SCHAUM'S OUTLINE SERIES

THEORY AND PROBLEMS OF

# LINEAR ALGEBRA

SI (metric) edition

SEYMOUR LIPSCHUTZ

INCLUDING OVER 600 FULLY SOLVED PROBLEMS

1888 McGraw Hill 100 YEARS 1988

METRIC EDITIONS
Schaum's Outline Series

# Chapter 4

## Vector Spaces and Subspaces

### INTRODUCTION

In Chapter 1 we studied the concrete structures $\mathbf{R}^n$ and $\mathbf{C}^n$ and derived various properties. Now certain of these properties will play the role of axioms as we define abstract "vector spaces" or, as they are sometimes called, "linear spaces". In particular, the conclusions (i) through (viii) of Theorem 1.1, page 3, become axioms $[A_1]$-$[A_4]$, $[M_1]$-$[M_4]$ below. We will see that, in a certain sense, we get nothing new. In fact, we prove in Chapter 5 that every vector space over $\mathbf{R}$ which has "finite dimension" (defined there) can be identified with $\mathbf{R}^n$ for some $n$.

The definition of a vector space involves an arbitrary field (see Appendix B) whose elements are called *scalars*. We adopt the following notation (unless otherwise stated or implied):

| | |
|---|---|
| $K$ | the field of scalars, |
| $a, b, c$ or $k$ | the elements of $K$, |
| $V$ | the given vector space, |
| $u, v, w$ | the elements of $V$. |

We remark that nothing essential is lost if the reader assumes that $K$ is the real field $\mathbf{R}$ or the complex field $\mathbf{C}$.

Lastly, we mention that the "dot product", and related notions such as orthogonality, is not considered as part of the fundamental vector space structure, but as an additional structure which may or may not be introduced. Such spaces shall be investigated in the latter part of the text.

**Definition:** Let $K$ be a given field and let $V$ be a nonempty set with rules of addition and scalar multiplication which assigns to any $u, v \in V$ a *sum* $u + v \in V$ and to any $u \in V, k \in K$ a *product* $ku \in V$. Then $V$ is called a *vector space over* $K$ (and the elements of $V$ are called *vectors*) if the following axioms hold:

$[A_1]$: For any vectors $u, v, w \in V$, $(u+v)+w = u+(v+w)$.

$[A_2]$: There is a vector in $V$, denoted by $0$ and called the *zero vector*, for which $u + 0 = u$ for any vector $u \in V$.

$[A_3]$: For each vector $u \in V$ there is a vector in $V$, denoted by $-u$, for which $u + (-u) = 0$.

$[A_4]$: For any vectors $u, v \in V$, $u + v = v + u$.

$[M_1]$: For any scalar $k \in K$ and any vectors $u, v \in V$, $k(u+v) = ku + kv$.

$[M_2]$: For any scalars $a, b \in K$ and any vector $u \in V$, $(a+b)u = au + bu$.

$[M_3]$: For any scalars $a, b \in K$ and any vector $u \in V$, $(ab)u = a(bu)$.

$[M_4]$: For the unit scalar $1 \in K$, $1u = u$ for any vector $u \in V$.

The above axioms naturally split into two sets. The first four are only concerned with the additive structure of $V$ and can be summarized by saying that $V$ is a *commutative group* (see Appendix B) under addition. It follows that any sum of vectors of the form

$$v_1 + v_2 + \cdots + v_m$$

requires no parenthesis and does not depend upon the order of the summands, the zero vector 0 is unique, the *negative* $-u$ of $u$ is unique, and the *cancellation law* holds:

$$u + w = v + w \quad \text{implies} \quad u = v$$

for any vectors $u, v, w \in V$. Also, *subtraction* is defined by

$$u - v = u + (-v)$$

On the other hand, the remaining four axioms are concerned with the "action" of the field $K$ on $V$. Observe that the labelling of the axioms reflects this splitting. Using these additional axioms we prove (Problem 4.1) the following simple properties of a vector space.

**Theorem 4.1:** Let $V$ be a vector space over a field $K$.

(i) For any scalar $k \in K$ and $0 \in V$, $k0 = 0$.

(ii) For $0 \in K$ and any vector $u \in V$, $0u = 0$.

(iii) If $ku = 0$, where $k \in K$ and $u \in V$, then $k = 0$ or $u = 0$.

(iv) For any scalar $k \in K$ and any vector $u \in V$, $(-k)u = k(-u) = -ku$.

## EXAMPLES OF VECTOR SPACES

We now list a number of important examples of vector spaces. The first example is a generalization of the space $\mathbf{R}^n$.

