subtracting a scalar multiple of any row to another row (written as $R_i \leftarrow R_i + \alpha R_j$, that is α multiples of the elements of the jth row are added to the corresponding elements of the ith row. The elements of the jth row remain unchanged, whereas, the elements of the ith row get changed).

These operations change the form of **A** but do not change the row-rank of **A** as they do not change the row-space of **A**. A matrix **B** is said to be *row equivalent* to a matrix **A**, if the matrix **B** can be obtained from the matrix **A** by a finite sequence of elementary row operations. Then, we usually write **B** ≈ **A**. We observe that

(i) every matrix is row equivalent to itself.

(ii) if **A** is row equivalent to **B**, then **B** is row equivalent to **A**.

(iii) if **A** is row equivalent to **B** and **B** is row equivalent to **C**, then **A** is row equivalent to **C**.

The above operations performed on columns (that is column in place of row) are called *elementary column operations*.

3.4.3 Echelon Form of a Matrix

An $m \times n$ matrix is called a *row echelon* matrix or in *row echelon* form if the number of zeros preceeding the first non-zero entry of a row increases row by row until a row having all zero entries (or no other elimination is possible) is obtained. Therefore, a matrix is in row echelon form if the following are satisfied.

(i) If the ith row contains all zeros, it is true for all subsequent rows. not even constant

(ii) If a column contains a non-zero entry of any row, then every subsequent entry in this column is zero, that is, if the ith and $(i + 1)$th rows are both non-zero rows, then the initial non-zero entry of the $(i + 1)$th row appears in a later column than that of the ith row.

(iii) Rows containing all zeros occur only after all non-zero rows.

For example, the following matrices are in row echelon form.

$$\begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 5 & 4 & 1 \\ 0 & 0 & 0 & 9 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let $\mathbf{A} = (a_{ij})$ be a given $m \times n$ matrix. Assume that $a_{11} \neq 0$. If $a_{11} = 0$, we interchange the first row with some other row to make the element in the (1, 1) position as non-zero. Using elementary row operations, we reduce the matrix $\mathbf{A}$ to its row echelon form (elements of first column below a_{11} are made zero, then elements in the second column below a_{22} are made zero and so on).

Similarly, we define the column echelon form of a matrix.

Rank of A The number of non-zero rows in the row echelon form of a matrix $\mathbf{A}$ gives the rank of the matrix $\mathbf{A}$ (that is, the dimension of the row-space of the matrix $\mathbf{A}$) and the set of the non-zero rows in the row echelon form gives the basis of the row-space.

Similar results hold for column echelon matrices.

Remark 12

(i) If $\mathbf{A}$ is a square matrix, then the row-echelon form is an upper triangular matrix and the column echelon form is a lower triangular matrix.

(ii) This approach can also be used to examine whether a given set of vectors are linearly independent or not. We form the matrix with each vector as its row (or column) and reduce it to the row (column) echelon form. The given vectors are linearly independent, if the row echelon form has no row with all its elements as zeros. The number of non-zero rows is the dimension of the given set of vectors and the set of vectors consisting of the non-zero rows is the basis.

Example 3.32 Reduce the following matrices to row echelon form and find their ranks.

(i) $\begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \\ -2 & 8 & 2 \end{bmatrix}$, (ii) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 5 \\ 1 & 5 & 5 & 7 \\ 8 & 1 & 14 & 17 \end{bmatrix}$.

Solution Let the given matrix be denoted by **A**. We have

(i) $$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \\ -2 & 8 & 2 \end{bmatrix} \begin{matrix} R_2 - 2R_1 \\ R_3 + 2R_1 \end{matrix} \approx \begin{bmatrix} 1 & 3 & 5 \\ 0 & -7 & -6 \\ 0 & 14 & 12 \end{bmatrix} R_3 + 2R_2 \approx \begin{bmatrix} 1 & 3 & 5 \\ 0 & -7 & -6 \\ 0 & 0 & 0 \end{bmatrix}.$$

This is the row echelon form of **A**. Since the number of non-zero rows in the row echelon form is 2, we get rank (**A**) = 2.

(ii) $$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 5 \\ 1 & 5 & 5 & 7 \\ 8 & 1 & 14 & 17 \end{bmatrix} \begin{matrix} R_2 - 2R_1 \\ R_3 - R_1 \\ R_4 - 8R_1 \end{matrix} \approx \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -2 & -3 \\ 0 & 3 & 2 & 3 \\ 0 & -15 & -10 & -15 \end{bmatrix} \begin{matrix} R_3 + R_2 \\ R_4 - 5R_2 \end{matrix}$$

$$\approx \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the number of non-zero rows in the row echelon form of **A** is 2, we get rank (**A**) = 2.

Example 3.33 Reduce the following matrices to column echelon form and find their ranks.

(i) $$\begin{bmatrix} 3 & 1 & 7 \\ 1 & 2 & 4 \\ 4 & -1 & 7 \\ 2 & 1 & 5 \end{bmatrix},$$ (ii) $$\begin{bmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & -3 & -3 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 3 & 3 \end{bmatrix}.$$

Solution Let the given matrix be denoted by **A**. We have

(i) $$\mathbf{A} = \begin{bmatrix} 3 & 1 & 7 \\ 1 & 2 & 4 \\ 4 & -1 & 7 \\ 2 & 1 & 5 \end{bmatrix} \begin{matrix} C_2 - C_1/3 \\ C_3 - 7C_1/3 \end{matrix} \approx \begin{bmatrix} 3 & 0 & 0 \\ 1 & 5/3 & 5/3 \\ 4 & -7/3 & -7/3 \\ 2 & 1/3 & 1/3 \end{bmatrix} C_3 - C_2 \approx \begin{bmatrix} 3 & 0 & 0 \\ 1 & 5/3 & 0 \\ 4 & -7/3 & 0 \\ 2 & 1/3 & 0 \end{bmatrix}.$$

Since the column echelon form of **A** has two non-zero columns, rank (**A**) = 2.

(ii) $$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 & 1 \\ -1 & 1 & -3 & -3 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 3 & 3 \end{bmatrix} \begin{matrix} C_2 - C_1 \\ C_3 + C_1 \\ C_4 - C_1 \end{matrix} \approx \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & -4 & -2 \\ 1 & -1 & 2 & 1 \\ 1 & -2 & 4 & 2 \end{bmatrix} \begin{matrix} C_3 + 2C_2 \\ C_4 + C_2 \end{matrix}$$

$$\approx \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}.$$

Since the column echelon form of **A** has 2 non-zero columns, rank $(\mathbf{A}) = 2$.

Example 3.34 Examine whether the following set of vectors is linearly independent. Find the dimension and the basis of the given set of vectors.

(i) $(1, 2, 3, 4), (2, 0, 1, -2), (3, 2, 4, 2)$,

(ii) $(1, 1, 0, 1), (1, 1, 1, 1), (-1, 1, 1, 1), (1, 0, 0, 1)$,

(iii) $(2, 3, 6, -3, 4), (4, 2, 12, -3, 6), (4, 10, 12, -9, 10)$.

Solution Let each given vector represent a row of a matrix **A**. We reduce **A** to row echelon form. If all the rows of the row echelon form have some non-zero elements, then the given set of vectors are linearly independent.

(i)
$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 1 & -2 \\ 3 & 2 & 4 & 2 \end{bmatrix} \begin{matrix} R_2 - 2R_1 \\ R_3 - 3R_1 \end{matrix} \approx \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -5 & -10 \\ 0 & -4 & -5 & -10 \end{bmatrix} R_3 - R_2$$

$$\approx \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -5 & -10 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since all the rows in the row echelon form of **A** are not non-zero, the given set of vectors are linearly dependent. Since the number of non-zero rows is 2, the dimension of the given set of vectors is 2. The basis can be taken as the set of vectors $\{(1\ \ 2\ \ 3\ \ 4), (0, -4, -5, -10)\}$.

(ii)
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} R_2 - R_1 \\ R_3 + R_1 \\ R_4 - R_1 \end{matrix} \approx \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & -1 & 0 & 0 \end{bmatrix} R_2 \sim R_3 \approx \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} R_4 + R_2/2$$

$$\approx \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 1 \end{bmatrix} R_4 - R_3/2 \approx \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Since all the rows in the row echelon form of **A** are non-zero, the given set of vectors are linearly independent and the dimension of the given set of vectors is 4. The set of vectors $\{(1, 1, 0, 1), (0, 2, 1, 2), (0, 0, 1, 0), (0, 0, 0, 1)\}$ or the given set itself forms the basis.

(iii) $\mathbf{A} = \begin{bmatrix} 2 & 3 & 6 & -3 & 4 \\ 4 & 2 & 12 & -3 & 6 \\ 4 & 10 & 12 & -9 & 10 \end{bmatrix} \begin{matrix} R_2 - 2R_1 \\ R_3 - 2R_1 \end{matrix} \approx \begin{bmatrix} 2 & 3 & 6 & -3 & 4 \\ 0 & -4 & 0 & 3 & -2 \\ 0 & 4 & 0 & -3 & 2 \end{bmatrix} R_3 + R_4$

$$\approx \begin{bmatrix} 2 & 3 & 6 & -3 & 4 \\ 0 & -4 & 0 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since all the rows in the row echelon form of **A** are not non-zero, the given set of vectors are linearly dependent. Since the number of non-zero rows is 2, the dimension of the given set of vectors is 2 and its basis can be taken as the set $\{(2, 3, 6, -3, 4), (0, -4, 0, 3, -2)\}$.

3.4.4 Gauss Elimination Method for Non-homogeneous Systems

Consider a non-homogeneous system of m equations in n unknowns

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.35}$$

where
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

We assume that at least one element of **b** is not zero. We write the augmented matrix of order $m \times (n + 1)$ as

$$(\mathbf{A} \mid \mathbf{b}) = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

and reduce it to the row echelon form by using elementary row operations. We need a maximum of $(m - 1)$ stages of eliminations to reduce the given augmented matrix to the equivalent row echelon form. This process may terminate at an earlier stage. We then have an equivalent system of the form

$$(\mathbf{A} \mid \mathbf{b}) = \left[\begin{array}{ccccccc|c} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1n} & & b_1 \\ 0 & \bar{a}_{22} & \cdots & \bar{a}_{2r} & \cdots & \bar{a}_{2n} & & \bar{b}_2 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & \cdots & a^*_{rr} & \cdots & a^*_{rn} & & b^*_r \\ 0 & 0 & \cdots & 0 & \cdots & 0 & & b^*_{r+1} \\ \vdots & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & & b^*_m \end{array}\right]$$

where $r \le m$ and $a_{11} \ne 0, \bar{a}_{22} \ne 0, \ldots, a^*_{rr} \ne 0$ are called *pivots*. We have the following cases:

(a) Let $r < m$ and one or more of the elements $b^*_{r+1}, b^*_{r+2}, \ldots, b^*_m$ are not zero. Then rank $(\mathbf{A}) \neq$ rank $(\mathbf{A} \mid \mathbf{b})$ and the system of equations has no solution.

(b) Let $m \geq n$ and $r = n$ (the number of columns in $\mathbf{A}$) and $b^*_{r+1}, b^*_{r+2}, \ldots, b^*_m$ are all zeros. In this case, rank $(\mathbf{A})$ = rank $(\mathbf{A} \mid \mathbf{b}) = n$ and the system of equations has a unique solution. We solve the nth equations for x_n, the $(n-1)$th equation for x_{n-1} and so on. This procedure is called the *back substitution method*.

For example, if we have 10 equations in 5 variables, then the augmented matrix is of order 10×6. When rank $(\mathbf{A})$ = rank $(\mathbf{A} \mid \mathbf{b}) = 5$, the system has a unique solution.

(c) Let $r < n$ and $b^*_{r+1}, b^*_{r+2}, \ldots, b^*_m$ are all zeros. In this case, r unknowns, $x_1, x_2, \ldots, x_r$ can be determined in terms of the remaining $(n - r)$ unknowns $x_{r+1}, x_{r+2}, \ldots, x_n$ by solving the rth equation for x_r, $(r-1)$th equation for x_{r-1} and so on. In this case, we obtain an $(n - r)$ parameter family of solutions, that is infinitely many solutions.

Remark 13

(a) We do not, normally use column elementary operations in solving the linear system of equations. When we interchange two columns, the order of the unknowns in the given system of equations is also changed. Keeping track of the order of unknowns is quite difficult.

(b) Gauss elimination method may be written as

$$(\mathbf{A} \mid \mathbf{b}) \xrightarrow[\text{row operations}]{\text{Elementary}} (\mathbf{B} \mid \mathbf{c}).$$

The matrix $\mathbf{B}$ is the row echelon form of the matrix $\mathbf{A}$ and $\mathbf{c}$ is the new right hand side column vector. We obtain the solution vector (if it exists) using the back substitution method.

(c) If $\mathbf{A}$ is a square matrix of order n, then $\mathbf{B}$ is an upper triangular matrix of order n.

(d) Gauss elimination method can be used to solve p systems of the form $\mathbf{Ax} = \mathbf{b}_1$, $\mathbf{Ax} = \mathbf{b}_2, \ldots,$ $\mathbf{Ax} = \mathbf{b}_p$ which have the same coefficient matrix but different right hand side column vectors. We form the augmented matrix as $(\mathbf{A} \mid \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p)$, which has m rows and $(n + p)$ columns. Using the elementary row operations, we obtain the row equivalent system $(\mathbf{B} \mid \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p)$, where $\mathbf{B}$ is the row echelon form of A. Now, we solve the systems $\mathbf{Bx} = \mathbf{c}_1$, $\mathbf{Bx} = \mathbf{c}_2, \ldots,$ $\mathbf{Bx} = \mathbf{c}_p$, using the back substitution method.

Example 3.35 Solve the following systems of equations (if possible) using Gauss elimination method.

(i) $\begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 2 \\ -1 & 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix}$, (ii) $\begin{bmatrix} 2 & 0 & 1 \\ 1 & -1 & 1 \\ 4 & -2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$,

(iii) $\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -1 \\ 5 & -2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$.

Solution We write the augmented matrix and reduce it to row echelon form by applying elementary row operations.

(i) $$(\mathbf{A}\,|\,\mathbf{b}) = \left[\begin{array}{ccc|c} 2 & 1 & -1 & 4 \\ 1 & -1 & 2 & -2 \\ -1 & 2 & -1 & 2 \end{array}\right] \begin{array}{l} R_2 - R_1/2 \\ R_3 + R_1/2 \end{array} \approx \left[\begin{array}{ccc|c} 2 & 1 & -1 & 4 \\ 0 & -3/2 & 5/2 & -4 \\ 0 & 5/2 & -3/2 & 4 \end{array}\right] R_3 + 5R_2/3$$

$$\approx \left[\begin{array}{ccc|c} 2 & 1 & -1 & 4 \\ 0 & -3/2 & 5/2 & -4 \\ 0 & 0 & 8/3 & -8/3 \end{array}\right].$$

Using the back substitution method, we obtain the solution as

$$\frac{8}{3}z = -\frac{8}{3}, \quad \text{or} \quad z = -1,$$

$$-\frac{3}{2}y + \frac{5}{2}z = -4, \quad \text{or} \quad y = 1,$$

$$2x + y - z = 4, \quad \text{or} \quad x = 1.$$

Therefore, the system of equations has the unique solution $x = 1$, $y = 1$, $z = -1$.

(ii) $$(\mathbf{A}\,|\,\mathbf{b}) = \left[\begin{array}{ccc|c} 2 & 0 & 1 & 3 \\ 1 & -1 & 1 & 1 \\ 4 & -2 & 3 & 3 \end{array}\right] \begin{array}{l} R_2 - R_1/2 \\ R_3 - 2R_1 \end{array} \approx \left[\begin{array}{ccc|c} 2 & 0 & 1 & 3 \\ 0 & -1 & 1/2 & -1/2 \\ 0 & -2 & 1 & -3 \end{array}\right] R_3 - 2R_2$$

$$\approx \left[\begin{array}{ccc|c} 2 & 0 & 1 & 3 \\ 0 & -1 & 1/2 & -1/2 \\ 0 & 0 & 0 & -2 \end{array}\right].$$

We find that rank $(\mathbf{A}) = 2$ and rank $(\mathbf{A}\,|\,\mathbf{b}) = 3$. Therefore, the system of equations has no solution.

(iii) $$(\mathbf{A}\,|\,\mathbf{b}) = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 2 & 1 & -1 & 2 \\ 5 & -2 & 2 & 5 \end{array}\right] \begin{array}{l} R_2 - 2R_1 \\ R_3 - 5R_1 \end{array} \approx \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -3 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right] R_3 - R_2$$

$$\approx \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The system is consistent and has infinite number of solutions. We find that the last equation is satisfied for all values of x, y, z. From the second equation we get $3y - 3z = 0$, or $y = z$. From the first equation, we get $x - y + z = 1$, or $x = 1$. Therefore, we obtain the solution $x = 1$, $y = z$ and z is arbitrary.

Example 3.36 Solve the following systems of equations using Gauss elimination method.

(i)
$$\begin{aligned} 4x - 3y - 9z + 6w &= 0 \\ 2x + 3y + 3z + 6w &= 6 \\ 4x - 21y - 39z - 6w &= -24, \end{aligned}$$

(ii)
$$\begin{aligned} x + 2y - 2z &= 1 \\ 2x - 3y + z &= 0 \\ 5x + y - 5z &= 1 \\ 3x + 14y - 12z &= 5. \end{aligned}$$

Solution We have

(i) $(\mathbf{A} \mid \mathbf{b})$

$$= \left[\begin{array}{cccc|c} 4 & -3 & -9 & 6 & 0 \\ 2 & 3 & 3 & 6 & 6 \\ 4 & -21 & -39 & -6 & -24 \end{array}\right] \begin{matrix} \\ R_2 - R_1/2 \\ R_3 - R_1 \end{matrix} \approx \left[\begin{array}{cccc|c} 4 & -3 & -9 & 6 & 0 \\ 0 & 9/2 & 15/2 & 3 & 6 \\ 0 & -18 & -30 & -12 & -24 \end{array}\right] R_3 + 4R_2$$

23387135
23387135
23387135

$$\approx \left[\begin{array}{cccc|c} 4 & -3 & -9 & 6 & 0 \\ 0 & 9/2 & 15/2 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The system of equations is consistent and has infinite number of solutions. Choose w as arbitrary. From the second equation, we obtain

$$\frac{9}{2} y + \frac{15}{2} z = 6 - 3w, \quad \text{or} \quad y = \frac{2}{9}\left(6 - 3w - \frac{15}{2} z\right) = \frac{1}{3}(4 - 5z - 2w).$$

From the first equation, we obtain

$$4x = 3y + 9z - 6w = 4 - 5z - 2w + 9z - 6w = 4 + 4z - 8w$$

or
$$x = 1 + z - 2w.$$

Thus, we obtain a two parameter family of solutions

Rank (A/B) Sid
Rank (A/B) Sid
(A/B) (A/B)

$$x = 1 + z - 2w \quad \text{and} \quad y = (4 - 5z - 2w)/3$$

where z and w are arbitrary.

$$\text{(ii)} \quad (\mathbf{A} \mid \mathbf{b}) = \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 2 & -3 & 1 & 0 \\ 5 & 1 & -5 & 1 \\ 3 & 14 & -12 & 5 \end{array}\right] \begin{matrix} R_2 - 2R_1 \\ R_3 - 5R_1 \\ R_4 - 3R_1 \end{matrix} \approx \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & -7 & 5 & -2 \\ 0 & -9 & 5 & -4 \\ 0 & 8 & -6 & 2 \end{array}\right] \begin{matrix} R_3 - 9R_2/7 \\ R_4 + 8R_2/7 \end{matrix}$$

$$\approx \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & -7 & 5 & -2 \\ 0 & 0 & -10/7 & -10/7 \\ 0 & 0 & -2/7 & -2/7 \end{array}\right] R_4 - 5R_3 \approx \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & -7 & 5 & -2 \\ 0 & 0 & -10/7 & -10/7 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The last equation is satisfied for all values of x, y, z. From the third equation, we obtain $z = 1$. Back substitution gives $y = 1$, $x = 1$. Hence, the system of equations has a unique solution $x = 1$, $y = 1$ and $z = 1$. Since $R_4 = (24R_1 - 7R_2 + R_3)/5$, the last equation is redundent.