**Example 4.1:** Let $K$ be an arbitrary field. The set of all $n$-tuples of elements of $K$ with vector addition and scalar multiplication defined by

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

and
$$k(a_1, a_2, \ldots, a_n) = (ka_1, ka_2, \ldots, ka_n)$$

where $a_i, b_i, k \in K$, is a vector space over $K$; we denote this space by $K^n$. The zero vector in $K^n$ is the $n$-tuple of zeros, $0 = (0, 0, \ldots, 0)$. The proof that $K^n$ is a vector space is identical to the proof of Theorem 1.1, which we may now regard as stating that $\mathbf{R}^n$ with the operations defined there is a vector space over $\mathbf{R}$.

**Example 4.2:** Let $V$ be the set of all $m \times n$ matrices with entries from an arbitrary field $K$. Then $V$ is a vector space over $K$ with respect to the operations of matrix addition and scalar multiplication, by Theorem 3.1.

**Example 4.3:** Let $V$ be the set of all polynomials $a_0 + a_1t + a_2t^2 + \cdots + a_nt^n$ with coefficients $a_i$ from a field $K$. Then $V$ is a vector space over $K$ with respect to the usual operations of addition of polynomials and multiplication by a constant.

**Example 4.4:** Let $K$ be an arbitrary field and let $X$ be any nonempty set. Consider the set $V$ of all functions from $X$ into $K$. The sum of any two functions $f, g \in V$ is the function $f + g \in V$ defined by

$$(f+g)(x) = f(x) + g(x)$$

and the product of a scalar $k \in K$ and a function $f \in V$ is the function $kf \in V$ defined by

$$(kf)(x) = k\,f(x)$$

Then $V$ with the above operations is a vector space over $K$ (Problem 4.5). The zero vector in $V$ is the zero function $\mathbf{0}$ which maps each $x \in X$ into $0 \in K$: $\mathbf{0}(x) = 0$ for every $x \in X$. Furthermore, for any function $f \in V$, $-f$ is that function in $V$ for which $(-f)(x) = -f(x)$, for every $x \in X$.

**Example 4.5:** Suppose $E$ is a field which contains a subfield $K$. Then $E$ can be considered to be a vector space over $K$, taking the usual addition in $E$ to be the vector addition and defining the scalar product $kv$ of $k \in K$ and $v \in E$ to be the product of $k$ and $v$ as element of the field $E$. Thus the complex field $\mathbf{C}$ is a vector space over the real field $\mathbf{R}$, and the real field $\mathbf{R}$ is a vector space over the rational field $\mathbf{Q}$.

## SUBSPACES

Let $W$ be a subset of a vector space V over a field $K$. $W$ is called a *subspace* of $V$ if $W$ is itself a vector space over $K$ with respect to the operations of vector addition and scalar multiplication on $V$. Simple criteria for identifying subspaces follow.

85, 86 sir

**Theorem 4.2:** $W$ is a subspace of $V$ if and only if

(i) $W$ is nonempty,

(ii) $W$ is closed under vector addition: $v, w \in W$ implies $v + w \in W$,

(iii) $W$ is closed under scalar multiplication: $v \in W$ implies $kv \in W$ for every $k \in K$.

**Corollary 4.3:** $W$ is a subspace of $V$ if and only if (i) $0 \in W$ (or $W \neq \emptyset$), and (ii) $v, w \in W$ implies $av + bw \in W$ for every $a, b \in K$.

sir

**Example 4.6:** Let $V$ be any vector space. Then the set $\{0\}$ consisting of the zero vector alone, and also the entire space $V$ are subspaces of $V$.

**Example 4.7:**

(i) Let $V$ be the vector space $\mathbf{R}^3$. Then the set $W$ consisting of those vectors whose third component is zero, $W = \{(a, b, 0) : a, b \in \mathbf{R}\}$, is a subspace of $V$.

(ii) Let $V$ be the space of all square $n \times n$ matrices (see Example 4.2). Then the set $W$ consisting of those matrices $A = (a_{ij})$ for which $a_{ij} = a_{ji}$, called *symmetric matrices*, is a subspace of $V$.

(iii) Let $V$ be the space of polynomials (see Example 4.3). Then the set $W$ consisting of polynomials with degree $\leq n$, for a fixed $n$, is a subspace of $V$.

(iv) Let $V$ be the space of all functions from a nonempty set $X$ into the real field $\mathbf{R}$. Then the set $W$ consisting of all bounded functions in $V$ is a subspace of $V$. (A function $f \in V$ is *bounded* if there exists $M \in \mathbf{R}$ such that $|f(x)| \leq M$ for every $x \in X$.)

**Example 4.8:** Consider any homogeneous system of linear equations in $n$ unknowns with, say, real coefficients:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

Recall that any particular solution of the system may be viewed as a point in $\mathbf{R}^n$. The set $W$ of all solutions of the homogeneous system is a subspace of $\mathbf{R}^n$ (Problem 4.16) called the *solution space*. We comment that the solution set of a nonhomogeneous system of linear equations in $n$ unknowns is not a subspace of $\mathbf{R}^n$.

**Example 4.9:** v.v.5 Let $U$ and $W$ be subspaces of a vector space $V$. We show that the intersection $U \cap W$ is also a subspace of $V$. Clearly $0 \in U$ and $0 \in W$ since $U$ and $W$ are subspaces; whence $0 \in U \cap W$. Now suppose $u, v \in U \cap W$. Then $u, v \in U$ and $u, v \in W$ and, since $U$ and $W$ are subspaces,

$$au + bv \in U \quad \text{and} \quad au + bv \in W$$

for any scalars $a, b \in K$. Accordingly, $au + bv \in U \cap W$ and so $U \cap W$ is a subspace of $V$.

The result in the preceding example generalizes as follows.

**Theorem 4.4:** The intersection of any number of subspaces of a vector space $V$ is a subspace of $V$.

87, 85

## LINEAR COMBINATIONS, LINEAR SPANS

Let $V$ be a vector space over a field $K$ and let $v_1, \ldots, v_m \in V$. Any vector in $V$ of the form

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m$$

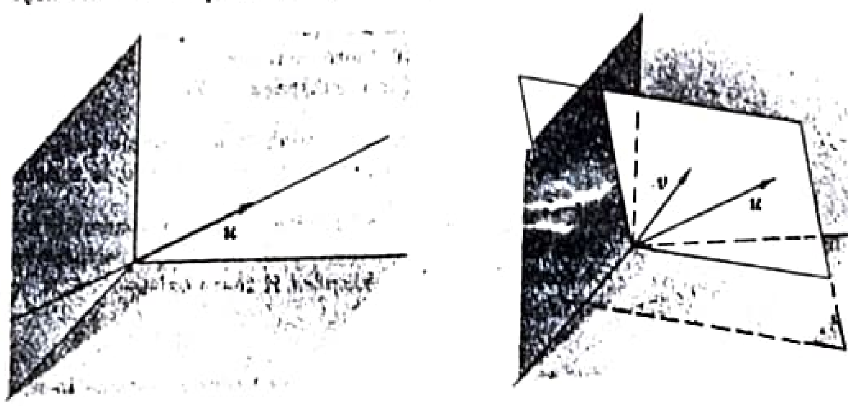
where the $a_i \in K$, is called a *linear combination* of $v_1, \ldots, v_m$. The following theorem applies.

**Theorem 4.5:** Let $S$ be a nonempty subset of $V$. The set of all linear combinations of vectors in $S$, denoted by $L(S)$, is a subspace of $V$ containing $S$. Furthermore, if $W$ is any other subspace of $V$ containing $S$, then $L(S) \subset W$.

In other words, $L(S)$ is the smallest subspace of $V$ containing $S$; hence it is called the subspace *spanned* or *generated* by $S$. For convenience, we define $L(\emptyset) = \{0\}$.

**Example 4.10:** Let $V$ be the vector space $\mathbf{R}^3$. The linear span of any nonzero vector $u$ consists of all scalar multiples of $u$; geometrically, it is the line through the origin and the point $u$. The linear space of any two vectors $u$ and $v$ which are not multiples of each other is the plane through the origin and the points $u$ and $v$.

31




**Example 4.11:** The vectors $e_1 = (1,0,0)$, $e_2 = (0,1,0)$ and $e_3 = (0,0,1)$ generate the vector space $\mathbf{R}^3$. For any vector $(a, b, c) \in \mathbf{R}^3$ is a linear combination of the $e_i$; specifically,

$$\begin{aligned}(a, b, c) &= a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) \\ &= ae_1 + be_2 + ce_3\end{aligned}$$

**Example 4.12:** The polynomials $1, t, t^2, t^3, \ldots$ generate the vector space $V$ of all polynomials (in $t$): $V = L(1, t, t^2, \ldots)$. For any polynomial is a linear combination of 1 and powers of $t$.

**Example 4.13:** Determine whether or not the vector $v = (3, 9, -4, -2)$ is a linear combination of the vectors $u_1 = (1, -2, 0, 3)$, $u_2 = (2, 3, 0, -1)$ and $u_3 = (2, -1, 2, 1)$, i.e. belongs to the space spanned by the $u_i$.

Set $v$ as a linear combination of the $u_i$ using unknowns $x$, $y$ and $z$; that is, set $v = xu_1 + yu_2 + zu_3$:

$$\begin{aligned}(3, 9, -4, -2) &= x(1, -2, 0, 3) + y(2, 3, 0, -1) + z(2, -1, 2, 1)\\ &= (x + 2y + 2z, -2x + 3y - z, 2z, 3x - y + z)\end{aligned}$$

Form the equivalent system of equations by setting corresponding components equal to each other, and then reduce to echelon form:

$$\begin{aligned} x + 2y + 2z &= 3\\ -2x + 3y - z &= 9\\ 2z &= -4\\ 3x - y + z &= -2 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4\\ -7y - 5z &= -11 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4\\ -2z &= 4 \end{aligned}$$

$$\text{or} \quad \begin{aligned} x + 2y + 2z &= 3\\ 7y + 3z &= 15\\ 2z &= -4 \end{aligned}$$

Note that the above system is consistent and so has a solution; hence $v$ is a linear combination of the $u_i$. Solving for the unknowns we obtain $x = 1$, $y = 3$, $z = -2$. Thus $v = u_1 + 3u_2 - 2u_3$.