3.4.5 Homogeneous System of Linear Equations

Consider the homogeneous system of equations

$$\mathbf{Ax} = \mathbf{0} \tag{3.36}$$

where $\mathbf{A}$ is an $m \times n$ matrix. The homogeneous system is always consistent since $\mathbf{x} = \mathbf{0}$ (trivial solution) is always a solution. In this case, rank $(\mathbf{A})$ = rank $(\mathbf{A} \mid \mathbf{0})$. Therefore, for the homogeneous system to have a non-trivial solution, we require that rank $(\mathbf{A}) < n$. If rank $(\mathbf{A}) = r < n$ we obtain an $(n - r)$ parameter family of solutions which form a vector space of dimension $(n - r)$ as $(n - r)$ parameters can be chosen arbitrarily.

The solution space of the homogeneous system is called the *null space* and its dimension is called the *nullity* of $\mathbf{A}$. Therefore, we obtain the result

$$\text{rank } (\mathbf{A}) + \text{nullity } (\mathbf{A}) = n \quad \text{(see Theorem 3.5).}$$

Remark 14

(a) If $\mathbf{x}_1$ and $\mathbf{x}_2$ are two solutions of a linear homogeneous system, then $\alpha \mathbf{x}_1 + \beta \mathbf{x}_2$ is also a solution of the homogeneous system for any scalars α, β. This result does not hold for non-homogeneous systems.

(b) A homogeneous system of m equations in n unknowns and $m < n$, always possesses a non-trivial solution.

Theorem 3.8 If a non-homogeneous system of linear equations $\mathbf{Ax} = \mathbf{b}$ has solutions, then all these solutions are of the form $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$, where $\mathbf{x}_0$ is any fixed solution of $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{x}_h$ is any solution of the corresponding homogeneous system.

Proof Let $\mathbf{x}$ be any solution and $\mathbf{x}_0$ be any fixed solution of $\mathbf{Ax} = \mathbf{b}$. Therefore, we have

$$\mathbf{Ax} = \mathbf{b} \quad \text{and} \quad \mathbf{Ax}_0 = \mathbf{b}.$$

Subtracting, we get

$$\mathbf{Ax} - \mathbf{Ax}_0 = \mathbf{0}, \quad \text{or} \quad \mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}.$$

Thus, the difference $\mathbf{x} - \mathbf{x}_0$ between any solution $\mathbf{x}$ of $\mathbf{Ax} = \mathbf{b}$ and any fixed solution $\mathbf{x}_0$ of $\mathbf{Ax} = \mathbf{b}$ is a solution of the homogeneous system $\mathbf{Ax} = \mathbf{0}$, say $\mathbf{x}_h$. Hence, the result.

Remark 15

If the non-homogeneous system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is an $m \times n$ matrix $(m \geq n)$ has a unique solution, that is, rank $(\mathbf{A}) = n$, then the corresponding homogeneous system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution, that is $\mathbf{x}_h = \mathbf{0}$.

Example 3.37 Solve the following homogeneous system of equations $\mathbf{Ax} = \mathbf{0}$, where $\mathbf{A}$ is given by

(i) $\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 3 & 2 \end{bmatrix}$, (ii) $\begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$, (iii) $\begin{bmatrix} 1 & 1 & -1 & 1 \\ 2 & 3 & 1 & 4 \\ 3 & 2 & -6 & 1 \end{bmatrix}$.

Find the rank $(\mathbf{A})$ and nullity $(\mathbf{A})$.

Solution We write the augmented matrix $(\mathbf{A} \mid \mathbf{0})$ and reduce it to row echelon form.

(i) $(\mathbf{A} \mid \mathbf{0}) = \left[\begin{array}{rr|r} 2 & 1 & 0 \\ 1 & -1 & 0 \\ 3 & 2 & 0 \end{array}\right] \begin{matrix} R_2 - R_1/2 \\ R_3 - 3R_1/2 \end{matrix} \approx \left[\begin{array}{rr|r} 2 & 1 & 0 \\ 0 & -3/2 & 0 \\ 0 & 1/2 & 0 \end{array}\right] R_3 + R_2/3 \approx \left[\begin{array}{rr|r} 2 & 1 & 0 \\ 0 & -3/2 & 0 \\ 0 & 0 & 0 \end{array}\right].$

Since, rank $(\mathbf{A}) = 2$ = number of unknowns, the system has only a trivial solution. Hence, nullity $(\mathbf{A}) = 0$.

(ii) $(\mathbf{A} \mid \mathbf{0}) = \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right] \begin{matrix} R_2 - R_1 \\ R_3 - R_1 \end{matrix} \approx \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & -3 & 4 & 0 \end{array}\right] R_3 - 3R_2 \approx \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -2 & 0 \end{array}\right].$

Since rank $(\mathbf{A}) = 3$ = number of unknowns, the homogeneous system has only a trivial solution. Therefore, nullity $(\mathbf{A}) = 0$.

(iii) $(\mathbf{A} \mid \mathbf{0}) = \left[\begin{array}{rrrr|r} 1 & 1 & -1 & 1 & 0 \\ 2 & 3 & 1 & 4 & 0 \\ 3 & 2 & -6 & 1 & 0 \end{array}\right] \begin{matrix} R_2 - 2R_1 \\ R_3 - 3R_1 \end{matrix} \approx \left[\begin{array}{rrrr|r} 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & -1 & -3 & -2 & 0 \end{array}\right] R_3 + R_2 \approx \left[\begin{array}{rrrr|r} 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$

Therefore, rank $(\mathbf{A}) = 2$ and the number of unknowns is 4. Hence, we obtain a two parameter family of solutions as $x_2 = -3x_3 - 2x_4$, $x_1 = -x_2 + x_3 - x_4 = 4x_3 + x_4$, where x_3 and x_4 are arbitrary. Therefore, nullity $(\mathbf{A}) = 2$.

3.4.6 Gauss-Jordan Method to Find the Inverse of a Matrix

Let $\mathbf{A}$ be a non-singular matrix of order n. Therefore, its inverse $\mathbf{B} = \mathbf{A}^{-1}$ exists and $\mathbf{AB} = \mathbf{I}$. Let the matrix $\mathbf{B}$ be written as $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n]$, where $\mathbf{b}_i$ is the ith column of $\mathbf{B}$.

From $\mathbf{AB} = \mathbf{I}$, we obtain

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n] = \mathbf{I} = [\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n]. \tag{3.37}$$

where $\mathbf{I}_i$ is the column vector with 1 in the ith position and zeros elsewhere. Using Gauss elimination method for solving n systems with the same coefficient matrix (see Remark 13(d)), we form the augmented matrix

$$(\mathbf{A} \mid \mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n) \text{ which is same as } (\mathbf{A} \mid \mathbf{I}),$$

where $\mathbf{I}$ is the identity matrix of order n. Using elementary row operations, we obtain

$$(\mathbf{A} \mid \mathbf{I}) \xrightarrow[\text{row operations}]{\text{Elementary}} (\mathbf{I} \mid \mathbf{B}). \tag{3.38}$$

Hence, $\mathbf{B} = \mathbf{A}^{-1}$. This method is called the *Gauss-Jordan* method. In the first step, all the elements below the pivot a_{11} are made zero. In the second step, all the elements above and below the second pivot $\bar{a}_{22}$ are made zero. At the kth step, all the elements above and below the pivot a_{kk}^{*} are made zero. The pivot in the (i, i) position can be made 1 at every step or when the elimination is completed.

Example 3.38 Using Gauss-Jordan method, find the inverse of the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$$

Solution We have

$$(\mathbf{A} \mid \mathbf{I}) = \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right].$$

The pivot element a_{11} is -1. We make it 1 by multiplying the first row by -1. Therefore,

$$(\mathbf{A} \mid \mathbf{I}) \approx \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \begin{matrix} R_2 - 3R_1 \\ R_3 + R_1 \end{matrix} \approx \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] R_2/2$$

$$\approx \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 7/2 & 3/2 & 1/2 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \begin{matrix} R_1 + R_2 \\ R_3 - 2R_2 \end{matrix} \approx \left[\begin{array}{ccc|ccc} 1 & 0 & 3/2 & 1/2 & 1/2 & 0 \\ 0 & 1 & 7/2 & 3/2 & 1/2 & 0 \\ 0 & 0 & -5 & -4 & -1 & 1 \end{array}\right] (-R_3)/5$$

$$\approx \left[\begin{array}{ccc|ccc} 1 & 0 & 3/2 & 1/2 & 1/2 & 0 \\ 0 & 1 & 7/2 & 3/2 & 1/2 & 0 \\ 0 & 0 & 1 & 4/5 & 1/5 & -1/5 \end{array}\right] \begin{matrix} R_1 - 3R_3/2 \\ R_2 - 7R_3/2 \end{matrix} \approx \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -7/10 & 2/10 & 3/10 \\ 0 & 1 & 0 & -13/10 & -2/10 & 7/10 \\ 0 & 0 & 1 & 4/5 & 1/5 & -1/5 \end{array}\right].$$

Hence,

$$\mathbf{A}^{-1} = \frac{1}{10}\begin{bmatrix} -7 & 2 & 3 \\ -13 & -2 & 7 \\ 8 & 2 & -2 \end{bmatrix}.$$

Exercise 3.3

Using the elementary row operations, determine the ranks of the following matrices.

1. $\begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$.

2. $\begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 2 \\ 5 & -5 & 11 \end{bmatrix}$.

3. $\begin{bmatrix} 2 & 1 & -2 \\ -1 & -1 & 1 \\ 3 & 1 & -2 \end{bmatrix}$.

4. $\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 5 \\ 1 & 4 & -13 & -5 \end{bmatrix}$.

5. $\begin{bmatrix} 1 & -2 & 1 & -1 \\ 1 & 1 & -2 & 3 \\ 4 & 1 & -5 & 8 \end{bmatrix}$.

6. $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 3 & -1 \\ 8 & 13 & 14 \end{bmatrix}$.

7. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 7 & 11 & 15 & 19 \\ 9 & 15 & 21 & 27 \end{bmatrix}$.

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

9. $\begin{bmatrix} 2 & 0 & -1 & 0 \\ 4 & 1 & 0 & 5 \\ 0 & 1 & 3 & 6 \\ 6 & 1 & -2 & 6 \end{bmatrix}$.

10. $\begin{bmatrix} 2 & 3 & 1 & 0 & 4 \\ 3 & 1 & 2 & -1 & 1 \\ 4 & -1 & 3 & -2 & -2 \\ 5 & 4 & 3 & -1 & 5 \end{bmatrix}$.

Using the elementary column operations, determine the ranks of the following matrices.

11. $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 3 \end{bmatrix}$. **12.** $\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$. **13.** $\begin{bmatrix} 1 & 2 & 3 & 2 \\ -1 & 1 & 3 & -5 \\ 2 & 3 & 4 & 5 \end{bmatrix}$.

14. $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \\ 5 & 4 & -5 \end{bmatrix}$. **15.** $\begin{bmatrix} 2 & 3 & 1 & 0 & 4 \\ 3 & 1 & 2 & -1 & 1 \\ 4 & -1 & 3 & -2 & -2 \\ 5 & 4 & 3 & -1 & 5 \end{bmatrix}$.

Determine whether the following set of vectors is linearly independent. Find also its dimension.

16. $\{(3, 2, 4), (1, 0, 2), (1, -1, -1)\}$. **17.** $\{(2, 2, 1), (1, -1, 1), (1, 0, 1)\}$.

18. $\{(2, 1, 0), (1, -1, 1), (4, 1, 2), (2, -3, 3)\}$. **19.** $\{(2, 2, 1), (2, i, -1), (1 + i, -i, 1)\}$.

20. $\{(1, 1, 1), (i, i, i), (1 + i, -1 - i, i)\}$. **21.** $\{(1, 1, 1, 1), (-1, 1, 1, -1), (1, 0, 1, 1), (1, 1, 0, 1)\}$.

22. $\{(1, 2, 3, 1), (2, 1, -1, 1), (4, 5, 5, 3), (5, 4, 1, 3)\}$.

23. $\{(1, 2, 3, 4), (0, 1, -1, 2), (1, 4, 1, 8), (3, 7, 8, 14)\}$.

24. $\{(1, 1, 0, 1), (1, 1, 1, 1), (4, 4, 1, 1), (1, 0, 0, 1)\}$.

25. $\{(2, 2, 0, 2), (4, 1, 4, 1), (3, 0, 4, 0)\}$.

Determine which of the following systems are consistent and find all the solutions for the consistent systems.

26. $\begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & 2 \\ 2 & 1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix}$. **27.** $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

28. $\begin{bmatrix} 1 & -4 & 7 \\ 3 & 8 & -2 \\ 7 & -8 & 26 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \\ 3 \end{bmatrix}$. **29.** $\begin{bmatrix} 1 & 1 & 1 \\ 3 & -9 & 2 \\ 5 & -3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ 6 \end{bmatrix}$.

30. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ 16 \\ 22 \end{bmatrix}$. **31.** $\begin{bmatrix} 2 & 0 & -3 \\ 0 & 2 & -3 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

32. $\begin{bmatrix} 5 & 3 & 14 \\ 0 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 2 \end{bmatrix}$. **33.** $\begin{bmatrix} 1 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \\ 1 & 7 & -5 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$.

34. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$. **35.** $\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Find all the solutions of the following homogeneous systems $\mathbf{Ax} = \mathbf{0}$, where $\mathbf{A}$ is given as the following.

36. $\begin{bmatrix} 3 & 1 & 2 \\ 1 & -2 & 3 \\ 1 & 5 & -4 \end{bmatrix}$. **37.** $\begin{bmatrix} 1 & 1 & 2 \\ 3 & 4 & -7 \\ -1 & -2 & 11 \end{bmatrix}$. **38.** $\begin{bmatrix} 3 & -11 & 5 \\ 4 & 1 & -10 \\ 4 & 9 & -6 \end{bmatrix}$.

39. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 6 & 12 \end{bmatrix}$. **40.** $\begin{bmatrix} 2 & -1 & -3 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & -7 & -13 & -1 \\ -1 & 5 & 9 & 1 \end{bmatrix}$. **41.** $\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$.

42. $\begin{bmatrix} 3 & 1 & 1 & 4 \\ 0 & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{bmatrix}$. **43.** $\begin{bmatrix} 1 & 1 & -3 & 2 \\ 2 & -1 & -2 & -3 \\ 3 & 0 & -5 & -1 \\ 5 & -1 & -7 & -4 \end{bmatrix}$. **44.** $\begin{bmatrix} 1 & -2 & 1 & -1 \\ 1 & 1 & -2 & 3 \\ 4 & 1 & -5 & 8 \\ 5 & -7 & 2 & -1 \end{bmatrix}$.

45. $\begin{bmatrix} 1 & 1 & -2 & -1 \\ 2 & 1 & 1 & -2 \\ 3 & 2 & -1 & -3 \\ 4 & 2 & 2 & -4 \end{bmatrix}$.

Using the Gauss-Jordan method find the inverses of the following matrices.

46. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}$. **47.** $\begin{bmatrix} 2 & 3 & 1 \\ 1 & 3 & 3 \\ 0 & 1 & 2 \end{bmatrix}$. **48.** $\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

49. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$. **50.** $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 4 & 4 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$.

3.5 Eigenvalue Problems

Let $\mathbf{A} = (a_{ij})$ be a square matrix of order n. The matrix $\mathbf{A}$ may be singular or non-singular. Consider the homogeneous system of equations

$$\mathbf{Ax} = \lambda\mathbf{x} \quad \text{or} \quad (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0} \tag{3.39}$$

where λ is a scalar and $\mathbf{I}$ is an identity matrix of order n. The homogeneous system of equations (3.39) always has a trivial solution. We need to find values of λ for which the homogeneous system (3.39) has non-trivial solutions. The values of λ, for which non-trivial solutions of the homogeneous system (3.39) exist, are called the *eigenvalues* or the *characteristic values* of $\mathbf{A}$ and the corresponding non-trivial solution vectors $\mathbf{x}$ are called the *eigenvectors* or the *characteristic vectors* of $\mathbf{A}$. If $\mathbf{x}$ is a non-trivial solution of the homogeneous system (3.39), then $\alpha\mathbf{x}$, where α is any constant is also a solution of the homogeneous system. Hence, an eigenvector is unique only upto a constant multiple. The problem of determining the eigenvalues and the corresponding eigenvectors of a square matrix **A** is called an *eigenvalue problem*.