Note that if the system of linear equations were not consistent, i.e. had no solution, then the vector $v$ would not be a linear combination of the $u_i$.

## ROW SPACE OF A MATRIX

Let $A$ be an arbitrary $m \times n$ matrix over a field $K$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

The rows of $A$,

$$R_1 = (a_{11}, a_{21}, \dots, a_{1n}), \quad \dots, \quad R_m = (a_{m1}, a_{m2}, \dots, a_{mn})$$

viewed as vectors in $K^n$, span a subspace of $K^n$ called the *row space* of $A$. That is,

$$\text{row space of } A = L(R_1, R_2, \dots, R_m)$$

Analogously, the columns of $A$, viewed as vectors in $K^m$, span a subspace of $K^m$ called the *column space* of $A$.

Now suppose we apply an elementary row operation on $A$,

$$\text{(i) } R_i \leftrightarrow R_j, \quad \text{(ii) } R_i \to kR_i,\ k \neq 0, \quad \text{or} \quad \text{(iii) } R_i \to kR_j + R_i$$

and obtain a matrix $B$. Then each row of $B$ is clearly a row of $A$ or a linear combination of rows of $A$. Hence the row space of $B$ is contained in the row space of $A$. On the other hand, we can apply the inverse elementary row operation on $B$ and obtain $A$; hence the row space of $A$ is contained in the row space of $B$. Accordingly, $A$ and $B$ have the same row space. This leads us to the following theorem.

**Theorem 4.6:** Row equivalent matrices have the same row space.

We shall prove (Problem 4.31), in particular, the following fundamental result concerning row reduced echelon matrices.

**Theorem 4.7:** Row reduced echelon matrices have the same row space if and only if they have the same nonzero rows.

[Thus every matrix is row equivalent to a unique row reduced echelon matrix called its *row canonical form.*]

We apply the above results in the next example.

**Example 4.14:** Show that the space $U$ generated by the vectors

$$u_1 = (1, 2, -1, 3), \quad u_2 = (2, 4, 1, -2), \quad \text{and} \quad u_3 = (3, 6, 3, -7)$$

and the space $V$ generated by the vectors

$$v_1 = (1, 2, -4, 11) \quad \text{and} \quad v_2 = (2, 4, -5, 14)$$

are equal; that is, $U = V$.

**Method 1.** Show that each $u_i$ is a linear combination of $v_1$ and $v_2$, and show that each $v_i$ is a linear combination of $u_1$, $u_2$ and $u_3$. Observe that we have to show that six systems of linear equations are consistent.

**Method 2.** Form the matrix $A$ whose rows are the $u_i$, and row reduce $A$ to row canonical form:

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{to } \begin{pmatrix} 1 & 2 & 0 & 1/3 \\ 0 & 0 & 1 & -8/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Now form the matrix $B$ whose rows are $v_1$ and $v_2$, and row reduce $B$ to row canonical form:

$$B = \begin{pmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & 0 & 1/3 \\ 0 & 0 & 1 & -8/3 \end{pmatrix}$$

Since the nonzero rows of the reduced matrices are identical, the row spaces of $A$ and $B$ are equal and so $U = V$.

## SUMS AND DIRECT SUMS

Let $U$ and $W$ be subspaces of a vector space $V$. The sum of $U$ and $W$, written $U + W$, consists of all sums $u + w$ where $u \in U$ and $w \in W$:

$$U + W = \{u + w : u \in U, w \in W\}$$

Note that $0 = 0 + 0 \in U + W$, since $0 \in U$, $0 \in W$. Furthermore, suppose $u + w$ and $u' + w'$ belong to $U + W$, with $u, u' \in U$ and $w, w' \in W$. Then

$$(u + w) + (u' + w') = (u + u') + (w + w') \in U + W$$

and, for any scalar $k$, $\quad k(u + w) = ku + kw \in U + W$

Thus we have proven the following theorem.

**Theorem 4.8:** The sum $U + W$ of the subspaces $U$ and $W$ of $V$ is also a subspace of $V$.

**Example 4.15:** Let $V$ be the vector space of 2 by 2 matrices over $\mathbf{R}$. Let $U$ consist of those matrices in $V$ whose second row is zero, and let $W$ consist of those matrices in $V$ whose second column is zero:

$$U = \left\{\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbf{R}\right\}, \qquad W = \left\{\begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} : a, c \in \mathbf{R}\right\}$$

Now $U$ and $W$ are subspaces of $V$. We have:

$$U + W = \left\{\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} : a, b, c \in \mathbf{R}\right\} \quad \text{and} \quad U \cap W = \left\{\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} : a \in \mathbf{R}\right\}$$

That is, $U + W$ consists of those matrices whose lower right entry is 0, and $U \cap W$ consists of those matrices whose second row and second column are zero.

**Definition:** 84 The vector space $V$ is said to be the *direct sum* of its subspaces $U$ and $W$, denoted by

$$V = U \oplus W$$

if every vector $v \in V$ can be written in one and only one way as $v = u + w$ where $u \in U$ and $w \in W$.

The following theorem applies.

**Theorem 4.9:** 89 The vector space $V$ is the direct sum of its subspaces $U$ and $W$ if and only if: (i) $V = U + W$, and (ii) $U \cap W = \{0\}$.

* **Example 4.16:** In the vector space $\mathbf{R}^3$, let $U$ be the $xy$ plane and let $W$ be the $yz$ plane:

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \quad \text{and} \quad W = \{(0, b, c) : b, c \in \mathbf{R}\}$$

Then $\mathbf{R}^3 = U + W$ since every vector in $\mathbf{R}^3$ is the sum of a vector in $U$ and a vector in $W$. However, $\mathbf{R}^3$ is not the direct sum of $U$ and $W$ since such sums are not unique; for example,

$$(3, 5, 7) = (3, 1, 0) + (0, 4, 7) \quad \text{and also} \quad (3, 5, 7) = (3, -4, 0) + (0, 9, 7)$$

* **Example 4.17:** In $\mathbf{R}^3$, let $U$ be the $xy$ plane and let $W$ be the $z$ axis:

$$U = \{(a, b, 0) : a, b \in \mathbf{R}\} \quad \text{and} \quad W = \{(0, 0, c) : c \in \mathbf{R}\}$$

Now any vector $(a, b, c) \in \mathbf{R}^3$ can be written as the sum of a vector in $U$ and a vector in $V$ in one and only one way:

$$(a, b, c) = (a, b, 0) + (0, 0, c)$$

Accordingly, $\mathbf{R}^3$ is the direct sum of $U$ and $W$, that is, $\mathbf{R}^3 = U \oplus W$.

## Solved Problems

### VECTOR SPACES

**4.1.** Prove Theorem 4.1: Let $V$ be a vector space over a field $K$.

(i) For any scalar $k \in K$ and $0 \in V$, $k0 = 0$.

(ii) For $0 \in K$ and any vector $u \in V$, $0u = 0$.

(iii) If $ku = 0$, where $k \in K$ and $u \in V$, then $k = 0$ or $u = 0$.

(iv) For any $k \in K$ and any $u \in V$, $(-k)u = k(-u) = -ku$.

(i) By axiom $[A_2]$ with $u = 0$, we have $0 + 0 = 0$. Hence by axiom $[M_1]$, $k0 = k(0 + 0) = k0 + k0$. Adding $-k0$ to both sides gives the desired result.

(ii) By a property of $K$, $0 + 0 = 0$. Hence by axiom $[M_2]$, $0u = (0 + 0)u = 0u + 0u$. Adding $-0u$ to both sides yields the required result.

(iii) Suppose $ku = 0$ and $k \neq 0$. Then there exists a scalar $k^{-1}$ such that $k^{-1}k = 1$; hence

$$u = 1u = (k^{-1}k)u = k^{-1}(ku) = k^{-1}0 = 0$$

(iv) Using $u + (-u) = 0$, we obtain $0 = k0 = k(u + (-u)) = ku + k(-u)$. Adding $-ku$ to both sides gives $-ku = k(-u)$.

Using $k + (-k) = 0$, we obtain $0 = 0u = (k + (-k))u = ku + (-k)u$. Adding $-ku$ to both sides yields $-ku = (-k)u$. Thus $(-k)u = k(-u) = -ku$.

**4.2.** **Show that for any scalar $k$ and any vectors $u$ and $v$, $k(u - v) = ku - kv$.**

Using the definition of subtraction $(u - v = u + (-v))$ and the result of Theorem 4.1(iv) $(k(-v) = -kv)$,

$$k(u - v) = k(u + (-v)) = ku + k(-v) = ku + (-kv) = ku - kv$$

**4.3.** **In the statement of axiom $[M_2]$, $(a + b)u = au + bu$, which operation does each plus sign represent?**

The $+$ in $(a + b)u$ denotes the addition of the two scalars $a$ and $b$; hence it represents the addition operation in the field $K$. On the other hand, the $+$ in $au + bu$ denotes the addition of the two vectors $au$ and $bu$; hence it represents the operation of vector addition. Thus each $+$ represents a different operation.

**4.4.** **In the statement of axiom $[M_3]$, $(ab)u = a(bu)$, which operation does each product represent?**

In $(ab)$ ie product $ab$ of the scalars $a$ and $b$ denotes multiplication in the field $K$, whereas the product of the scalar $ab$ and the vector $u$ denotes scalar multiplication.