3.5.1 Eigenvalues and Eigenvectors

If the homogeneous system (3.39) has a non-trivial solution, then the rank of the coefficient matrix $(\mathbf{A} - \lambda \mathbf{I})$ is less than n, that is, the coefficient matrix must be singular. Therefore,

$$det\,(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{3.40}$$

Expanding the determinant given in Eq. (3.40), we obtain a polynomial of degree n in λ, which is of the form

$$P_n(\lambda) = |\,\mathbf{A} - \lambda\,\mathbf{I}\,| = (-1)^n[\lambda^n - c_1\,\lambda^{n-1} + c_2\,\lambda^{n-2} - \dots + (-1)^n c_n] = 0$$

or $$\lambda^n - c_1\,\lambda^{n-1} + c_2\,\lambda^{n-2} - \dots + (-1)^n c_n = 0. \tag{3.41}$$

where $c_1, c_2, \dots, c_n$ can be expressed in terms of the elements a_{ij} of the matrix **A**. This equation is called the *characteristic* equation of the matrix **A**. The polynomial equation $P_n(\lambda) = 0$ has n roots which can be real or complex, simple or repeated. The roots $\lambda_1, \lambda_2, \dots, \lambda_n$ of the polynomial equation $P_n(\lambda) = 0$ are called the *eigenvalues*. By using the relation between the roots and the coefficients, we can write

$$\begin{aligned} &\lambda_1 + \lambda_2 + \dots + \lambda_n = c_1 = a_{11} + a_{22} + \dots + a_{nn} \\ &\lambda_1\lambda_2 + \lambda_1\lambda_3 + \dots + \lambda_{n-1}\lambda_n = c_2 \\ &\vdots \\ &\lambda_1\lambda_2 \dots \lambda_n = c_n. \end{aligned}$$

If we set $\lambda = 0$ in Eq. (3.40), we get

$$|\,\mathbf{A}\,| = (-1)^{2n} c_n = c_n = \lambda_1\,\lambda_2 \dots \lambda_n. \tag{3.42}$$

Therefore, we get

$$\text{sum of eigenvalues} = trace\,(\mathbf{A}) \quad \text{and} \quad \text{product of eigenvalues} = |\,\mathbf{A}\,|.$$

The set of eigenvalues is called the *spectrum* of **A** and the largest eigenvalue in magnitude is called the *spectral radius* of **A** and is denoted by ρ (**A**). If $|\,\mathbf{A}\,| = 0$, that is the matrix is singular, then from Eq. (3.42), we find that one of the eigenvalues must be zero. Conversely, if one of the eigenvalues is zero, then $|\,\mathbf{A}\,| = 0$. Note that if **A** is a diagonal or an upper triangular or a lower triangular matrix, then the diagonal elements of the matrix **A** are the eigenvalues of **A**.

After determining the eigenvalues λ_i's, we solve the homogeneous system $(\mathbf{A} - \lambda_i\,\mathbf{I})\mathbf{x} = \mathbf{0}$ for each λ_i, $i = 1, 2, \dots, n$ to obtain the corresponding eigenvectors.

Properties of eigenvalues and eigenvectors

Let λ be an eigenvalue of **A** and **x** be its corresponding eigenvector. Then we have the following results.

1. α **A** has eigenvalue $\alpha\lambda$ and the corresponding eigenvector is **x**.

$$\mathbf{A}\mathbf{x} = \lambda\,\mathbf{x} \Rightarrow \alpha\mathbf{A}\mathbf{x} = (\alpha\lambda)\mathbf{x}.$$

2. $\mathbf{A}^m$ has eigenvalue λ^m and the corresponding eigenvector is $\mathbf{x}$ for any positive integer m.
Premultiplying both sides of $\mathbf{Ax} = \lambda \mathbf{x}$ by $\mathbf{A}$, we get

$$\mathbf{AAx} = \mathbf{A}\lambda\mathbf{x} = \lambda\mathbf{Ax} = \lambda(\lambda\mathbf{x}) \quad \text{or} \quad \mathbf{A}^2\mathbf{x} = \lambda^2\mathbf{x}.$$

Therefore, $\mathbf{A}^2$ has the eigenvalue λ^2 and the corresponding eigenvector is $\mathbf{x}$. Premultiplying successively m times, we obtain the result.

3. $\mathbf{A} - k\mathbf{I}$ has the eigenvalue $\lambda - k$, for any scalar k and the corresponding eigenvector is $\mathbf{x}$.

$$\mathbf{Ax} = \lambda\mathbf{x} \Rightarrow \mathbf{Ax} - k\,\mathbf{Ix} = \lambda\mathbf{x} - k\mathbf{x}$$

or
$$(\mathbf{A} - k\mathbf{I})\mathbf{x} = (\lambda - k)\,\mathbf{x}.$$

4. $\mathbf{A}^{-1}$ (if it exists) has the eigenvalue $1/\lambda$ and the corresponding eigenvector is $\mathbf{x}$.
Premultiplying both sides of $\mathbf{Ax} = \lambda\mathbf{x}$ by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}\mathbf{Ax} = \lambda\mathbf{A}^{-1}\mathbf{x} \quad \text{or} \quad \mathbf{A}^{-1}\mathbf{x} = (1/\lambda)\mathbf{x}.$$

5. $(\mathbf{A} - k\mathbf{I})^{-1}$ has the eigenvalue $1/(\lambda - k)$ and the corresponding eigenvector is $\mathbf{x}$ for an scalar k.

6. $\mathbf{A}$ and $\mathbf{A}^T$ have the same eigenvalues, since a determinant can be expanded by rows columns.

7. For a real matrix $\mathbf{A}$, if $\alpha + i\beta$ is an eigenvalue, then its conjugate $\alpha - i\beta$ is also an eigenval (since the characteristic equation has real coefficients). When the matrix $\mathbf{A}$ is complex, th property does not hold.

We now present an important result which gives the relationship of a matrix $\mathbf{A}$ and its characteris equation.

Theorem 3.9 **(Cayley-Hamilton theorem)** Every square matrix $\mathbf{A}$ satisfies its own characteris equation, that is

$$\mathbf{A}^n - c_1\mathbf{A}^{n-1} + \ldots + (-1)^{n-1}c_{n-1}\mathbf{A} + (-1)^n c_n\mathbf{I} = \mathbf{0}. \tag{3.4}$$

Proof The cofactors of the elements of the determinant $|\mathbf{A} - \lambda\mathbf{I}|$ are polynomials in λ of degr $(n-1)$ or less. Therefore, the elements of the adjoint matrix (transpose of the cofactor matrix) are al polynomials in λ of degree $(n-1)$ or less. Hence, we can express the adjoint matrix as a polynom in λ whose coefficients $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$ are square matrices of order n having elements as functio of the elements of the matrix $\mathbf{A}$. Thus, we can write

$$adj\,(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{B}_1\lambda^{n-1} + \mathbf{B}_2\lambda^{n-2} + \ldots + \mathbf{B}_{n-1}\lambda + \mathbf{B}_n.$$

We also have

$$(\mathbf{A} - \lambda\mathbf{I})\,adj\,(\mathbf{A} - \lambda\mathbf{I}) = |\mathbf{A} - \lambda\mathbf{I}|\,\mathbf{I}.$$

Therefore, we can write for any λ

$$(\mathbf{A} - \lambda\mathbf{I})(\mathbf{B}_1\lambda^{n-1} + \mathbf{B}_2\lambda^{n-2} + \ldots + \mathbf{B}_{n-1}\lambda + \mathbf{B}_n) = \lambda^n\mathbf{I} - c_1\lambda^{n-1}\mathbf{I} + \ldots + (-1)^{n-1}c_{n-1}\lambda\mathbf{I} + (-1)^n c$$

Comparing the coefficients of various powers of λ, we obtain

$$-\mathbf{B}_1 = \mathbf{I}$$
$$\mathbf{AB}_1 - \mathbf{B}_2 = -c_1\mathbf{I}$$
$$\mathbf{AB}_2 - \mathbf{B}_3 = c_2\mathbf{I}$$

$$\ldots$$

$$\mathbf{A}\mathbf{B}_{n-1} - \mathbf{B}_n = (-1)^{n-1} c_{n-1}\mathbf{I}$$

$$\mathbf{A}\mathbf{B}_n = (-1)^n c_n \mathbf{I}.$$

Premultiplying these equations by $\mathbf{A}^n, \mathbf{A}^{n-1}, \ldots, \mathbf{A}, \mathbf{I}$ respectively and adding, we get

$$\mathbf{A}^n - c_1\mathbf{A}^{n-1} + \ldots + (-1)^{n-1} c_{n-1}\mathbf{A} + (-1)^n c_n \mathbf{I} = \mathbf{0}$$

which proves the theorem.

Remark 16a We have for any non-zero vector $\mathbf{x}$

$$\mathbf{I}\mathbf{x} = 1\mathbf{x}$$

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$\mathbf{A}^2\mathbf{x} = \lambda^2\mathbf{x}$$

$$\ldots$$

$$\mathbf{A}^n\mathbf{x} = \lambda^n\mathbf{x}.$$

Multiplying these equations by $(-1)^n c_n, (-1)^{n-1} c_{n-1}, \ldots, (-1)c_1$, 1 respectively and adding, we get

$$(-1)^n c_n \mathbf{I}\mathbf{x} + (-1)^{n-1} c_{n-1}\mathbf{A}\mathbf{x} + \ldots + (-1)^1 c_1 \mathbf{A}^{n-1}\mathbf{x} + \mathbf{A}^n\mathbf{x}$$

$$= (-1)^n c_n \mathbf{x} + (-1)^{n-1} c_{n-1}\lambda\mathbf{x} + \ldots + (-1)^1 c_1 \lambda^{n-1}\mathbf{x} + \lambda^n\mathbf{x}$$

or

$$[\mathbf{A}^n - c_1\mathbf{A}^{n-1} + \ldots + (-1)^{n-1} c_{n-1}\mathbf{A} + (-1)^n c_n \mathbf{I}]\mathbf{x}$$

$$= [\lambda^n - c_1\lambda^{n-1} + \ldots + (-1)^{n-1} c_{n-1}\lambda + (-1)^n c_n]\mathbf{x} = 0\mathbf{x} = \mathbf{0}.$$

Since $\mathbf{x} \neq \mathbf{0}$, there is a possibility that

$$\mathbf{A}^n - c_1\mathbf{A}^{n-1} + \ldots + (-1)^{n-1} c_{n-1}\mathbf{A} + (-1)^n c_n \mathbf{I} = \mathbf{0}.$$

Remark 16b

(a) We can use Eq. (3.43) to find $\mathbf{A}^{-1}$ (if it exists) in terms of the powers of the matrix $\mathbf{A}$.

Premultiplying both sides in Eq. (3.43) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}\mathbf{A}^n - c_1\mathbf{A}^{-1}\mathbf{A}^{n-1} + \ldots + (-1)^{n-1} c_{n-1}\mathbf{A}^{-1}\mathbf{A} + (-1)^n c_n \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$$

or

$$\mathbf{A}^{-1} = -\frac{(-1)^n}{c_n}[\mathbf{A}^{n-1} - c_1\mathbf{A}^{n-2} + \ldots + (-1)^{n-1} c_{n-1}\mathbf{I}] \tag{3.44}$$

(b) We can use Eq. (3.43) to obtain $\mathbf{A}^n$ in terms of lower powers of $\mathbf{A}$ as

$$\mathbf{A}^n = c_1\mathbf{A}^{n-1} - c_2\mathbf{A}^{n-2} + \ldots + (-1)^{n-1} c_n \mathbf{I}. \tag{3.45}$$

Example 3.39 Verify Cayley-Hamilton theorem for the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Also (i) obtain $\mathbf{A}^{-1}$ and $\mathbf{A}^3$, (ii) find eigenvalues of $\mathbf{A}$, $\mathbf{A}^2$ and verify that eigenvalues of $\mathbf{A}^2$ are squares of those of **A**, (iii) find the spectral radius of **A**.

Solution The characteristic equation of **A** is given by

$$|\mathbf{A}-\lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ -1 & 1-\lambda & 2 \\ 1 & 2 & 1-\lambda \end{vmatrix} = (1-\lambda)\{(1-\lambda)^2-4\} - 2\{-(1-\lambda)-2\}$$

$$= (1-\lambda)(\lambda^2-2\lambda-3) - 2(\lambda-3) = -\lambda^3+3\lambda^2-\lambda+3 = 0.$$

Now,
$$\mathbf{A}^2 = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 4 & 4 \\ 0 & 3 & 4 \\ 0 & 6 & 5 \end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{A}^2\mathbf{A} = \begin{bmatrix} -1 & 4 & 4 \\ 0 & 3 & 4 \\ 0 & 6 & 5 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 10 & 12 \\ 1 & 11 & 10 \\ -1 & 16 & 17 \end{bmatrix}.$$

We have

$$-\mathbf{A}^3+3\mathbf{A}^2-\mathbf{A}+3\mathbf{I} = -\begin{bmatrix} -1 & 10 & 12 \\ 1 & 11 & 10 \\ -1 & 16 & 17 \end{bmatrix} + 3\begin{bmatrix} -1 & 4 & 4 \\ 0 & 3 & 4 \\ 0 & 6 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} + 3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}. \qquad (3.46$$

Hence, **A** satisfies the characteristic equation $-\lambda^3+3\lambda^2-\lambda+3=0$.

(i) From Eq. (3.46), we get

$$\mathbf{A}^{-1} = \frac{1}{3}[\mathbf{A}^2-3\mathbf{A}+\mathbf{I}] = \frac{1}{3}\left[\begin{pmatrix} -1 & 4 & 4 \\ 0 & 3 & 4 \\ 0 & 6 & 5 \end{pmatrix} - \begin{pmatrix} 3 & 6 & 0 \\ -3 & 3 & 6 \\ 3 & 6 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$$

$$= \frac{1}{3}\begin{bmatrix} -3 & -2 & 4 \\ 3 & 1 & -2 \\ -3 & 0 & 3 \end{bmatrix}.$$

From Eq. (3.46), we get

$$\mathbf{A}^3 = 3\mathbf{A}^2-\mathbf{A}+3\mathbf{I} = \begin{bmatrix} -3 & 12 & 12 \\ 0 & 9 & 12 \\ 0 & 18 & 15 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 10 & 12 \\ 1 & 11 & 10 \\ -1 & 16 & 17 \end{bmatrix}.$$

(ii) Eigenvalues of **A** are the roots of

$$\lambda^3 - 3\lambda^2 + \lambda - 3 = 0 \quad \text{or} \quad (\lambda - 3)(\lambda^2 + 1) = 0 \quad \text{or} \quad \lambda = 3, i, -i.$$

The characteristic equation of $\mathbf{A}^2$ is given by

$$\begin{vmatrix} -1-\lambda & 4 & 4 \\ 0 & 3-\lambda & 4 \\ 0 & 6 & 5-\lambda \end{vmatrix} = (-1-\lambda)[(3-\lambda)(5-\lambda) - 24] = 0$$

or
$$(\lambda + 1)(\lambda^2 - 8\lambda - 9) = 0 \quad \text{or} \quad (\lambda + 1)(\lambda - 9)(\lambda + 1) = 0.$$

The eigenvalues of $\mathbf{A}^2$ are $9, -1, -1$ which are the squares of the eigenvalues of $\mathbf{A}$.

(iii) The spectral radius of $\mathbf{A}$ is given by

$$\rho(\mathbf{A}) = \text{largest eigenvalue in magnitude} = \max_i |\lambda_i| = 3.$$

Example 3.40 If $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, then show that $\mathbf{A}^n = \mathbf{A}^{n-2} + \mathbf{A}^2 - \mathbf{I}$ for $n \geq 3$. Hence, find $\mathbf{A}^{50}$.

Solution The characteristic equation of $\mathbf{A}$ is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 1) = 0, \quad \text{or} \quad \lambda^3 - \lambda^2 - \lambda + 1 = 0.$$

Using Cayley-Hamilton theorem, we get

$$\mathbf{A}^3 - \mathbf{A}^2 - \mathbf{A} + \mathbf{I} = \mathbf{0}, \quad \text{or} \quad \mathbf{A}^3 - \mathbf{A}^2 = \mathbf{A} - \mathbf{I}.$$

Premultiplying both sides successively by $\mathbf{A}$, we obtain

$$\begin{aligned} \mathbf{A}^3 - \mathbf{A}^2 &= \mathbf{A} - \mathbf{I} \\ \mathbf{A}^4 - \mathbf{A}^3 &= \mathbf{A}^2 - \mathbf{A} \\ &\dots \\ \mathbf{A}^{n-1} - \mathbf{A}^{n-2} &= \mathbf{A}^{n-3} - \mathbf{A}^{n-4} \\ \mathbf{A}^n - \mathbf{A}^{n-1} &= \mathbf{A}^{n-2} - \mathbf{A}^{n-3}. \end{aligned}$$

Adding these equations, we get

$$\mathbf{A}^n - \mathbf{A}^2 = \mathbf{A}^{n-2} - \mathbf{I}, \quad \text{or} \quad \mathbf{A}^n = \mathbf{A}^{n-2} + \mathbf{A}^2 - \mathbf{I}, \; n \geq 3.$$

Using this equation recursively, we obtain

$$\begin{aligned} \mathbf{A}^n &= (\mathbf{A}^{n-4} + \mathbf{A}^2 - \mathbf{I}) + \mathbf{A}^2 - \mathbf{I} = \mathbf{A}^{n-4} + 2(\mathbf{A}^2 - \mathbf{I}) \\ &= (\mathbf{A}^{n-6} + \mathbf{A}^2 - \mathbf{I}) + 2(\mathbf{A}^2 - \mathbf{I}) = \mathbf{A}^{n-6} + 3(\mathbf{A}^2 - \mathbf{I}) \\ &\dots \\ &= \mathbf{A}^{n-(n-2)} + \frac{1}{2}(n-2)(\mathbf{A}^2 - \mathbf{I}) = \frac{n}{2}\mathbf{A}^2 - \frac{1}{2}(n-2)\mathbf{I}. \end{aligned}$$

Substituting $n = 50$, we get

$$\mathbf{A}^{50} = 25\mathbf{A}^2 - 24\mathbf{I} = 25\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} - 24\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 25 & 1 & 0 \\ 25 & 0 & 1 \end{bmatrix}.$$

Example 3.41 Find the eigenvalues and the corresponding eigenvectors of the following matrices

(i) $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$, (ii) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, (iii) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$.

Solution

(i) The characteristic equation of $\mathbf{A}$ is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 4 \\ 3 & 2-\lambda \end{vmatrix} = 0 \quad \text{or} \quad \lambda^2 - 3\lambda - 10 = 0, \quad \text{or} \quad \lambda = -2, 5.$$

Corresponding to the eigenvalue $\lambda = -2$, we have

$$(\mathbf{A} + 2\mathbf{I})\mathbf{x} = \begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad 3x_1 + 4x_2 = 0 \quad \text{or} \quad x_1 = -\frac{4}{3}x_2.$$

Hence, the eigenvector $\mathbf{x}$ is given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4x_2/3 \\ x_2 \end{bmatrix} = x_2\begin{bmatrix} -4/3 \\ 1 \end{bmatrix}.$$

Since an eigenvector is unique upto a constant multiple, we can take the eigenvector as $[-4, 3]^T$.

Corresponding to the eigenvalue $\lambda = 5$, we have

$$(\mathbf{A} - 5\mathbf{I})\mathbf{x} = \begin{pmatrix} -4 & 4 \\ 3 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad x_1 - x_2 = 0, \quad \text{or} \quad x_1 = x_2.$$

Therefore, the eigenvector is given by $\mathbf{x} = (x_1, x_2)^T = x_1(1, 1)^T$ or simply $(1, 1)^T$.

(ii) The characteristic equation of $\mathbf{A}$ is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = 0, \quad \text{or} \quad \lambda^2 - 2\lambda + 2 = 0, \quad \text{or} \quad \lambda = 1 \pm i.$$

Corresponding to the eigenvalue $\lambda = 1 + i$, we have

$$[\mathbf{A} - (1+i)\mathbf{I}]\mathbf{x} = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$-ix_1 + x_2 = 0 \quad \text{and} \quad -x_1 - ix_2 = 0.$$

Both the equations reduce to $-x_1 - ix_2 = 0$. Choosing $x_2 = 1$, we get $x_1 = -i$. Therefore, the eigenvector is $\mathbf{x} = [-i, 1]^T$.

Corresponding to the eigenvalue $\lambda = 1 - i$, we have

$$[\mathbf{A} - (1 - i)\mathbf{I}]\mathbf{x} = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or $$ix_1 + x_2 = 0 \quad \text{and} \quad -x_1 + ix_2 = 0.$$

Both the equations reduce to $-x_1 + ix_2 = 0$. Choosing $x_2 = 1$, we get $x_1 = i$. Therefore, the eigenvector is $\mathbf{x} = [i, 1]^T$.

Remark 17

For a real matrix **A**, the eigenvalues and the corresponding eigenvectors can be complex.

(iii) The characteristic equation of **A** is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{vmatrix} = 0 \quad \text{or} \quad (1-\lambda)(2-\lambda)(3-\lambda) = 0$$

or $\lambda = 1, 2, 3$.

Corresponding to the eigenvalue $\lambda = 1$, we have

$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{cases} x_2 + x_3 = 0 \\ x_1 + x_3 = 0. \end{cases}$$

We obtain two equations in three unknowns. One of the variables x_1, x_2, x_3 can be chosen arbitrarily. Taking $x_3 = 1$, we obtain the eigenvector as $[-1, -1, 1]^T$.

Corresponding to the eigenvalue $\lambda = 2$, we have

$$(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or $x_1 = 0$, $x_3 = 0$ and x_2 arbitrary. Taking $x_2 = 1$, we obtain the eigenvector as $[0, 1, 0]^T$.

Corresponding to the eigenvalue $\lambda = 3$, we have

$$(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{cases} x_1 = 0 \\ x_2 + x_3 = 0. \end{cases}$$

Choosing $x_3 = 1$, we obtain the eigenvector as $[0, -1, 1]^T$.

Example 3.42 Find the eigenvalues and the corresponding eigenvectors of the following matrices

(i) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, (ii) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, (iii) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Solution In each of the above problems, we obtain the characteristic equation as $(1 - \lambda)^3 = 0$. Therefore, the eigenvalues are $\lambda = 1, 1, 1$, a repeated value. Since a 3×3 matrix has 3 eigenvalues, it is important to know, whether the given matrix has 3 linearly independent eigenvectors or it has lesser number of linearly independent eigenvectors.

Corresponding to the eigenvalue $\lambda = 1$, we obtain the following eigenvectors.

(i)
$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} x_2 = 0, \\ x_3 = 0 \\ x_1 \text{ arbitrary.} \end{cases}$$

Choosing $x_1 = 1$, we obtain the solution as $[1, 0, 0]^T$.
Hence, **A** has only one independent eigenvector.

(ii)
$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{or} \quad \begin{cases} x_2 = 0 \\ x_1, x_3 \text{ arbitrary.} \end{cases}$$

Taking $x_1 = 0, x_3 = 1$ and $x_1 = 1, x_3 = 0$, we obtain two linearly independent solutions

$$\mathbf{x}_1 = [0, 0, 1]^T, \mathbf{x}_2 = [1, 0, 0]^T.$$

In this case **A** has two linearly independent eigenvectors.

(iii)
$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This system is satisfied for arbitrary values of all the three variables. Hence, we obtain three linearly independent eigenvectors, which can be written as

$$\mathbf{x}_1 = [1, 0, 0]^T, \mathbf{x}_2 = [0, 1, 0]^T, \mathbf{x}_3 = [0, 0, 1]^T.$$

We now state some important results regarding the relationship between the eigenvalues of a matrix and the corresponding linearly independent eigenvectors.

1. Eigenvectors corresponding to distinct eigenvalues are linearly independent.
2. If λ is an eigenvalue of multiplicity m of a square matrix **A** of order n, then the number of linearly independent eigenvectors associated with λ is given by

$$p = n - r, \quad \text{where} \quad r = \text{rank } (\mathbf{A} - \lambda \mathbf{I}),\ 1 \le p \le m.$$

Remark 18

In Example 3.41, all the eigenvalues are distinct and therefore, the corresponding eigenvectors are linearly independent. In Example 3.42 the eigenvalue $\lambda = 1$ is of multiplicity 3. We find that in

(i) Example 3.42(i), the rank of the matrix **A** – **I** is 2 and we obtain one linearly independent eigenvector.

(ii) Example 3.42(ii), the rank of the matrix **A** – **I** is 1 and we obtain two linearly independent eigenvectors.

(iii) Example 3.42(iii), the rank of the matrix $\mathbf{A} - \mathbf{I}$ is 0 and we obtain three linearly independent eigenvectors.

3.5.2 Similar and Diagonalizable Matrices

Similar matrices

Let **A** and **B** be square matrices of the same order. The matrix **A** is said to be similar to the matrix **B** if there exists an invertible matrix **P** such that

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} \quad \text{or} \quad \mathbf{P}\mathbf{A} = \mathbf{B}\mathbf{P}. \tag{3.47}$$

Postmultiplying both sides in Eq. (3.47) by $\mathbf{P}^{-1}$, we get

$$\mathbf{P}\mathbf{A}\mathbf{P}^{-1} = \mathbf{B}.$$

Therefore, **A** is similar to **B** if and only if **B** is similar to **A**. The matrix **P** is called the *similarity matrix*. We now prove a result regarding eigenvalues of similar matrices.

Theorem 3.10 Similar matrices have the same characteristic equation (and hence the same eigenvalues). Further, if $\mathbf{x}$ is an eigenvector of **A** corresponding to the eigenvalue λ, then $\mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of **B** corresponding to the eigenvalue λ, where **P** is the similarity matrix.

Proof Let λ be an eigenvalue and $\mathbf{x}$ be the corresponding eigenvector of **A**. That is

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Premultiplying both sides by an invertible matrix $\mathbf{P}^{-1}$, we obtain

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{x} = \lambda\mathbf{P}^{-1}\mathbf{x}.$$

Set $\mathbf{x} = \mathbf{P}\mathbf{y}$. We get

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{y} = \lambda\mathbf{P}^{-1}\mathbf{P}\mathbf{y}, \quad \text{or} \quad (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\mathbf{y} = \lambda\,\mathbf{y} \quad \text{or} \quad \mathbf{B}\mathbf{y} = \lambda\mathbf{y}.$$

where $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Therefore, **B** has the same eigenvalues as **A**, that is the characteristic equation of **B** is same as the characteristic equation of **A**. Now, **A** and **B** are similar matrices. Therefore, similar matrices have the same characteristic equation (and hence the same eigenvalues). Also $\mathbf{x} = \mathbf{P}\mathbf{y}$, that is eigenvectors of **A** and **B** are related by $\mathbf{x} = \mathbf{P}\mathbf{y}$ or $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$.

Remark 19

(a) Theorem 3.10 states that if two matrices are similar, then they have the same characteristic equation and hence the same eigenvalues. However, the converse of this theorem is not true. Two matrices which have the same characteristic equation need not always be similar.

(b) If **A** is similar to **B** and **B** is similar to **C**, then **A** is similar to **C**.
Let there be two invertible matrices **P** and **Q** such that

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P} \quad \text{and} \quad \mathbf{B} = \mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}.$$

Then $$\mathbf{A} = \mathbf{P}^{-1}\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}\mathbf{P} = \mathbf{R}^{-1}\mathbf{C}\mathbf{R}, \quad \text{where} \quad \mathbf{R} = \mathbf{Q}\mathbf{P}.$$

Example 3.43 Examine whether **A** is similar to **B**, where

(i) $\mathbf{A} = \begin{bmatrix} 5 & 5 \\ -2 & 0 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}$.

(ii) $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Solution The given matrices are similar if there exists an invertible matrix **P** such that

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{BP} \quad \text{or} \quad \mathbf{PA} = \mathbf{BP}.$$

Let $\mathbf{P} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We shall determine a, b, c, d such that **PA** = **BP** and then check whether **P** is non-singular.

(i) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 5 & 5 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, or $\begin{bmatrix} 5a-2b & 5a \\ 5c-2d & 5c \end{bmatrix} = \begin{bmatrix} a+2c & b+2d \\ -3a+4c & -3b+4d \end{bmatrix}$.

Equating the corresponding elements, we obtain the system of equations

$$\begin{aligned} 5a - 2b &= a + 2c, &\quad \text{or} \quad && 4a - 2b - 2c = 0 \\ 5a &= b + 2d, &\quad \text{or} \quad && 5a - b - 2d = 0 \\ 5c - 2d &= -3a + 4c, &\quad \text{or} \quad && 3a + c - 2d = 0 \\ 5c &= -3b + 4d, &\quad \text{or} \quad && 3b + 5c - 4d = 0. \end{aligned}$$

A solution to this system of equations is $a = 1, b = 1, c = 1, d = 2$.

Therefore, we get $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, which is a non-singular matrix. Hence the matrices **A** and **B** are similar.

(ii) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, or $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ c & d \end{bmatrix}$.

Equating the corresponding elements, we get

$$a = a + c,\ b = b + d \quad \text{or} \quad c = d = 0.$$

Therefore, $\mathbf{P} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$, which is a singular matrix.

Since an invertible matrix **P** does not exist, the matrices **A** and **B** are not similar.

It can be verified that the eigenvalues of **A** are 1, 1 whereas the eigenvalues of **B** are 0, 2.

In practice, it is usually difficult to obtain a non-singular matrix **P** which satisfies the equation $\mathbf{A} = \mathbf{P}^{-1}\mathbf{BP}$ for any two matrices **A** and **B**. However, it is possible to obtain the matrix **P** when **A** or **B** is a diagonal matrix. Thus, our interest is to find a similarity matrix **P** such that for a given matrix **A**, we have

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP} \quad \text{or} \quad \mathbf{PDP}^{-1} = \mathbf{A}$$

where **D** is a diagonal matrix. If such a matrix exists, then we say that the matrix **A** is *diagonalizable*.

Diagonalizable matrices

A matrix $\mathbf{A}$ is diagonalizable, if it is similar to a diagonal matrix, that is there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix. Since, similar matrices have the same eigenvalues, the diagonal elements of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$. A necessary and sufficient condition for the existence of $\mathbf{P}$ is given in the following theorem.

Theorem 3.11 A square matrix $\mathbf{A}$ of order n is diagonalizable if and only if it has n linearly independent eigenvectors.

Proof We shall prove the case that if $\mathbf{A}$ has n linearly independent eigenvectors, then $\mathbf{A}$ is diagonalizable. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be n linearly independent eigenvectors corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (not necessarily distinct) of the matrix $\mathbf{A}$ in the same order, that is the eigenvector $\mathbf{x}_j$ corresponds to the eigenvalue λ_j, $j = 1, 2, \ldots, n$. Let

$$\mathbf{P} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \quad \text{and} \quad \mathbf{D} = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

be the diagonal matrix with eigenvalues of $\mathbf{A}$ as its diagonal elements. The matrix $\mathbf{P}$ is called the *modal matrix* of $\mathbf{A}$ and $\mathbf{D}$ is called the *spectral matrix* of $\mathbf{A}$. We have

$$\begin{aligned} \mathbf{AP} &= \mathbf{A}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = (\mathbf{Ax}_1, \mathbf{Ax}_2, \ldots, \mathbf{Ax}_n) \\ &= (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \ldots, \lambda_n\mathbf{x}_n) = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\mathbf{D} = \mathbf{PD}. \end{aligned} \tag{3.48}$$

Since the columns of $\mathbf{P}$ are linearly independent, the rank of $\mathbf{P}$ is n and therefore the matrix P is invertible. Premultiplying both sides in Eq. (3.48) by $\mathbf{P}^{-1}$, we obtain

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{P}^{-1}\mathbf{PD} = \mathbf{D} \tag{3.49}$$

which implies that $\mathbf{A}$ is similar to $\mathbf{D}$. Therefore, the matrix of eigenvectors $\mathbf{P}$ reduces a matrix $\mathbf{A}$ to its diagonal form.

Postmultiplying both sides in Eq. (3.48) by $\mathbf{P}^{-1}$, we obtain

$$\mathbf{A} = \mathbf{PDP}^{-1}. \tag{3.50}$$

Remark 20

(a) A square matrix $\mathbf{A}$ of order n has always n linearly indendent eigenvectors when its eigenvalues are distinct. The matrix may also have n linearly independent eigenvectors even when some eigenvalues are repeated (see Example 3.42(iii)). Therefore, there is no restriction imposed on the eigenvalues of the matrix $\mathbf{A}$ in Theorem 3.11.

(b) From Eq. (3.50), we obtain

$$\mathbf{A}^2 = \mathbf{AA} = (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1}) = \mathbf{PD}^2\mathbf{P}^{-1}.$$

Repeating the pre-multiplication (post-multiplication) m times, we get

$$\mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1} \text{ for any positive integer } m.$$

Therefore, if $\mathbf{A}$ is diagonalizable, so is $\mathbf{A}^m$.

(c) If $\mathbf{D}$ is a diagonal matrix of order n, and

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{bmatrix}, \text{ then } \mathbf{D}^m = \begin{bmatrix} \lambda_1^m & & & \mathbf{0} \\ & \lambda_2^m & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n^m \end{bmatrix}$$

for any positive integer m. If $Q(\mathbf{D})$ is a polynomial in $\mathbf{D}$, then we get

$$Q(\mathbf{D}) = \begin{bmatrix} Q(\lambda_1) & & & \mathbf{0} \\ & Q(\lambda_2) & & \\ & & \ddots & \\ \mathbf{0} & & & Q(\lambda_n) \end{bmatrix}.$$

Now, let a matrix $\mathbf{A}$ be diagonalizable. Then, we have

$$\mathbf{A} = \mathbf{PDP}^{-1} \quad \text{and} \quad \mathbf{A}^m = \mathbf{PD}^m\mathbf{P}^{-1}$$

for any positive integer m. Hence, we obtain

$$Q(\mathbf{A}) = \mathbf{P}Q(\mathbf{D})\mathbf{P}^{-1}$$

for any matrix polynomial $Q(\mathbf{A})$.

Example 3.44 Show that the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

is diagonalizable. Hence, find $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix. Then, obtain the matrix $\mathbf{B} = \mathbf{A}^2 + 5\mathbf{A} + 3\mathbf{I}$.

Solution The characteristic equation of $\mathbf{A}$ is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 3-\lambda & 1 & -1 \\ -2 & 1-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{vmatrix} = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0, \quad \text{or} \quad \lambda = 1, 2, 3.$$

Since the matrix $\mathbf{A}$ has three distinct eigenvalues, it has three linearly independent eigenvectors and hence it is diagonalizable.

The eigenvector corresponding to the eigenvalue $\lambda = 1$ is the solution of the system

$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \text{ The solution is } \mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

The eigenvector corresponding to the eigenvalue $\lambda = 2$ is the solution of the system

$$(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \begin{bmatrix} 1 & 1 & -1 \\ -2 & -1 & 2 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \text{ The solution is } \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

The eigenvector corresponding to the eigenvalue $\lambda = 3$ is the solution of the system

$$(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \begin{bmatrix} 0 & 1 & -1 \\ -2 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \text{ The solution is } \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Hence, the modal matrix is given by

$$\mathbf{P} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{P}^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

It can be verified that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}\,(1, 2, 3)$.

We have $\mathbf{D} = \text{diag}\,(1, 2, 3)$, $\mathbf{D}^2 = \text{diag}\,(1, 4, 9)$.

Therefore,
$$\mathbf{A}^2 + 5\mathbf{A} + 3\mathbf{I} = \mathbf{P}(\mathbf{D}^2 + 5\mathbf{D} + 3\mathbf{I})\mathbf{P}^{-1}$$

Now,
$$\mathbf{D}^2 + 5\mathbf{D} + 3\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} + \begin{bmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 15 \end{bmatrix} + \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 27 \end{bmatrix}.$$

Hence, we obtain

$$\mathbf{A}^2 + 5\mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 27 \end{bmatrix}\begin{bmatrix} -1 & -1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 25 & 8 & -8 \\ -18 & 9 & 18 \\ -2 & 8 & 19 \end{bmatrix}.$$

Example 3.45 Examine whether the matrix $\mathbf{A}$, where $\mathbf{A}$ is given by

(i) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$.   (ii) $\mathbf{A} = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$.

is diagonalizable. If so, obtain the matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a diagonal matrix.

Solution

(i) The characteristic equation of the matrix $\mathbf{A}$ is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 2 & 2 \\ 0 & 2-\lambda & 1 \\ -1 & 2 & 2-\lambda \end{vmatrix}$$

$$= (1-\lambda)[(2-\lambda)(2-\lambda) - 2] - [2 - 2(2-\lambda)] = (1-\lambda)(2-\lambda)(2-\lambda) = 0,$$

or $\lambda = 1, 2, 2$. We first find the eigenvectors corresponding to the repeated eigenvalue $\lambda = 2$. We have the system

$$(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \begin{bmatrix} -1 & 2 & 2 \\ 0 & 0 & 1 \\ -1 & 2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Since the rank of the coefficient matrix is 2, it has one linearly independent eigenvector. We obtain another linearly independent eigenvector corresponding to the eigenvalue $\lambda = 1$. Since the matrix **A** has only two linearly independent eigenvectors, the matrix is not diagonalizable.

(ii) The characteristic equation of the matrix **A** is given by

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = 0 \text{ or } \lambda^3 + \lambda^2 - 21\lambda - 45 = 0, \text{ or } \lambda = 5, -3, -3.$$

Eigenvector corresponding to the eigenvalue $\lambda = 5$ is the solution of the system

$$(\mathbf{A} - 5\mathbf{I})\mathbf{x} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A solution of this system is $[1, 2, -1]^T$.

Eigenvectors corresponding to $\lambda = -3$ are the solutions of the system

$$(\mathbf{A} + 3\mathbf{I})\mathbf{x} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ or } x_1 + 2x_2 - 3x_3 = 0.$$

The rank of the coefficient matrix is 1. Therefore, the system has two linearly independent solutions. We use the equation $x_1 + 2x_2 - 3x_3 = 0$ to find two linearly independent eigenvectors. Taking $x_3 = 0, x_2 = 1$, we obtain the eigenvector $[-2, 1, 0]^T$ and taking $x_2 = 0, x_3 = 1$, we obtain the eigenvector $[3, 0, 1]^T$. The given 3×3 matrix has three linearly independent eigenvectors. Therefore, the matrix **A** is diagonalizable. The modal matrix **P** is given by

$$\mathbf{P} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{P}^{-1} = \frac{1}{8}\begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix}.$$

It can be verified that $\mathbf{P}^{-1}\mathbf{AP} = \text{diag}(5, -3, -3)$.