In $a(bu)$ the product $bu$ of the scalar $b$ and the vector $u$ denotes scalar multiplication; also, the product of the scalar $a$ and the vector $bu$ denotes scalar multiplication.

**4.5.** **Let $V$ be the set of all functions from a nonempty set $X$ into a field $K$. For any functions $f, g \in V$ and any scalar $k \in K$, let $f + g$ and $kf$ be the functions in $V$ defined as follows:**

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (kf)(x) = kf(x), \qquad \forall x \in X$$

**(The symbol $\forall$ means "for every".) Prove that $V$ is a vector space over $K$.**

Since $X$ is nonempty, $V$ is also nonempty. We now need to show that all the axioms of a vector space hold.

$[A_1]$: Let $f, g, h \in V$. To show that $(f + g) + h = f + (g + h)$, it is necessary to show that the function $(f + g) + h$ and the function $f + (g + h)$ both assign the same value to each $x \in X$. Now,

$$((f + g) + h)(x) = (f + g)(x) + h(x) = (f(x) + g(x)) + h(x), \qquad \forall x \in X$$

$$(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)), \qquad \forall x \in X$$

But $f(x)$, $g(x)$ and $h(x)$ are scalars in the field $K$ where addition of scalars is associative; hence

$$(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$$

Accordingly, $(f + g) + h = f + (g + h)$.

$[A_2]$: Let $0$ denote the zero function: $0(x) = 0$, $\forall x \in X$. Then for any function $f \in V$,

$$(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x), \qquad \forall x \in X$$

Thus $f + 0 = f$, and $0$ is the zero vector in $V$.

$[A_3]$: For any function $f \in V$, let $-f$ be the function defined by $(-f)(x) = -f(x)$. Then,

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0 = \mathbf{0}(x), \quad \forall x \in X$$

Hence $f + (-f) = \mathbf{0}$.

$[A_4]$: Let $f, g \in V$. Then

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x), \quad \forall x \in X$$

Hence $f + g = g + f$. (Note that $f(x) + g(x) = g(x) + f(x)$ follows from the fact that $f(x)$ and $g(x)$ are scalars in the field $K$ where addition is commutative.)

$[M_1]$: Let $f, g \in V$ and $k \in K$. Then

$$\begin{aligned}(k(f + g))(x) &= k((f + g)(x)) = k(f(x) + g(x)) = kf(x) + kg(x)\\ &= (kf)(x) + (kg)(x) = (kf + kg)(x), \quad \forall x \in X\end{aligned}$$

Hence $k(f + g) = kf + kg$. (Note that $k(f(x) + g(x)) = kf(x) + kg(x)$ follows from the fact that $k$, $f(x)$ and $g(x)$ are scalars in the field $K$ where multiplication is distributive over addition.)

$[M_2]$: Let $f \in V$ and $a, b \in K$. Then

$$\begin{aligned}((a + b)f)(x) &= (a + b)f(x) = af(x) + bf(x) = (af)(x) + bf(x)\\ &= (af + bf)(x), \quad \forall x \in X\end{aligned}$$

Hence $(a + b)f = af + bf$.

$[M_3]$: Let $f \in V$ and $a, b \in K$. Then,

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a(bf)(x) = (a(bf))(x), \quad \forall x \in X$$

Hence $(ab)f = a(bf)$.

$[M_4]$: Let $f \in V$. Then, for the unit $1 \in K$, $(1f)(x) = 1f(x) = f(x)$, $\forall x \in X$. Hence $1f = f$.

Since all the axioms are satisfied, $V$ is a vector space over $K$.

**4.6.** Let $V$ be the set of ordered pairs of real numbers: $V = \{(a, b) : a, b \in \mathbf{R}\}$. Show that $V$ is not a vector space over $R$ with respect to each of the following operations of addition in $V$ and scalar multiplication on $V$:

(i) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (ka, b)$;

(ii) $(a, b) + (c, d) = (a, b)$ and $k(a, b) = (ka, kb)$;

(iii) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (k^2a, k^2b)$.

In each case show that one of the axioms of a vector space does not hold.

(i) Let $r = 1$, $s = 2$, $v = (3, 4)$. Then

$$(r + s)v = 3(3, 4) = (9, 4)$$

$$rv + sv = 1(3, 4) + 2(3, 4) = (3, 4) + (6, 4) = (9, 8)$$

Since $(r + s)v \neq rv + sv$, axiom $[M_2]$ does not hold.

(ii) Let $v = (1, 2)$, $w = (3, 4)$. Then

$$v + w = (1, 2) + (3, 4) = (1, 2)$$

$$w + v = (3, 4) + (1, 2) = (3, 4)$$

Since $v + w \neq w + v$, axiom $[A_4]$ does not hold.

(iii) Let $r = 1$, $s = 2$, $v = (3, 4)$. Then

$$(r + s)v = 3(3, 4) = (27, 36)$$

$$rv + sv = 1(3, 4) + 2(3, 4) = (3, 4) + (12, 16) = (15, 20)$$

Thus $(r + s)v \neq rv + sv$, and so axiom $[M_2]$ does not hold.