Example 3.46 The eigenvectors of a 3×3 matrix **A** corresponding to the eigenvalues 1, 1, 3 are $[1, 0, -1]^T$, $[0, 1, -1]^T$ and $[1, 1, 0]^T$ respectively. Find the matrix **A**.

Solution We have

$$\text{modal matrix } \mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix} \text{ and the spectral matrix } \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

We find that $$\mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Therefore, $$\mathbf{A} = \mathbf{P}\,\mathbf{D}\,\mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

3.5.3 Special Matrices

In this section we define some special matrices and study the properties of the eigenvalues and eigenvectors of these matrices. These matrices have applications in many areas. We first give some definitions.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ be two vectors of dimension n in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then we define the following:

Inner Product (dot product) of vectors Let $\mathbf{x}$ and $\mathbf{y}$ be two vectors in $\mathbb{R}^n$. Then

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i \tag{3.51}$$

is called the *inner product* of the vectors $\mathbf{x}$ and $\mathbf{y}$ and is a scalar. The inner product is also denoted by $< \mathbf{x}, \mathbf{y} >$. In this case $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$. Note that $\mathbf{x} \cdot \mathbf{x} \geq 0$ and $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{C}^n$, then the inner product of these vectors is defined as

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\,\overline{\mathbf{y}} = \sum_{i=1}^{n} x_i \overline{y}_i \quad \text{and} \quad \mathbf{y} \cdot \mathbf{x} = \mathbf{y}^T\,\overline{\mathbf{x}} = \sum_{i=1}^{n} y_i \overline{x}_i$$

where $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are complex conjugate vectors of $\mathbf{x}$ and $\mathbf{y}$ respectively. Note that $\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}}$. It can be easily verified that

$$(\alpha\mathbf{x} + \beta\mathbf{y}) \cdot \mathbf{z} = \alpha(\mathbf{x} \cdot \mathbf{z}) + \beta(\mathbf{y} \cdot \mathbf{z})$$

for any vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and scalars α, β.

Length (norm of a vector) Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

is called the *length* or the *norm* of the vector $\mathbf{x}$.

Unit vector The vector $\mathbf{x}$ is called a *unit vector* if $\|\mathbf{x}\| = 1$. If $\mathbf{x} \neq \mathbf{0}$, then the vector $\mathbf{x}/\|\mathbf{x}\|$ is always a unit vector.

Orthogonal vectors The vectors $\mathbf{x}$ and $\mathbf{y}$ for which $\mathbf{x} \cdot \mathbf{y} = 0$ are said to be *orthogonal vectors*.

Orthonormal vectors The vectors of $\mathbf{x}$ and $\mathbf{y}$ for which

$$\mathbf{x} \cdot \mathbf{y} = 0 \quad \text{and} \quad \| \mathbf{x} \| = 1, \| \mathbf{y} \| = 1$$

are called orthonormal vectors. If $\mathbf{x}$, $\mathbf{y}$ are any vectors and $\mathbf{x} \cdot \mathbf{y} = 0$, then $\mathbf{x}/\| \mathbf{x} \|$, $\mathbf{y}/\| \mathbf{y} \|$ a orthonormal.

For example, the set of vectors

(i) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ form an orthonormal set in $\mathbb{R}^3$.

(ii) $\begin{pmatrix} 3i \\ -4i \\ 0 \end{pmatrix}, \begin{pmatrix} -4i \\ 3i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix}$ form an orthogonal set in $\mathbb{C}^3$ and $\begin{pmatrix} 3i/5 \\ -4i/5 \\ 0 \end{pmatrix}, \begin{pmatrix} -4i/5 \\ 3i/5 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ (1+i)/\sqrt{2} \end{pmatrix}$

form an orthonormal set in $\mathbb{C}^3$.

Orthonormal and unitary system of vectors Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be m vectors in $\mathbb{R}^n$. Then, th set of vectors forms an *orthonormal system* of vectors, if

$$\mathbf{x}_i \cdot \mathbf{x}_j = \mathbf{x}_i^T \mathbf{x}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be m vectors in $\mathbb{C}^n$. Then, this set of vectors forms an *unitary system* of vector if

$$\mathbf{x}_i \cdot \mathbf{x}_j = \mathbf{x}_i^T \,\overline{\mathbf{x}}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

In section 3.2.2, we have defined symmetric, skew-symmetric, Hermitian and skew-Hermitian matrices We now define a few more special matrices.

Orthogonal matrices A real matrix $\mathbf{A}$ is *orthogonal* if $\mathbf{A}^{-1} = \mathbf{A}^T$. A simple example is

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Unitary matrices A complex matrix $\mathbf{A}$ is *unitary* if $\mathbf{A}^{-1} = (\overline{\mathbf{A}})^T$, or $(\overline{\mathbf{A}})^{-1} = \mathbf{A}^T$. If $\mathbf{A}$ is real, then unitary matrix is same as orthogonal matrix. We note the following.

1. If $\mathbf{A}$ and $\mathbf{B}$ are Hermitian matrices, then $\alpha\mathbf{A} + \beta\mathbf{B}$ is also Hermitian for any real scalars α, β, since

$$(\overline{\alpha\mathbf{A} + \beta\mathbf{B}})^T = (\alpha\overline{\mathbf{A}} + \beta\overline{\mathbf{B}})^T = \alpha\overline{\mathbf{A}}^T + \beta\overline{\mathbf{B}}^T = \alpha\mathbf{A} + \beta\mathbf{B}.$$

2. Eigenvalues and eigenvectors of $\overline{\mathbf{A}}$ are the conjugates of the eigenvalues and eigenvectors of $\mathbf{A}$, since

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad \text{gives} \quad \overline{\mathbf{A}}\,\overline{\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}}.$$

3. The inverse of a unitary (orthogonal) matrix is unitary (orthogonal). We have $\mathbf{A}^{-1} = \overline{\mathbf{A}}^T$. Let $\mathbf{B} = \mathbf{A}^{-1}$. Then

$$\mathbf{B}^{-1} = \mathbf{A} = (\overline{\mathbf{A}}^T)^{-1} = [(\overline{\mathbf{A}})^{-1}]^T = [\overline{(\mathbf{A}^{-1})}]^T = \overline{\mathbf{B}}^T.$$

We now establish some important results.

Theorem 3.12 An orthogonal set of vectors is linearly independent.

Proof Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be an orthogonal set of vectors, that is $\mathbf{x}_i \cdot \mathbf{x}_j = 0$, $i \neq j$. Consider the vector equation

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \ldots + \alpha_m \mathbf{x}_m = \mathbf{0} \tag{3.52}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_m$ are scalars. Taking the inner product of the vector $\mathbf{x}$ in Eq. (3.52) with $\mathbf{x}_1$, we get

$$\mathbf{x} \cdot \mathbf{x}_1 = (\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \ldots + \alpha_m \mathbf{x}_m) \cdot \mathbf{x}_1 = \mathbf{0} \cdot \mathbf{x}_1 = 0$$

or

$$\alpha_1(\mathbf{x}_1 \cdot \mathbf{x}_1) = 0 \quad \text{or} \quad \alpha_1 \| \mathbf{x}_1 \|^2 = 0.$$

Since $\| \mathbf{x}_1 \|^2 \neq 0$, we get $\alpha_1 = 0$. Similarly, taking the inner products of $\mathbf{x}$ with $\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_m$ successively, we find that $\alpha_2 = \alpha_3 = \ldots = \alpha_m = 0$. Therefore, the set of orthogonal vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ is linearly independent.

Theorem 3.13 The eigenvalues of

(i) an Hermitian matrix are real.
(ii) a skew-Hermitian matrix are zero or pure imaginary.
(iii) an unitary matrix are of magnitude 1.

Proof Let λ be an eigenvalue and $\mathbf{x}$ be the corresponding eigenvector of the matrix $\mathbf{A}$. We have $\mathbf{Ax} = \lambda \mathbf{x}$. Premultiplying both sides by $\overline{\mathbf{x}}^T$, we get

$$\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x} = \lambda \overline{\mathbf{x}}^T \mathbf{x} \quad \text{or} \quad \lambda = \frac{\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}}{\overline{\mathbf{x}}^T \mathbf{x}}. \tag{3.53}$$

Note that $\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$ and $\overline{\mathbf{x}}^T \mathbf{x}$ are scalars. Also, the denominator $\overline{\mathbf{x}}^T \mathbf{x}$ is always real and positive. Therefore, the behaviour of λ is governed by the scalar $\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$.

(i) Let $\mathbf{A}$ be an Hermitian matrix, that is $\overline{\mathbf{A}} = \mathbf{A}^T$. Now,

$$(\overline{\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}}) = \mathbf{x}^T \overline{\mathbf{A}}\, \overline{\mathbf{x}} = \mathbf{x}^T \mathbf{A}^T \overline{\mathbf{x}} = (\mathbf{x}^T \mathbf{A}^T \overline{\mathbf{x}})^T = \overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$$

since $\overline{\mathbf{x}}^T \mathbf{A}^T \mathbf{x}$ is a scalar. Therefore, $\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$ is real. From Eq. (3.53), we conclude that λ is real.

(ii) Let $\mathbf{A}$ be a skew-Hermitian matrix, that is $\mathbf{A}^T = -\overline{\mathbf{A}}$. Now,

$$(\overline{\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}}) = \mathbf{x}^T \overline{\mathbf{A}}\, \overline{\mathbf{x}} = -\mathbf{x}^T \mathbf{A}^T \overline{\mathbf{x}} = -(\mathbf{x}^T \mathbf{A}^T \overline{\mathbf{x}})^T = -\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$$

since $\mathbf{x}^T \mathbf{A}^T \overline{\mathbf{x}}$ is a scalar. Therefore, $\overline{\mathbf{x}}^T \mathbf{A} \mathbf{x}$ is zero or pure imaginary. From Eq. (3.53), we conclude that λ is zero or pure imaginary.

(iii) Let $\mathbf{A}$ be an unitary matrix, that is $\mathbf{A}^{-1} = (\overline{\mathbf{A}})^T$. Now, from

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x} \quad \text{or} \quad \overline{\mathbf{A}}\, \overline{\mathbf{x}} = \overline{\lambda}\, \overline{\mathbf{x}} \tag{3.54}$$

we get

$$(\overline{\mathbf{A}}\, \overline{\mathbf{x}})^T = (\overline{\lambda}\, \overline{\mathbf{x}})^T \quad \text{or} \quad \overline{\mathbf{x}}^T \overline{\mathbf{A}}^T = \overline{\lambda}\, \overline{\mathbf{x}}^T.$$

or $$\overline{\mathbf{x}}^T\mathbf{A}^{-1} = \overline{\lambda}\,\overline{\mathbf{x}}^T. \tag{3.55}$$

Using Eqs. (3.54) and (3.55), we can write

$$(\overline{\mathbf{x}}^T\mathbf{A}^{-1})(\mathbf{A}\mathbf{x}) = (\overline{\lambda}\,\overline{\mathbf{x}}^T)(\lambda\mathbf{x}) = |\lambda|^2\overline{\mathbf{x}}^T\mathbf{x}$$

or $$\overline{\mathbf{x}}^T\mathbf{x} = |\lambda|^2\overline{\mathbf{x}}^T\mathbf{x}$$

Since $\mathbf{x} \neq \mathbf{0}$, we have $\overline{\mathbf{x}}^T\mathbf{x} \neq 0$. Therefore, $|\lambda|^2 = 1$, or $\lambda = \pm 1$. Hence, the result.

Remark 21

From Theorem 3.13, we conclude that the eigenvalues of

(i) a symmetric matrix are real.

(ii) a skew-symmetric matrix are zero or pure imaginary.

(iii) an orthogonal matrix are of magnitude 1 and are real or complex conjugate pairs.

Theorem 3.14 The column vectors (and also row vectors) of an unitary matrix form an unitary system of vectors.

Proof Let $\mathbf{A}$ be an unitary matrix of order n, with column vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Then

$$\mathbf{A}^{-1}\mathbf{A} = \overline{\mathbf{A}}^T\mathbf{A} = \begin{bmatrix} \overline{\mathbf{x}}_1^T \\ \overline{\mathbf{x}}_2^T \\ \vdots \\ \overline{\mathbf{x}}_n^T \end{bmatrix}[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = \begin{bmatrix} \overline{\mathbf{x}}_1^T\mathbf{x}_1 & \overline{\mathbf{x}}_1^T\mathbf{x}_2 & \ldots & \overline{\mathbf{x}}_1^T\mathbf{x}_n \\ \overline{\mathbf{x}}_2^T\mathbf{x}_1 & \overline{\mathbf{x}}_2^T\mathbf{x}_2 & \ldots & \overline{\mathbf{x}}_2^T\mathbf{x}_n \\ \vdots & & & \\ \overline{\mathbf{x}}_n^T\mathbf{x}_1 & \overline{\mathbf{x}}_n^T\mathbf{x}_2 & \ldots & \overline{\mathbf{x}}_n^T\mathbf{x}_n \end{bmatrix} = \mathbf{I}$$

Therefore, $$\overline{\mathbf{x}}_i^T\mathbf{x}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

Hence, the column vectors of $\mathbf{A}$ form an unitary system. Since the inverse of an unitary matrix is also an unitary matrix and the columns of $\mathbf{A}^{-1}$ are the conjugate of the rows of $\mathbf{A}$, we conclude that the row vectors of $\mathbf{A}$ also form an unitary system.

Remark 22

(a) From Theorem 3.14, we conclude that the column vectors (and also the row vectors) of an orthogonal matrix form an orthonormal system of vectors.

(b) A symmetric matrix of order n has n linearly independent eigenvectors and hence is diagonalizable.

Example 3.47 Show that the matrices $\mathbf{A}$ and $\mathbf{A}^T$ have the same eigenvalues and for distinct eigenvalues the eigenvectors corresponding to $\mathbf{A}$ and $\mathbf{A}^T$ are mutually orthogonal.

Solution We have

$$|\mathbf{A} - \lambda\mathbf{I}| = |(\mathbf{A}^T)^T - \lambda\mathbf{I}^T| = |[\mathbf{A}^T - \lambda\mathbf{I}]^T| = |\mathbf{A}^T - \lambda\mathbf{I}|.$$

Since $\mathbf{A}$ and $\mathbf{A}^T$ have the same characteristic equation, they have the same eigenvalues.

Let λ and μ be two distinct eigenvalues of $\mathbf{A}$. Let $\mathbf{x}$ be the eigenvector corresponding to the eigenvalue λ for $\mathbf{A}$ and $\mathbf{y}$ be the eigenvector corresponding to the eigenvalue μ for $\mathbf{A}^T$. We hav $\mathbf{Ax} = \lambda\mathbf{x}$. Premultiplying by $\mathbf{y}^T$, we get

$$\mathbf{y}^T\mathbf{Ax} = \lambda\,\mathbf{y}^T\mathbf{x}. \tag{3.56}$$

We also have $\quad \mathbf{A}^T\mathbf{y} = \mu\mathbf{y}, \;\text{ or }\; (\mathbf{A}^T\mathbf{y})^T = (\mu\mathbf{y})^T \;\text{ or }\; \mathbf{y}^T\mathbf{A} = \mu\mathbf{y}^T.$

Postmultiplying by $\mathbf{x}$, we get

$$\mathbf{y}^T\mathbf{Ax} = \mu\,\mathbf{y}^T\mathbf{x} \tag{3.57}$$

Subtracting Eqs. (3.56) and (3.57), we obtain

$$(\lambda - \mu)\,\mathbf{y}^T\mathbf{x} = 0.$$

Since $\lambda \neq \mu$, we obtain $\mathbf{y}^T\mathbf{x} = 0$. Therefore, the vectors $\mathbf{x}$ and $\mathbf{y}$ are mutually orthogonal.

3.5.4 Quadratic Forms

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ be an arbitrary vector in $\mathbb{R}^n$. A real *quadratic form* is an homogeneou expression of the form

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{3.58}$$

in which the total power in each term is 2. Expanding, we can write

$$\begin{aligned} Q &= a_{11}x^2 + (a_{12} + a_{21})\, x_1x_2 + \ldots + (a_{1n} + a_{n1})x_1x_n \\ &\quad + a_{22}x_2^2 + (a_{23} + a_{32})x_2x_3 + \ldots + (a_{2n} + a_{n2})x_2x_n \\ &\quad + \ldots + a_{nn}x_n^2 \\ &= \mathbf{x}^{\mathrm{T}}\,\mathbf{A}\,\mathbf{x} \end{aligned} \tag{3.59}$$

using the definition of matrix multiplication. Now, set $b_{ij} = (a_{ij} + a_{ji})/2$. The matrix $\mathbf{B} = (b_{ij})$ is symmetric since $b_{ij} = b_{ji}$. Further, $b_{ij} + b_{ji} = a_{ij} + a_{ji}$. Hence, Eq. (3.59) can be written as

$$Q = \mathbf{x}^T\mathbf{Bx}$$

where $\mathbf{B}$ is a symmetric matrix and $b_{ij} = (a_{ij} + a_{ji})/2$.

For example, for $n = 2$, we have

$$b_{11} = a_{11},\; b_{12} = b_{21} = (a_{12} + a_{21})/2 \quad\text{and}\quad b_{22} = a_{22}.$$

Example 3.48 Obtain the symmetric matrix $\mathbf{B}$ for the quadratic form

(i) $Q = 2x_1^2 + 3x_1x_2 + x_2^2$.

(ii) $Q = x_1^2 + 2x_1x_2 - 4x_1x_3 + 6x_2x_3 - 5x_2^2 + 4x_3^2$.

Solution

(i) $\quad a_{11} = 2,\; a_{12} + a_{21} = 3$ and $a_{22} = 1$. Therefore,

$$b_{11} = a_{11} = 2,\; b_{12} = b_{21} = \frac{1}{2}(a_{12} + a_{21}) = \frac{3}{2} \text{ and } b_{22} = a_{22} = 1.$$

Therefore, $$\mathbf{B} = \begin{bmatrix} 2 & 3/2 \\ 3/2 & 1 \end{bmatrix}.$$

(ii) $a_{11} = 1$, $a_{12} + a_{21} = 2$, $a_{13} + a_{31} = -4$, $a_{23} + a_{32} = 6$, $a_{22} = -5$, $a_{33} = 4$. Therefore,

$b_{11} = a_{11} = 1$, $b_{12} = b_{21} = \frac{1}{2}(a_{12} + a_{21}) = 1$, $b_{13} = b_{31} = \frac{1}{2}(a_{13} + a_{31}) = -2$,

$b_{23} = b_{32} = \frac{1}{2}(a_{23} + a_{32}) = 3$, $b_{22} = a_{22} = -5$, $b_{33} = a_{33} = 4$.

Therefore, $$\mathbf{B} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -5 & 3 \\ -2 & 3 & 4 \end{bmatrix}.$$

If **A** is a complex matrix, then the quadratic form is defined as

$$Q = \sum_{i=1}^{n} \sum_{i=1}^{n} a_{ij}\, \bar{x}_i x_j = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} \tag{3.61}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is an arbitrary vector in $\mathbb{C}^n$. However, this quadratic form is usually define for an Hermitian matrix **A**. Then, it is called a *Hermitian form* and is always real.

For example, consider the Hermitian matrix $\mathbf{A} = \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix}$. The quadratic form becomes

$$Q = \bar{\mathbf{x}}^T \mathbf{A}\mathbf{x} = [\bar{x}_1, \bar{x}_2] \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= |x_1|^2 + (1+i)\bar{x}_1 x_2 + (1-i) x_1 \bar{x}_2 + 2|x_2|^2$$

$$= |x_1|^2 + (\bar{x}_1 x_2 + x_1 \bar{x}_2) + i(\bar{x}_1 x_2 - x_1 \bar{x}_2) + 2|x_2|^2.$$

Now, $\bar{x}_1 x_2 + x_1 \bar{x}_2$ is real and $\bar{x}_1 x_2 - x_1 \bar{x}_2$ is imaginary. For example if $x_1 = p_1 + iq_1$, $x_2 = p_2 + iq_2$ we obtain

$$\bar{x}_1 x_2 + x_1 \bar{x}_2 = 2(p_1 p_2 + q_1 q_2) \text{ and } \bar{x}_1 x_2 - x_1 \bar{x}_2 = 2i(p_1 q_2 - p_2 q_1).$$

We can also write

$$(\bar{x}_1 x_2 + \bar{x}_2 x_1) + i(\bar{x}_1 x_2 - x_1 \bar{x}_2) = 2[(p_1 p_2 + q_1 q_2) - (p_1 q_2 - p_2 q_1)]$$

$$= 2 \operatorname{Re}[(1+i)\bar{x}_1 x_2]$$

Therefore, $Q = |x_1|^2 + 2 \operatorname{Re}[(1+i)\bar{x}_1 x_2] + |x_2|^2$.

Positive definite matrices

Let $\mathbf{A} = (a_{ij})$ be a square matrix. Then, the matrix **A** is said to be *positive definite* if

$$Q = \bar{\mathbf{x}}^T \mathbf{A}\mathbf{x} > 0 \text{ for any vector } \mathbf{x} \neq \mathbf{0} \text{ and } \bar{\mathbf{x}}^T \mathbf{A}\mathbf{x} = 0, \text{ if and only if } \mathbf{x} = \mathbf{0}.$$

If $\mathbf{A}$ is real, then $\mathbf{x}$ can be taken as real.

Positive definite matrices have the following properties.

1. The eigenvalues of a positive definite matrix are all real and positive.

   This is easily proved when $\mathbf{A}$ is a real matrix. From Eq. (3.53), we have

$$\lambda = (\mathbf{x}^T\mathbf{A}\mathbf{x})/(\mathbf{x}^T\mathbf{x})$$

   Since $\mathbf{x}^T\mathbf{x} > 0$ and $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$, we obtain $\lambda > 0$. If $\mathbf{A}$ is Hermitian, then $\overline{\mathbf{x}}^T\mathbf{A}\mathbf{x}$ is real and λ is real (see Theorem 3.13). Therefore, if the Hermitian form $Q > 0$, then the eigenvalues are real and positive.

2. All the leading minors of $\mathbf{A}$ are positive.

Remark 23

(a) If $\mathbf{A}$ is Hermitian and strictly diagonally dominant ($|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|, i = 1, 2, \ldots, n$) with positive real elements on the diagonal, then $\mathbf{A}$ is positive definite.

(b) If $\overline{\mathbf{x}}^T\mathbf{A}\mathbf{x} \geq 0$, then the matrix $\mathbf{A}$ is called *semi-positive definite.*

(c) A matrix $\mathbf{A}$ is called *negative definite* if $(-\mathbf{A})$ is positive definite. All the eigenvalues of a negative definite matrix are real and negative.

Example 3.49 Examine which of the following matrices are positive definite.

(a) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$, (b) $\mathbf{A} = \begin{bmatrix} 3 & -2i \\ 2i & 4 \end{bmatrix}$, (c) $\mathbf{A} = \begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 3 \end{bmatrix}$.

Solution

(a) (i)

$$Q = \mathbf{x}^T\mathbf{A}\mathbf{x} = [x_1, x_2]\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 3x_1x_2 + 4x_2^2$$

$$= 3\left(x_1 + \frac{1}{2}x_2\right)^2 + \frac{13}{4}x_2^2 > 0 \quad \text{for all} \quad \mathbf{x} \neq \mathbf{0}.$$

(ii) eigenvalues of $\mathbf{A}$ are 2 and 5 which are both positive.

(iii) leading minors $|3| = 3$, $\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} = 10$ are both positive.

Hence, the matrix $\mathbf{A}$ is positive definite (it is not necessary to show all the three parts).

(b)

$$Q = \overline{\mathbf{x}}^T\mathbf{A}\mathbf{x} = [\overline{x}_1, \overline{x}_2]\begin{bmatrix} 3 & -2i \\ 2i & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\overline{x}_1, \overline{x}_2]\begin{bmatrix} 3x_1 - 2ix_2 \\ 2ix_1 + 4x_2 \end{bmatrix}$$

$$= 3x_1\overline{x}_1 - 2i\overline{x}_1x_2 + 2ix_1\overline{x}_2 + 4x_2\overline{x}_2.$$

Taking $x_1 = p_1 + iq_1$ and $x_2 = p_2 + iq_2$ and simplifying, we get

$$Q = 3(p_1^2 + q_1^2) + 4(p_2^2 + q_2^2) + 4(p_1q_2 - p_2q_1)$$

$$= p_1^2 + q_1^2 + 2p_2^2 + 2q_2^2 + 2(p_2 - q_1)^2 + 2(p_1 + q_2)^2 > 0.$$

Therefore, the given matrix is positive definite.

Note that **A** is Hermitian, strictly diagonally dominant $(3 > |-2i|, 4 > |2i|)$ with positive rea diagonal entries. Therefore, **A** is positive definite (see Remark 23(a).)

(c) $$Q = \bar{\mathbf{x}}^T\mathbf{A}\mathbf{x} = [\bar{x}_1, \bar{x}_2, \bar{x}_3]\begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [\bar{x}_1, \bar{x}_2, \bar{x}_3]\begin{bmatrix} x_1 + ix_3 \\ x_2 \\ -ix_1 + 3x_3 \end{bmatrix}$$

$$= x_1\bar{x}_1 + i\bar{x}_1x_3 + x_2\bar{x}_2 - ix_1\bar{x}_3 + 3x_3\bar{x}_3$$

$$= |x_1|^2 + |x_2|^2 + 3|x_3|^2 + i(\bar{x}_1x_3 - x_1\bar{x}_3)$$

Taking $x_1 = p_1 + iq_1$, $x_2 = p_2 + iq_2$, $x_3 = p_3 + iq_3$ and simplifying, we obtain

$$Q = (p_1^2 + q_1^2) + (p_2^2 + q_2^2) + 3(p_3^2 + q_3^2) - 2(p_1q_3 - p_3q_1)$$

$$= (p_1 - q_3)^2 + (p_3 + q_1)^2 + (p_2^2 + q_2^2) + 2(p_3^2 + q_3^2) > 0.$$

Therefore, the matrix **A** is positive definite. It can be verified that the eigenvalues of are 1, 2, 2 which are all positive.

Example 3.50 Let **A** be a real square matrix. Show that the matrix $\mathbf{A}^T\mathbf{A}$ has real and positiv eigenvalues.

Solution Since $(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{A}$, the matrix $\mathbf{A}^T\mathbf{A}$ is symmetric. Therefore, the eigenvalues of $\mathbf{A}^T\mathbf{A}$ ar all real. Now,

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) = \mathbf{y}^T\mathbf{y}, \text{ where } \mathbf{A}\mathbf{x} = \mathbf{y}.$$

Since $\mathbf{y}^T\mathbf{y} > 0$ for any vector $\mathbf{y} \neq \mathbf{0}$, the matrix $\mathbf{A}^T\mathbf{A}$ is positive definite and hence all the eigenvalue of $\mathbf{A}^T\mathbf{A}$ are positive. Therefore, all the eigenvalues of $\mathbf{A}^T\mathbf{A}$ are real and positive.

Exercise 3.4

Verify the Cayley-Hamilton theorem for the matrix **A**. Find $\mathbf{A}^{-1}$, if it exists, where **A** is as given in Problems to 6.

1. $\begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$.
2. $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}$
3. $\begin{bmatrix} 1 & -2 & 1 \\ 2 & 3 & -2 \\ 3 & 1 & -1 \end{bmatrix}$.

4. $\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$.
5. $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & -1 \\ -2 & -1 & 1 \end{bmatrix}$.
6. $\begin{bmatrix} 1 & i & i \\ i & 1 & i \\ i & i & 1 \end{bmatrix}$.

Find all the eigenvalues and the corresponding eigenvectors of the matrices given in Problems **7** to **18**. Which of the matrices are diagonalizable?

7. $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$.
8. $\begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$.
9. $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$.

10. $\begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ -1 & 3 & 4 \end{bmatrix}$.

11. $\begin{bmatrix} 1 & 1 & i \\ 1 & 0 & i \\ -i & -i & 1 \end{bmatrix}$.

12. $\begin{bmatrix} 0 & i & i \\ i & 0 & i \\ i & i & 0 \end{bmatrix}$.

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

14. $\begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}$.

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

16. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 2 & 1 \\ 4 & 3 & 1 & 2 \end{bmatrix}$.

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

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

Show that the matrices given in Problems **19** to **24** are diagonalizable. Find the matrix **P** such that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix.

19. $\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}$.

20. $\begin{bmatrix} -3 & -2 & 1 \\ -2 & 0 & 4 \\ -6 & -3 & 5 \end{bmatrix}$.

21. $\begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & -3 & 0 \end{bmatrix}$.

22. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$.

23. $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$.

24. $\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$.

Find the matrix **A** whose eigenvalues and the corresponding eigenvectors are as given in Problems **25** to **30**.

25. Eigenvalues: 2, 2, 4; Eigenvectors: $(-2, 1, 0)^T$, $(-1, 0, 1)^T$, $(1, 0, 1)^T$.
26. Eigenvalues: 1, – 1, 2; Eigenvectors: $(1, 1, 0)^T$, $(1, 0, 1)^T$, $(3, 1, 1)^T$.
27. Eigenvalues: 1, 2, 3; Eigenvectors: $(1, 2, 1)^T$, $(2, 3, 4)^T$, $(1, 4, 9)^T$.
28. Eigenvalues: 1, 1, 1; Eigenvectors: $(-1, 1, 1)^T$, $(1, -1, 1)^T$, $(1, 1, -1)^T$.
29. Eigenvalues: 0, – 1, 1; Eigenvectors: $(-1, 1, 0)^T$, $(1, 0, -1)^T$, $(1, 1, 1)^T$.
30. Eigenvalues: 0, 0, 3; Eigenvectors: $(1, 2, -1)^T$, $(-2, 1, 0)^T$, $(3, 0, 1)^T$.
31. Let a 4×4 matrix **A** have eigenvalues 1, – 1, 2, – 2. Find the value of the determinant of the matrix $\mathbf{B} = 2\mathbf{A} + \mathbf{A}^{-1} - \mathbf{I}$.
32. Let a 3×3 matrix **A** have eigenvalues 1, 2, – 1. Find the trace of the matrix $\mathbf{B} = \mathbf{A} - \mathbf{A}^{-1} + \mathbf{A}^2$.
33. Show that the matrices **A** and $\mathbf{P}^{-1}\mathbf{AP}$ have the same eigenvalues.
34. Let **A** and **B** be square matrices of the same order. Then, show that **AB** and **BA** have the same eigenvalues but different eigenvectors.
35. Show that the matrices $\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{BA}^{-1}$ have the same eigenvalues but different eigenvectors.
36. An $n \times n$ matrix **A** is *nilpotent* if for some positive integer k, $\mathbf{A}^k = \mathbf{0}$. Show that all the eigenvalues of a nilpotent matrix are zero.
37. If **A** is an $n \times n$ diagonalizable matrix and $\mathbf{A}^2 = \mathbf{A}$, then show that each eigenvalue of **A** is 0 or 1.

38. Show that the matrix $\mathbf{A} = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$, $a \neq b$, is transformed to a diagonal matrix $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, where $\mathbf{P}$ is of the form $\mathbf{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and $\tan 2\theta = \dfrac{2h}{a-b}$.

39. Let $\mathbf{A}$ be similar to $\mathbf{B}$. Then show that (i) $\mathbf{A}^{-1}$ is similar to $\mathbf{B}^{-1}$, (ii) $\mathbf{A}^m$ is similar to $\mathbf{B}^m$ for any positive integer m, (iii) $|\mathbf{A}| = |\mathbf{B}|$.

40. Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric matrices of the same order. Then, show that $\mathbf{AB}$ is symmetric if and only if $\mathbf{AB} = \mathbf{BA}$.

41. For any square matrix $\mathbf{A}$, show that $\mathbf{A}^T\mathbf{A}$ is symmetric.

42. Let $\mathbf{A}$ be a non-singular matrix. Show that $\mathbf{A}^T\mathbf{A}^{-1}$ is symmetric if and only if $\mathbf{A}^2 = (\mathbf{A}^T)^2$.

43. If $\mathbf{A}$ is a symmetric matrix and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, then show that $\mathbf{P}$ is an orthogonal matrix.

44. Show that the product of two orthogonal matrices of the same order is also an orthogonal matrix.

45. Find the conditions that a matrix $\mathbf{A} = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix}$ is orthogonal.

46. If $\mathbf{A}$ is an orthogonal matrix, show that $|\mathbf{A}| = \pm 1$.

47. Prove that the eigenvectors of a symmetric matrix corresponding to distinct eigenvalues are orthogonal.

48. A matrix $\mathbf{A}$ is called a *normal matrix* if $\mathbf{A}\overline{\mathbf{A}}^T = \overline{\mathbf{A}}^T\mathbf{A}$. Show that the Hermitian, skew-Hermitian and unitary matrices are normal.

49. If a matrix $\mathbf{A}$ can be diagonalized using an orthogonal matrix then show that $\mathbf{A}$ is symmetric.

50. Suppose that a matrix $\mathbf{A}$ is both unitary and Hermitian. Then, show that $\mathbf{A} = \mathbf{A}^{-1}$.

51. If $\mathbf{A}$ is a symmetric matrix and $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for every real vector $\mathbf{x} \neq \mathbf{0}$, then show that $\overline{\mathbf{z}}^T\mathbf{A}\mathbf{z}$ is real and positive for any complex vector $\mathbf{z} \neq \mathbf{0}$.

52. Show that an unitary transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is an unitary matrix preserves the value of the inner product.

53. Prove that a real 2×2 symmetric matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite if and only if $a > 0$ (1×1 leading minor) and $ac - b^2 > 0$ (2×2 leading minor).

54. Show that the matrix $\begin{bmatrix} 2 & 1 & 3 \\ -3 & 4 & -1 \\ -1 & 1 & 2 \end{bmatrix}$ is positive definite.

55. Show that the matrix $\begin{bmatrix} -3 & -2 & 1 \\ -2 & 0 & 4 \\ -6 & -3 & 5 \end{bmatrix}$ is not positive definite.

Find the symmetric or the Hermitian matrix $\mathbf{A}$ for the quadratic forms given in Problems **56** to **60**.

56. $x_1^2 - 2x_1x_2 + 4x_2x_3 - x_2^2 + x_3^2$.

57. $3x_1^2 + 2x_1x_2 - 4x_1x_3 + 8x_2x_3 + x_2^2$.

58. $x_1^2 + 2ix_1x_2 - 8x_1x_3 + 4ix_2x_3 + 4x_3^2$.

59. $x_1^2 - (2+4i)x_1x_2 - (4-6i)x_2x_3 + x_2^2$.

60. $2x_1^2 - 3x_2^2 + (6+8i)x_1x_2 + (4-2i)x_2x_3$.

3.6 Answers and Hints

Exercise 3.1

3. $\mathbf{A}^{-1} = \frac{1}{5}\begin{bmatrix} -3 & 2 & 2 \\ 2 & -3 & 2 \\ 2 & 2 & -3 \end{bmatrix}$.

4. $\mathbf{A}^{-1} = \frac{1}{11}\begin{bmatrix} -3 & 4 & 5 \\ 9 & -1 & -4 \\ 5 & -3 & -1 \end{bmatrix}$.

8. (i) $|\mathbf{A}\ adj\ (\mathbf{A})| = \text{diag}\ (|\mathbf{A}|, |\mathbf{A}|, \ldots, |\mathbf{A}|) = |\mathbf{A}|^n$ (use property 10 of determinants). Therefore, $|Adj\ (\mathbf{A})| = |\mathbf{A}|^{n-1}$.

   (ii) Let $\mathbf{B} = adj\ (\mathbf{A})$. Since $\mathbf{B}^{-1} = adj\ (\mathbf{B})/|\mathbf{B}|$, we have $\mathbf{B}\ adj\ (\mathbf{B}) = |\mathbf{B}|\ \mathbf{I}$. Therefore,
$$adj\ (\mathbf{A})\ adj\ (adj\ (\mathbf{A})) = |\ adj\ (\mathbf{A})\ |\mathbf{I} = |\mathbf{A}|^{n-1}\mathbf{I}.$$
   Premultiplying by $\mathbf{A}$ and using $adj\ (\mathbf{A}) = \mathbf{A}^{-1}|\mathbf{A}|$, we get
$$\mathbf{A}[\mathbf{A}^{-1}|\mathbf{A}|]\ adj\ (adj\ (\mathbf{A})) = |\mathbf{A}|^{n-1}\mathbf{AI} \quad \text{or} \quad adj\ (adj\ (\mathbf{A})) = |\mathbf{A}|^{n-2}\mathbf{A}.$$

9. $|\mathbf{AA}^{-1}| = |\mathbf{A}|\,|\mathbf{A}^{-1}| = |\mathbf{I}|$ or $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$.

10. $(\mathbf{BAB}^T)^T = \mathbf{BA}^T\mathbf{B}^T = \mathbf{BAB}^T$.

13. $\mathbf{AB} = \mathbf{BA} \Rightarrow \mathbf{B}^{-1}\mathbf{AB} = \mathbf{A} \Rightarrow \mathbf{B}^{-1}\mathbf{A} = \mathbf{AB}^{-1}$. Similarly, $\mathbf{A}^{-1}\mathbf{B} = \mathbf{BA}^{-1}$.

    (i) $(\mathbf{AB}^{-1})^T = (\mathbf{B}^{-1})^T\mathbf{A}^T = (\mathbf{B}^T)^{-1}\mathbf{A}^T = \mathbf{B}^{-1}\mathbf{A} = \mathbf{AB}^{-1}$.

    (ii) $(\mathbf{A}^{-1}\mathbf{B})^T = \mathbf{B}^T(\mathbf{A}^{-1})^T = \mathbf{B}^T(\mathbf{A}^T)^{-1} = \mathbf{BA}^{-1} = \mathbf{A}^{-1}\mathbf{B}$.