## SUBSPACES

**4.7.** Prove Theorem 4.2: $W$ is a subspace of $V$ if and only if (i) $W$ is nonempty, (ii) $v, w \in W$ implies $v + w \in W$, and (iii) $v \in W$ implies $kv \in W$ for every scalar $k \in K$.

Suppose $W$ satisfies (i), (ii) and (iii). By (i), $W$ is nonempty; and by (ii) and (iii), the operations of vector addition and scalar multiplication are well defined for $W$. Moreover, the axioms $[A_1]$, $[A_4]$, $[M_1]$, $[M_2]$, $[M_3]$ and $[M_4]$ hold in $W$ since the vectors in $W$ belong to $V$. Hence we need only show that $[A_2]$ and $[A_3]$ also hold in $W$. By (i), $W$ is nonempty, say $u \in W$. Then by (iii), $0u = 0 \in W$ and $v + 0 = v$ for every $v \in W$. Hence $W$ satisfies $[A_2]$. Lastly, if $v \in W$ then $(-1)v = -v \in W$ and $v + (-v) = 0$; hence $W$ satisfies $[A_3]$. Thus $W$ is a subspace of $V$.

Conversely, if $W$ is a subspace of $V$ then clearly (i), (ii) and (iii) hold.

**4.8.** Prove Corollary 4.3: $W$ is a subspace of $V$ if and only if (i) $0 \in W$ and (ii) $v, w \in W$ implies $av + bw \in W$ for all scalars $a, b \in K$.

Suppose $W$ satisfies (i) and (ii). Then, by (i), $W$ is nonempty. Furthermore, if $v, w \in W$ then, by (ii), $v + w = 1v + 1w \in W$; and if $v \in W$ and $k \in K$ then, by (ii), $kv = kv + 0v \in W$. Thus by Theorem 4.2, $W$ is a subspace of $V$.

Conversely, if $W$ is a subspace of $V$ then clearly (i) and (ii) hold in $W$.

**4.9.** Let $V = \mathbf{R}^3$. Show that $W$ is a subspace of $V$ where:

(i) $W = \{(a, b, 0) : a, b \in \mathbf{R}\}$, i.e. $W$ is the $xy$ plane consisting of those vectors whose third component is 0;

(ii) $W = \{(a, b, c) : a + b + c = 0\}$, i.e. $W$ consists of those vectors each with the property that the sum of its components is zero.

(i) $0 = (0, 0, 0) \in W$ since the third component of 0 is 0. For any vectors $v = (a, b, 0)$, $w = (c, d, 0)$ in $W$, and any scalars (real numbers) $k$ and $k'$,

$$\begin{aligned} kv + k'w &= k(a, b, 0) + k'(c, d, 0) \\ &= (ka, kb, 0) + (k'c, k'd, 0) = (ka + k'c,\ kb + k'd,\ 0) \end{aligned}$$

Thus $kv + k'w \in W$, and so $W$ is a subspace of $V$.

(ii) $0 = (0, 0, 0) \in W$ since $0 + 0 + 0 = 0$. Suppose $v = (a, b, c)$, $w = (a', b', c')$ belong to $W$, i.e. $a + b + c = 0$ and $a' + b' + c' = 0$. Then for any scalars $k$ and $k'$,

$$\begin{aligned} kv + k'w &= k(a, b, c) + k'(a', b', c') \\ &= (ka, kb, kc) + (k'a', k'b', k'c') \\ &= (ka + k'a',\ kb + k'b',\ kc + k'c') \end{aligned}$$

and furthermore,

$$\begin{aligned} (ka + k'a') + (kb + k'b') + (kc + k'c') &= k(a + b + c) + k'(a' + b' + c') \\ &= k0 + k'0 = 0 \end{aligned}$$

Thus $kv + k'w \in W$, and so $W$ is a subspace of $V$.

**4.10.** Let $V = \mathbf{R}^3$. Show that $W$ is not a subspace of $V$ where:

(i) $W = \{(a, b, c) : a \geq 0\}$, i.e. $W$ consists of those vectors whose first component is nonnegative;

(ii) $W = \{(a, b, c) : a^2 + b^2 + c^2 \leq 1\}$, i.e. $W$ consists of those vectors whose length does not exceed 1;

(iii) $W = \{(a, b, c) : a, b, c \in \mathbf{Q}\}$, i.e. $W$ consists of those vectors whose components are rational numbers.

In each case, show that one of the properties of, say, Theorem 4.2 does not hold.

(i) $v = (1, 2, 3) \in W$ and $k = -5 \in \mathbf{R}$. But $kv = -5(1, 2, 3) = (-5, -10, -15)$ does not belong to $W$ since $-5$ is negative. Hence $W$ is not a subspace of $V$.