    (iii) $(\mathbf{A}^{-1}\mathbf{B}^{-1})^T = [(\mathbf{BA})^{-1}]^T = [(\mathbf{AB})^{-1}]^T = (\mathbf{A}^T)^{-1}(\mathbf{B}^T)^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$.

14. Premultiply both sides by (i) $\mathbf{I} - \mathbf{A}$, (ii) $\mathbf{I} + \mathbf{A}$.

15. $(\mathbf{PAQ})^{-1} = \mathbf{Q}^{-1}\mathbf{A}^{-1}\mathbf{P}^{-1} = \mathbf{I} \Rightarrow \mathbf{A}^{-1}\mathbf{P}^{-1} = \mathbf{Q} \Rightarrow \mathbf{A}^{-1} = \mathbf{QP}$.

16. Use $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots) = \mathbf{I}$.

17. $(\mathbf{ABC})(\mathbf{ABC})^{-1} = \mathbf{I}$. Premultiply successively by $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$ and $\mathbf{C}^{-1}$.

21. Multiply C_1 by a, C_2 by b, C_3 by c and take out a from R_1, b from R_2, c from R_3.

27. $\begin{vmatrix} \sin\alpha & \cos\alpha & 0 \\ \sin\beta & \cos\beta & 0 \\ \sin\gamma & \cos\gamma & 0 \end{vmatrix}\begin{vmatrix} \cos a & \cos b & \sin c \\ \sin a & \sin b & \sin c \\ 0 & 0 & 0 \end{vmatrix} = 0$

28. 1, 2, 3.

29. 1, 1, 1.

30. 1, 1, 1.

31. 1, 2, 1.

32. (i) $k \neq 2$ and $k \neq -3$, (ii) $k = 2$, or $k = -3$.

33. $\theta = \pi/6$, or $\theta = \sin^{-1}[(9 - \sqrt{161})/4]$.

35. (i) $\lambda \neq 3$, μ arbitrary, (ii) $\lambda = 3$, $\mu = 10$, (iii) $\lambda = 3$, $\mu \neq 10$.

36. 2.

37. 1.

38. 2.

39. $|\mathbf{A}| = (p-q)(q-r)(r-p)(p+q+r)$; rank $(\mathbf{A})$ is

    (i) 3, if $p \neq q \neq r$ and $p + q + r \neq 0$;

(ii) 2, if $p \neq q \neq r$ and $p + q + r = 0$,

(iii) 2, if exactly two of p, q and r are identical;

(iv) 1, if $p = q = r$.

40. (a) 2; (b) $|\mathbf{A}| = (a_1a_2 + b_1b_2 + c_1c_2)^2$, rank ($\mathbf{A}$) is

(i) 4, if $a_1a_2 + b_1b_2 + c_1c_2 \neq 0$;

(ii) 2, if $a_1a_2 + b_1b_2 + c_1c_2 = 0$, since all determinants of third order have the value zero.

42. Conisder $(\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A} + \mathbf{A}^2 - \ldots + (-1)^{n-1}\mathbf{A}^{n-1}) = \mathbf{I} + (-1)^{n-1}\mathbf{A}^n$. In the limit $n \to \infty$, $\mathbf{A}^n \to 0$. Therefore,

$$(\mathbf{I} + \mathbf{A})(\mathbf{I} - \mathbf{A} + \mathbf{A}^2 - \ldots) = \mathbf{I}.$$

43. (i) $Trace\ (\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha \sum_{i=1}^{n} a_{ii} + \beta \sum_{i=1}^{n} b_{ii} = \alpha\ Trace\ (\mathbf{A}) + \beta\ Trace(\mathbf{B})$,

(ii) $Trace\ (\mathbf{AB}) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{i=1}^{n} \sum_{m=1}^{n} b_{im} a_{mi} = Trace\ (\mathbf{BA})$,

(iii) If the result is true, then $Trace\ (\mathbf{AB} - \mathbf{BA}) = Trace\ (\mathbf{I})$ which gives $0 = n$ which is not possible.

44. Result is true for $p = 0$ and 1. Let it be true for $p = k$ and show that it is true for $p = k + 1$. Note that when $\mathbf{BC} = \mathbf{CB}$ and $\mathbf{C}^2 = \mathbf{0}$, we have $\mathbf{CB}^{k+1} = \mathbf{B}^{k+1}\mathbf{C}$ and $\mathbf{CB}^k\mathbf{C} = \mathbf{0}$.

45. Apply the operation $C_1 \leftarrow C_1 + C_2 + \ldots + C_n$ and then the operation $R_i \leftarrow R_i - R_1$, $i = 2, 3, \ldots, n$.

46. None.
47. Symmetric.
48. Skew-symmetric.
49. Hermitian.
50. None
51. Skew-Hermitian.
52. None.
53. Skew-Hermitian.
54. Hermitian.
55. None.

Exercise 3.2

1. Yes.
2. No, 1, 4, 5, 6.
3. No, 1, 4, 5, 6.
4. No, when the scalar α is irrational, property 6 is not satisfied. If the field of scalars is taken only as rationals, then it defines a vector space.
5. Yes, since $\mathbf{1} + \mathbf{x} = 1x = x = \mathbf{x}$ and $\mathbf{x} + \mathbf{1} = x1 = x = \mathbf{x}$, the zero vector $\mathbf{0}$ is $\mathbf{1} = 1$. Define $-\mathbf{x} = 1/x$. Then, $\mathbf{x} + (-\mathbf{x}) = x(1/x) = 1 = \mathbf{1} = \mathbf{0}$. Therefore, negative vector is its reciprocal.
6. No, 8, 10.
7. No, 2, 3, 8, 10.
8. Yes (same arguments as in Problem 5.). $(\alpha + \beta)\, x = x^{\alpha+\beta} = x^{\alpha}x^{\beta} = x^{\alpha} + x^{\beta} = \alpha x + \beta x$.
9. (i) Yes, (ii) No, 1, 6.
10. (i) Yes, (ii) No, 1, 4, 6.
11. (i) Yes, (ii) No, when $\mathbf{x}, \mathbf{y} \in W$, $\mathbf{x} + \mathbf{y} \notin W$,
(iii) No, when $\mathbf{x}, \mathbf{y} \in W$, $\mathbf{x} + \mathbf{y} \notin W$, (iv) Yes.
12. (i) No, when $\mathbf{A} \in W$, $\alpha\mathbf{A} \notin W$ for α negative,
(ii) No, sum of two non-singular matrices need not be non-singular,
(iii) Yes,
(iv) No, $\alpha\mathbf{A}$ and $\mathbf{A} + \mathbf{B}$ need not belong to W, ($\mathbf{A} = \mathbf{I}$, $\mathbf{A}^2 = \mathbf{I} = \mathbf{A}$ but $2\mathbf{A} \neq (2\mathbf{A})^2$).
13. (i) Yes, (ii) No; let $\alpha = i$. Then $\alpha A = iA \notin W$.
14. (i) No; for $P, Q \in W$, $P + Q \notin W$, (ii) Yes.

(iii) No; for $P, Q \in W$, $\alpha P \notin W$ and also $P + Q \notin W$,

(iv) No, for $P, Q \notin W$ having real roots, $P + Q$ need not have real roots. For example, take $P = 2t^2 - 1$, $Q = -t^2 + 3$.

15. (i) Yes,

(ii) No, $\mathbf{x}, \mathbf{y} \in W$, $\mathbf{x} + \mathbf{y} \notin W$. For example, if $\mathbf{x} = (x_1, x_1, x_1 - 1)$, $\mathbf{y} = (y_1, y_1, y_1 - 1)$; $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_1 + y_1, x_1 + y_1 - 2) \notin W$,

(iii) No, $\mathbf{x} \in W$, $\alpha\mathbf{x} \notin W$, for α negative,

(iv) No, $\mathbf{x} \in W$, $\alpha\mathbf{x} \notin W$, (v) No, $\mathbf{x} \in W$, $\alpha\mathbf{x} \notin W$, (for α a rational number).

16. (i) $\mathbf{u} + 2\mathbf{v} - \mathbf{w}$, (ii) $2\mathbf{u} + \mathbf{v} - \mathbf{w}$,

(iii) $(-33\,\mathbf{u} - 11\,\mathbf{v} + 23\,\mathbf{w})/16$.

17. (i) $\mathbf{u} - 2\mathbf{v} + 2\mathbf{w}$, (ii) $3\mathbf{u} + \mathbf{v} - \mathbf{w}$, (iii) not possible.

18. (i) $3P_1(t) - 2P_2(t) - P_3(t)$, (ii) $4P_1(t) - P_2(t) + 3P_3(t)$.

19. Let $S = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$. Then, $\mathbf{x} = (a, b, c)^T = \alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w}$, where $\alpha = (a + b)/2$, $\beta = (a - b)/2$ and $\gamma = c$.

20. Let $S = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$. Then, $\mathbf{E} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha\mathbf{A} + \beta\mathbf{B} + \gamma\mathbf{C} + \delta\mathbf{D}$, where $\alpha = (-a - b + 2c - 2d)/3$,

$\beta = (5a + 2b - 4c + 4d)/3$, $\gamma = (-4a - b + 5c - 2d)/3$ and $\delta = (-2a + b + c - d)/3$.

21. (i) independent, (ii) dependent, (iii) dependent,

(iv) independent, (v) dependent.

22. (i) independent, (ii) dependent, (iii) dependent,

(iv) independent, (v) independent.

24. $(-4, 7, 9) = (1, 2, 3) + 2(-1, 3, 4) - (3, 1, 2)$. The vectors in S are linearly dependent.

25. $t^2 + t + 1 = [-t + (t^2 - 1) + 2(t^2 + 2t + 2)]/3$. The elements in S are linearly independent.

26. (i) dimension: 2, a basis : $\{(1, 0, 0, -1), (0, 1, -1, 0)\}$,

(ii) dimension: 3, a basis: $\{(1, 0, 0, 1), (0, 1, 0, 1), (0, 0, 1, 1)\}$,

(iii) dimension: 3, a basis: $\{(1, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$,

(iv) dimension: 1, a basis: $\{(1, 1, 1, 1)\}$.

27. The given vectors must be linearly independent.

(i) $k \neq 0, 1, -4/3$, (ii) $k \neq 0$, (iii) $k \neq 0$, (iv) $k \neq -8$.

28. (i) dimension: 4, basis: $\{\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}\}$ where $\mathbf{E}_{rs}$ is the standard basis of order 2,

(ii) dimension: 3, basis: $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$,

(iii) dimension: 1, basis: $\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\}$,

(iv) a 2×2 skew-Hermitian matrix (diagonal elements are 0 or pure imaginary) is given by

$$\mathbf{A} = \begin{pmatrix} ia_1 & b_1 + ib_2 \\ -b_1 + ib_2 & ia_2 \end{pmatrix} = \begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix} + i\begin{pmatrix} a_1 & b_2 \\ b_2 & a_2 \end{pmatrix} = \mathbf{B} + i\,\mathbf{C}$$

where $\mathbf{B}$ is a skew-symmetric and $\mathbf{C}$ is a symmetric matrix, dimension: 4, basis: $\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$,

(v) dimension: 3, basis: $\left\{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\}$.

(vi) dimension: 3, basis: $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$.

29. (i) dimension: 3, basis: $\{\mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{33}\}$,

(ii) dimension: 6, basis: $\{\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{13}, \mathbf{E}_{22}, \mathbf{E}_{23}, \mathbf{E}_{33}\}$,

(iii) dimension: 6, basis: $\{\mathbf{E}_{11}, \mathbf{E}_{21}, \mathbf{E}_{22}, \mathbf{E}_{31}, \mathbf{E}_{32}, \mathbf{E}_{33}\}$.

where $\mathbf{E}_{rs}$ is the standard basis of order 3.

30. (i) n^2, (ii) n, (iii) $n(n + 1)/2$, (iv) $n(n - 1)/2$.

31. Not linear, $T(\mathbf{x}) + T(\mathbf{y}) \neq T(\mathbf{x} + \mathbf{y})$.

32. Linear.

33. Not linear, $T(\mathbf{x}) + T(\mathbf{y}) \neq T(\mathbf{x} + \mathbf{y})$.

34. Not linear, $T(1, 0) = 3$, $T(0, 1) = 2$, $T(1,1) = 0 \neq T(1, 0) + T(0, 1)$.

35. Not linear, $T(\mathbf{x}) + T(\mathbf{y}) \neq T(\mathbf{x} + \mathbf{y})$.

36. $ker(T) = (0, 0, 0)^T$, $ran(T) = x(1, 0, 1)^T + y(1, 0, -1)^T + z(0, 1, 0)^T$.
$dim\,(ker(T)) = 0$, $dim\,(ran(T)) = 3$.

37. $ker(T) = (0, 0)^T$, $ran(T) = x(2, -1, 3)^T + y(1, 1, 4)^T$. $dim\,(ker(T)) = 0$, $dim\,(ran(T)) = 2$.

38. $ker(T) = w(1, -2, 0, 1)^T$,

$$ran(T) = x(1, 0, 0)^T + y(1, 0, 1)^T + z(0, 1, 0)^T + w(1, 0, 2)^T$$
$$= r(1, 0, 0)^T + s(1, 0, 1)^T + z(0, 1, 0),$$

where $r = x - w$, $s = y + 2w$. $dim\,(ker(T)) = 1$, $dim\,(ran(T)) = 3$.

39. $ker(T) = x(-3, 1)^T$, $ran(T) =$ real number. $dim\,(ker(T)) = 1$, $dim\,(ran(T)) = 1$.

40. $ker(T) = x(1, -3, 0)^T + z(0, 0, 1)^T$, $ran(T) =$ real number. $dim\,(ker(T)) = 2$, $dim\,(ran(T)) = 1$.

41. $ker\,(T) = x(1, 1)^T$, $ran(T) = x(1, 1)^T - y(1, 1)^T = r(1, 1)^T$, where $r = x - y$.
$dim\,(ker(T)) = 1$, $dim\,(ran(T)) = 1$.

42. $ker(T) = x(1, 2, -3)^T$, $ran(T) = x(2, 3)^T + y(-1, 0)^T + z(0, 1)^T$ or $ran(T) = r(-1, 0)^T + s(0, 1)^T$, where $r = y + 2x$, $s = z + 3x$. $dim\,(ker(T)) = 1$, $dim\,(ran(T)) = 2$.

43. (i) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, (ii) $\mathbf{A} = \begin{bmatrix} -5 & -8 & -7 \\ 3 & 5 & 4 \end{bmatrix}$.

44. (i) $\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 2 & 2 \end{bmatrix}$, (ii) $\mathbf{A} = \begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & -1/2 \\ 1 & 1 & 1/2 \end{bmatrix}$.

45. $\mathbf{A} = \begin{bmatrix} -1/2 & -1/2 & -3/2 \\ -1/2 & -3/2 & -1/2 \\ 0 & -1 & -1 \end{bmatrix}$.

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

47. We have $T[\mathbf{v}_1, \mathbf{v}_2] = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]\,\mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 2 & 3 \\ 0 & 1 \end{bmatrix}$.

Now, any vector $\mathbf{x} = (x_1, x_2)^T$ in $\mathbb{R}^2$ with respect to the given basis can be written as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

We obtain $\alpha = (-4x_1 + 3x_2)/2$, $\beta = (2x_1 - x_2)/2$. Hence, we have

$$T\mathbf{x} = \alpha T\mathbf{v}_1 + \beta T\mathbf{v}_2 = \alpha \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4\alpha + 5\beta \\ 2\alpha + 3\beta \\ \beta \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -6x_1 + 7x_2 \\ -2x_1 + 3x_2 \\ 2x_1 - x_2 \end{bmatrix}.$$

48. $$T\mathbf{x} = \begin{bmatrix} -x_1 + 2x_2 + 8x_3 \\ -2x_1 + 3x_2 + 12x_3 \end{bmatrix}$$

49. $$T\,P_1(t) = (4x_2 - 5x_1) + 7\,(x_2 - x_1)\,t + (2x_1 - x_2)\,t^2.$$

50. (i) Two degrees of freedom, dimension is 2, a basis is {[3, 1, 0], [– 2, 0, 1]}.
(ii) One degree of freedom, dimension is 1, a basis is {(– 5, 4, 23)}.

Exercise 3.3

1. 3. **2.** 2. **3.** 3. **4.** 2.
5. 2. **6.** 2. **7.** 2. **8.** 3.
9. 4. **10.** 2. **11.** 2. **12.** 3.
13. 2. **14.** 2. **15.** 2.
16. Independent, 3. **17.** Independent, 3. **18.** Dependent, 3.
19. Independent, 3. **20.** Dependent, 2. **21.** Dependent, 3.
22. Dependent, 2. **23.** Dependent, 2. **24.** Independent, 4.
25. Dependent, 2. **26.** [1, 2, 2].

27. $[1+\alpha, -2\alpha, \alpha]$, α arbitrary.

28. Inconsistent.

29. $[1, 1, 1]$.

30. $[1, 3, 3]$.

31. $[3/2, 3/2, 1]$.

32. $[-1, -1/2, 3/4]$.

33. $[(5+\alpha-4\beta)/3, (1+2\alpha+\beta)/3, \alpha, \beta]$, α, β arbitrary.

34. $[2-\alpha, 1, \alpha, 1]$, α arbitrary.

35. $[-1/4, 1/4, 1/4, 1/4]$.

36. $[-\alpha, \alpha, \alpha]$, α arbitrary.

37. $[-15\alpha, 13\alpha, \alpha]$, α arbitrary.

38. $[0, 0, 0]$.

39. $[-2\alpha/3, 7\alpha/3, -8\alpha/3, \alpha]$, α arbitrary.

40. $[2(\beta-\alpha)/3, -(5\beta+\alpha)/3, \beta, \alpha]$, α, β arbitrary.

41. $[0, 0, 0, 0]$.

42. $[(2\beta-5\alpha)/4, -(10\beta+\alpha)/4, \beta, \alpha]$, α, β arbitrary.

43. $[(\alpha+5\beta)/3, (4\beta-7\alpha)/3, \beta, \alpha]$, α, β arbitrary.

44. $[(3\beta-5\alpha)/3, (3\beta-4\alpha)/3, \beta, \alpha]$, α, β arbitrary.

45. $[\alpha-3\beta, 5\beta, \beta, \alpha]$, α, β arbitrary.

46. $\begin{bmatrix} 1 & 1 & -1 \\ 1 & -3 & 2 \\ -1 & 2 & -1 \end{bmatrix}$.

47. $\begin{bmatrix} 3 & -5 & 6 \\ -2 & 4 & -5 \\ 1 & -2 & 3 \end{bmatrix}$.

48. $\frac{1}{4}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

49. $\frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$.

50. $\begin{bmatrix} -1 & -1/3 & 1/3 & 1 \\ 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & 1/3 & -1/3 & 0 \end{bmatrix}$.

Exercise 3.4

1. $P(\lambda) = \lambda^3 - 9\lambda^2 - 9\lambda + 81 = 0$; $\mathbf{A}^{-1} = \frac{1}{81}\begin{bmatrix} 1 & 16 & -20 \\ 16 & 13 & 4 \\ -20 & 4 & -5 \end{bmatrix}$.

2. $P(\lambda) = \lambda^3 - 8\lambda^2 + 20\lambda - 16 = 0$; $\mathbf{A}^{-1} = \frac{1}{16}\begin{bmatrix} 6 & -4 & -2 \\ 0 & 8 & 0 \\ -2 & -4 & 6 \end{bmatrix}$.

3. $P(\lambda) = \lambda^3 - 3\lambda^2 + 2\lambda = 0$; Inverse does not exist.

4. $P(\lambda) = \lambda^3 - \lambda^2 - 4\lambda + 4 = 0$; $\mathbf{A}^{-1} = \frac{1}{4}\begin{bmatrix} 4 & 4 & -4 \\ -2 & -1 & 3 \\ -2 & 1 & 1 \end{bmatrix}$.

5. $P(\lambda) = \lambda^3 - 5\lambda^2 + 9\lambda - 13 = 0$; $\mathbf{A}^{-1} = \frac{1}{13}\begin{bmatrix} 2 & -3 & -7 \\ 1 & 5 & 3 \\ 5 & -1 & 2 \end{bmatrix}$.

6. $P(\lambda) = \lambda^3 - 3\lambda^2 + 6\lambda - 4 + 2i = 0$; $\mathbf{A}^{-1} = -\frac{1+3i}{10}\begin{bmatrix} i-1 & 1 & 1 \\ 1 & i-1 & 1 \\ 1 & 1 & i-1 \end{bmatrix}$.

7. $\lambda = 1$: $(1, 1, -1)^T$; $\lambda = 2, 2$: $(2, 1, 0)^T$; not diagonalizable.
8. $\lambda = -1$: $(0, -1, 1)^T$; $\lambda = i$: $(1 + i, 1, 1)^T$; $\lambda = -i$: $[1 - i, 1, 1]^T$, diagonalizable.
9. $\lambda = 1, 1, 1$: $[0, 3, -2]^T$; not diagonalizable.
10. $\lambda = 1, 1$: $[0, 1, -1]^T$; $\lambda = 7$: $(6, 7, 5)^T$; not diagonalizable.
11. $\lambda = 0$: $[i, 0, -1]^T$; $\lambda = 1 + \sqrt{3}$: $[1, \sqrt{3} - 1, -i]^T$;
$\lambda = 1 - \sqrt{3}$: $[1, -(\sqrt{3} + 1), -i]^T$; diagonalizable.
12. $\lambda = -i, -i$: $[1, 0, -1]^T$, $[1, -1, 0]^T$; $\lambda = 2i$: $[1, 1, 1]^T$; diagonalizable.
13. $\lambda = 0, 0, 0, 0$: $[1, 0, 0, 0]^T$; not diagonalizable.
14. $\lambda = 0, 0$: $[1, 0, 0, 1]^T$, $[1, -1, -1, 0]^T$; $\lambda = 2$: $[1, 1, 0, 0]^T$;
$\lambda = -2$: $[1, 0, 1, 1]^T$; diagonalizable.
15. $\lambda = -1, -1, -1$: $[1, -1, 0, 0]^T$, $[1, 0, -1, 0]^T$, $[1, 0, 0, -1]^T$;
$\lambda = 3$: $[1, 1, 1, 1]^T$; diagonalizable.
16. $\lambda = -4$: $[1, 1, -1, -1]^T$; $\lambda = 10$: $[1, 1, 1, 1]^T$; $\lambda = \sqrt{2}$: $[\sqrt{2} - 1, 1 - \sqrt{2}, -1, 1]^T$,
$\lambda = -\sqrt{2}$: $[-(1 + \sqrt{2}), 1 + \sqrt{2}, -1, 1]^T$; diagonalizable.
17. $\lambda = -1, -1$: $[1, 0, 0, 0, -1]^T$, $[0, 1, 0, -1, 0]^T$; $\lambda = 1, 1, 1$: $[1, 0, 0, 0, 1]^T$, $[0, 1, 0, 1, 0]^T$, $[0, 0, 1, 0, 0]^T$; diagonalizable.
18. $\lambda = 1, w, w^2, w^3, w^4$, w is fifth root of unity. Let $\xi_j = w^j$, $j = 0, 1, 2, 3, 4$. $\lambda = \xi_j$: $[1, \xi_j, \xi_j^2, \xi_j^3, \xi_j^4]^T$, $j = 0, 1, 2, 3, 4$; diagonalizable.
19. $\lambda = 2, 2$: $[1, 0, -1]^T$, $[-2, 1, 0]^T$; $\lambda = 4$: $[1, 0, 1]^T$.

$$\mathbf{P} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 2 & -1 \\ 0 & 2 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

20. $\lambda = 1$: $[1, -2, 0]^T$; $\lambda = -1$: $[3, -2, 2]^T$; $\lambda = 2$: $[-1, 3, 1]^T$.

$$\mathbf{P} = \begin{bmatrix} 1 & 3 & -1 \\ -2 & -2 & 3 \\ 0 & 2 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} -8 & -5 & 7 \\ 2 & 1 & -1 \\ -4 & -2 & 4 \end{bmatrix}.$$

21. $\lambda = 0$: $[3, 1, -2]^T$; $\lambda = 2i$: $[3 + i, 1 + 3i, -4]^T$; $\lambda = -2i$: $[3 - i, 1 - 3i, -4]^T$.

$$\mathbf{P} = \begin{bmatrix} 3 & 3+i & 3-i \\ 1 & 1+3i & 1-3i \\ -2 & -4 & -4 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{32}\begin{bmatrix} 24 & -8 & 16 \\ 2i-6 & 2-6i & -8 \\ -2i-6 & 2+6i & -8 \end{bmatrix}.$$

22. $\lambda = 0$: $[1, 0, -1]^T$; $\lambda = 1$: $[-1, -1, 1]^T$, $\lambda = 2$: $[1, 1, 0]^T$.

$$\mathbf{P} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix}; \mathbf{P}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

23. $\lambda = 1$: $[3, -1, 3]^T$; $\lambda = 2, 2$: $[2, 0, 1]^T$, $[2, 1, 0]^T$.

$$\mathbf{P} = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}; \mathbf{P}^{-1} = \begin{bmatrix} -1 & 2 & 2 \\ 3 & -6 & -5 \\ -1 & 3 & 2 \end{bmatrix}.$$

24. $\lambda = 1: [1, -1, -1]^T$; $\lambda = 2: [0, 1, 1]^T$, $\lambda = -2: [8 - 5, 7]^T$.

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 8 \\ -1 & 1 & -5 \\ -1 & 1 & 7 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{12}\begin{bmatrix} 12 & 8 & -8 \\ 12 & 15 & -3 \\ 0 & -1 & 1 \end{bmatrix}.$$

25. $\mathbf{P} = \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 0 & 2 & 0 \\ -1 & -2 & 1 \\ 1 & 2 & 1 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}.$

26. $\mathbf{P} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \begin{bmatrix} -1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \begin{bmatrix} 6 & -5 & -7 \\ 1 & 0 & -1 \\ 3 & -3 & -4 \end{bmatrix}.$

27. $\mathbf{P} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 4 & 9 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{12}\begin{bmatrix} -11 & 14 & -5 \\ 14 & -8 & 2 \\ -5 & 2 & 1 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \frac{1}{12}\begin{bmatrix} 30 & -12 & 6 \\ 2 & 4 & 14 \\ -34 & 4 & 38 \end{bmatrix}.$

28. $\mathbf{P} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{4}\begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

29. $\mathbf{P} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 2 & -1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \frac{1}{3}\begin{bmatrix} 0 & 0 & 3 \\ 1 & 1 & 1 \\ 2 & 2 & -1 \end{bmatrix}.$

30. $\mathbf{P} = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}; \mathbf{P}^{-1} = \frac{1}{8}\begin{bmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{bmatrix}; \mathbf{A} = \mathbf{PDP}^{-1} = \frac{1}{8}\begin{bmatrix} 9 & 18 & 45 \\ 0 & 0 & 0 \\ 3 & 6 & 15 \end{bmatrix}.$

31. Eigenvalues of **B** are $2\lambda_j + (1/\lambda_j) - 1$, $j = 1, 2, 3, 4$ or $2, -4, 7/2, -11/2$. $|\mathbf{B}|$ = product of eigenvalues of $\mathbf{B} = 154$.

32. Eigenvalues of **B** are $\lambda_j + \lambda_j^2 - (1/\lambda_j)$, $j = 1, 2, 3$ or $1, 11/2, 1$. Trace of **B** = sum of eigenvalues of $\mathbf{B} = 15/2$.

33. Premultiply $\mathbf{Ax} = \lambda\mathbf{x}$ by $\mathbf{P}^{-1}$ and substitute $\mathbf{x} = \mathbf{Py}$.

34. Let λ be an eigenvalue and **x** be the corresponding eigenvector of **AB**, that is $\mathbf{ABx} = \lambda\mathbf{x}$. Premultiply by $\mathbf{A}^{-1}$ and substitute $\mathbf{x} = \mathbf{Ay}$. We get $\mathbf{BAy} = \lambda\mathbf{y}$. Therefore, λ is also an eigenvalue of **BA** and eigenvectors are related by $\mathbf{x} = \mathbf{Ay}$.

35. Let λ be an eigenvalue and **x** be the corresponding eigenvector of $\mathbf{A}^{-1}\mathbf{B}$, that is $\mathbf{A}^{-1}\mathbf{Bx} = \lambda\mathbf{x}$. Premultiply by **A** and set $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. We obtain $\mathbf{BA}^{-1}\mathbf{y} = \lambda\mathbf{y}$. Therefore, λ is also an eigenvalue of $\mathbf{BA}^{-1}$ with the corresponding eigenvector $\mathbf{y} = \mathbf{Ax}$.

36. From $\mathbf{Ax} = \lambda\mathbf{x}$, we obtain $\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x} = \mathbf{0}$. Therefore, $\lambda^k = 0$ or $\lambda = 0$, since $\mathbf{x} \neq \mathbf{0}$.

37. Since **A** is a diagonalizable matrix, there exists a non-singular matrix **P** such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ and the eigenvalues of **A** and **D** are same. We have $\mathbf{P}^{-1}\mathbf{A}^2\mathbf{P} = \mathbf{D}^2$. Since $\mathbf{A}^2 = \mathbf{A}$, we get $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}^2$. Therefore, we obtain $\mathbf{D}^2 - \mathbf{D} = \mathbf{0}$. Thus $\mathbf{D} = \mathbf{0}$ or $\mathbf{D} = \mathbf{I}$. Hence, the eigenvalues of **A** are 0 or 1.

38. Simplify the right hand side and set the off-diagonal element to zero.

39. Since $\mathbf{A}$ and $\mathbf{B}$ are similar, we have $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$. From this equation, show that $\mathbf{A}^{-1} = \mathbf{P}^{-1}\mathbf{B}^{-1}\mathbf{P}$ and $\mathbf{A}^m = \mathbf{P}^{-1}\mathbf{B}^m\mathbf{P}$. Also $|\mathbf{A}| = |\mathbf{P}^{-1}||\mathbf{B}||\mathbf{P}| = |\mathbf{B}|$.

40. We have $\mathbf{A} = \mathbf{A}^T$ and $\mathbf{B} = \mathbf{B}^T$. Therefore, $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T = \mathbf{BA}$.

41. $(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{A}$.

42. Let $\mathbf{A}^T\mathbf{A}^{-1}$ be a symmetric matrix. We have $(\mathbf{A}^T\mathbf{A}^{-1})^T = (\mathbf{A}^{-1})^T\mathbf{A} = \mathbf{A}^T\mathbf{A}^{-1}$, or $(\mathbf{A}^{-1})^T\mathbf{A}^2 = \mathbf{A}^T$ or $\mathbf{A}^2 = (\mathbf{A}^T)^2$. Now, let $\mathbf{A}^2 = (\mathbf{A}^T)^2$. We have $\mathbf{AA} = \mathbf{A}^T\mathbf{A}^T \Rightarrow \mathbf{A} = \mathbf{A}^{-1}\mathbf{A}^T\mathbf{A}^T \Rightarrow \mathbf{A}(\mathbf{A}^T)^{-1} = \mathbf{A}^{-1}\mathbf{A}^T$, or $\mathbf{A}(\mathbf{A}^{-1})^T = (\mathbf{A}^{-1}\mathbf{A}^T)^T = \mathbf{A}^{-1}\mathbf{A}^T$. Therefore, $\mathbf{A}^T\mathbf{A}^{-1}$ is symmetric.

43. Since $\mathbf{A}$ is symmetric, we have

$\mathbf{I} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^{-1}\mathbf{A}^T = (\mathbf{PDP}^{-1})^{-1}(\mathbf{PDP}^{-1})^T = (\mathbf{PD}^{-1}\mathbf{P}^{-1})\,[(\mathbf{P}^{-1})^T\mathbf{DP}^T]$, since $\mathbf{D}^T = \mathbf{D}$. This result is true only when $\mathbf{P}^{-1}(\mathbf{P}^{-1})^T = \mathbf{I}$, or $\mathbf{P}^{-1} = \mathbf{P}^T$.

44. Let $\mathbf{A}$ and $\mathbf{B}$ be the orthogonal matrices, that is $\mathbf{A}^{-1} = \mathbf{A}^T$ and $\mathbf{B}^{-1} = \mathbf{B}^T$. Then $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T = \mathbf{B}^{-1}\mathbf{A}^{-1} = (\mathbf{AB})^{-1}$.

45. $\mathbf{A}^{-1} = \mathbf{A}^T$ gives $\mathbf{AA}^T = \mathbf{I}$. We obtain the conditions as $l_i^2 + m_i^2 + n_i^2 = 1$, $i = 1, 2, 3$ and $l_1l_2 + m_1m_2 + n_1n_2 = 0$, $l_1l_3 + m_1m_3 + n_1n_3 = 0$, $l_2l_3 + m_2m_3 + n_2n_3 = 0$.

46. Since $\mathbf{A}$ is an orthogonal matrix, we have $\mathbf{A}^{-1} = \mathbf{A}^T$. Hence, $|\mathbf{A}^{-1}| = |\mathbf{A}^T| = |\mathbf{A}|$ or $1/|\mathbf{A}| = |\mathbf{A}| \Rightarrow |\mathbf{A}|^2 = 1$ or $|\mathbf{A}| = \pm 1$.

47. Let λ and μ be two distinct eigenvalues and $\mathbf{x}$, $\mathbf{y}$ be the corresponding eigenvectors. We have $\mathbf{Ax} = \lambda\mathbf{x}$ and $\mathbf{Ay} = \mu\mathbf{y}$. From the first equation, we get $\mathbf{x}^T\mathbf{A}^T = \lambda\mathbf{x}^T$ or $\mathbf{x}^T\mathbf{A} = \lambda\mathbf{x}^T$. Postmultiplying by $\mathbf{y}$, we obtain $\mathbf{x}^T\mathbf{Ay} = \lambda\mathbf{x}^T\mathbf{y}$. From the second equation, we get $\mathbf{x}^T\mathbf{Ay} = \mu\mathbf{x}^T\mathbf{y}$. Subtracting the two results, we obtain $(\lambda - \mu)\mathbf{x}^T\mathbf{y} = 0$, which gives $\mathbf{x}^T\mathbf{y} = 0$ since $\lambda \neq \mu$.

49. There exists an orthogonal matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$. Now, $\mathbf{A} = \mathbf{PDP}^{-1} = \mathbf{PDP}^T$, since $\mathbf{P}$ is orthogonal. We have $\mathbf{A}^T = (\mathbf{PDP}^T)^T = \mathbf{PD}^T\mathbf{P}^T = \mathbf{PDP}^T = \mathbf{A}$, since a diagonal matrix is always symmetric.

51. Let $\mathbf{z} = \mathbf{U} + i\,\mathbf{V}$, where $\mathbf{U} \neq \mathbf{0}$, $\mathbf{V} \neq \mathbf{0}$ be real vectors. Then

$$\overline{\mathbf{z}}^T\mathbf{Az} = (\mathbf{U}^T\mathbf{AU} + \mathbf{V}^T\mathbf{AV}) + i(\mathbf{U}^T\mathbf{AV} - \mathbf{V}^T\mathbf{AU}) = \mathbf{U}^T\mathbf{AU} + \mathbf{V}^T\mathbf{AV} > 0$$

since $\mathbf{U}^T\mathbf{AV} = (\mathbf{U}^T\mathbf{AV})^T = \mathbf{V}^T\mathbf{A}^T\mathbf{U} = \mathbf{V}^T\mathbf{AU}$.

52. Let the vectors $\mathbf{a}$, $\mathbf{b}$ be transformed to vectors $\mathbf{u}$, $\mathbf{v}$ respectively. Then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{u}}^T \cdot \mathbf{v} = (\overline{\mathbf{A}}\,\overline{\mathbf{a}})^T(\mathbf{Ab}) = \overline{\mathbf{a}}^T\overline{\mathbf{A}}^T\mathbf{Ab} = \overline{\mathbf{a}}^T\mathbf{b} = \mathbf{a} \cdot \mathbf{b}.$$

53. $$\mathbf{x}^T\mathbf{Ax} = [x_1, x_2]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2$$

$$= a[(x_1 + bx_2/a)^2 + x_2^2(ac - b^2)/a^2] > 0 \text{ for all } x_1, x_2.$$

Therefore, $a > 0$, $ac - b^2 > 0$.

54. $$\mathbf{x}^T\mathbf{Ax} = [x_1, x_2, x_3]\begin{bmatrix} 2 & 1 & 3 \\ -3 & 4 & -1 \\ -1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= 2x_1^2 - 2x_1x_2 + 2x_1x_3 + 4x_2^2 + 2x_3^2 = (x_1 - x_2)^2 + (x_1 + x_3)^2 + 3x_2^2 + x_3^2 > 0.$$

55. All the leading minors are not positive. If can also be verified that all the eigenvalues are not positive.

56. $\begin{bmatrix} 1 & -1 & 0 \\ -1 & -1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$.

57. $\begin{bmatrix} 3 & 1 & -2 \\ 1 & 1 & 4 \\ -2 & 4 & 0 \end{bmatrix}$.

58. $\begin{bmatrix} 1 & i & -4 \\ -i & 0 & 2i \\ -4 & -2i & 4 \end{bmatrix}$.

59. $\begin{bmatrix} 1 & -1-2i & 0 \\ -1+2i & 1 & -2+3i \\ 0 & -2-3i & 0 \end{bmatrix}$.

60. $\begin{bmatrix} 2 & 3+4i & 0 \\ 3-4i & -3 & 2-i \\ 0 & 2+i & 0 \end{bmatrix}$.