(ii) $v = (1,0,0) \in W$ and $w = (0,1,0) \in W$. But $v + w = (1,0,0) + (0,1,0) = (1,1,0)$ does not belong to $W$ since $1^2 + 1^2 + 0^2 = 2 > 1$. Hence $W$ is not a subspace of $V$.

(iii) $v = (1,2,3) \in W$ and $k = \sqrt{2} \in \mathbf{R}$. But $kv = \sqrt{2}(1,2,3) = (\sqrt{2}, 2\sqrt{2}, 3\sqrt{2})$ does not belong to $W$ since its components are not rational numbers. Hence $W$ is not a subspace of $V$.

**4.11.** Let $V$ be the vector space of all square $n \times n$ matrices over a field $K$. Show that $W$ is a subspace of $V$ where:

(i) $W$ consists of the symmetric matrices, i.e. all matrices $A = (a_{ij})$ for which $a_{ji} = a_{ij}$;

(ii) $W$ consists of all matrices which commute with a given matrix $T$; that is, $W = \{A \in V : AT = TA\}$.

(i) $0 \in W$ since all entries of $0$ are $0$ and hence equal. Now suppose $A = (a_{ij})$ and $B = (b_{ij})$ belong to $W$, i.e. $a_{ji} = a_{ij}$ and $b_{ji} = b_{ij}$. For any scalars $a, b \in K$, $aA + bB$ is the matrix whose $ij$-entry is $aa_{ij} + bb_{ij}$. But $aa_{ji} + bb_{ji} = aa_{ij} + bb_{ij}$. Thus $aA + bB$ is also symmetric, and so $W$ is a subspace of $V$.

(ii) $0 \in W$ since $0T = 0 = T0$. Now suppose $A, B \in W$; that is, $AT = TA$ and $BT = TB$. For any scalars $a, b \in K$,

$$\begin{aligned}(aA + bB)T &= (aA)T + (bB)T = a(AT) + b(BT) = a(TA) + b(TB) \\ &= T(aA) + T(bB) = T(aA + bB)\end{aligned}$$

Thus $aA + bB$ commutes with $T$, i.e. belongs to $W$; hence $W$ is a subspace of $V$.

**4.12.** Let $V$ be the vector space of all $2 \times 2$ matrices over the real field $\mathbf{R}$. Show that $W$ is not a subspace of $V$ where:

(i) $W$ consists of all matrices with zero determinant;

(ii) $W$ consists of all matrices $A$ for which $A^2 = A$.

(i) (Recall that $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.) The matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ belong to $W$ since $\det(A) = 0$ and $\det(B) = 0$. But $A + B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ does not belong to $W$ since $\det(A + B) = 1$. Hence $W$ is not a subspace of $V$.

(ii) The unit matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ belongs to $W$ since

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

But $2I = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ does not belong to $W$ since

$$(2I)^2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \neq 2I$$

Hence $W$ is not a subspace of $V$.

**4.13.** Let $V$ be the vector space of all functions from the real field $\mathbf{R}$ into $\mathbf{R}$. Show that $W$ is a subspace of $V$ where:

(i) $W = \{f : f(3) = 0\}$, i.e. $W$ consists of those functions which map 3 into 0;

(ii) $W = \{f : f(7) = f(1)\}$, i.e. $W$ consists of those functions which assign the same value to 7 and 1;

(iii) $W$ consists of the odd functions, i.e. those functions $f$ for which $f(-x) = -f(x)$.

Here $0$ denotes the zero function: $0(x) = 0$, for every $x \in R$.

(i) $0 \in W$ since $0(3) = 0$. Suppose $f, g \in W$ i.e. $f(3) = 0$ and $g(3) = 0$. Then for any real numbers $a$ and $b$,

$$(af + bg)(3) = af(3) + bg(3) = a0 + b0 = 0$$

Hence $af + bg \in W$, and so $W$ is a subspace of $V$.

(ii) $0 \in W$ since $0(7) = 0 = 0(1)$. Suppose $f, g \in W$, i.e. $f(7) = f(1)$ and $g(7) = g(1)$. Then, for any real numbers $a$ and $b$,

$$(af + bg)(7) = af(7) + bg(7) = af(1) + bg(1) = (af + bg)(1)$$

Hence $af + bg \in W$, and so $W$ is a subspace of $V$.

(iii) $0 \in W$ since $0(-x) = 0 = -0 = -0(x)$. Suppose $f, g \in W$, i.e. $f(-x) = -f(x)$ and $g(-x) = -g(x)$. Then for any real numbers $a$ and $b$,

$$(af + bg)(-x) = af(-x) + bg(-x) = -af(x) - bg(x) = -(af(x) + bg(x)) = -(af + bg)(x)$$

Hence $af + bg \in W$, and so $W$ is a subspace of $V$.

**4.14.** Let $V$ be the vector space of all functions from the real field $\mathbf{R}$ into $\mathbf{R}$. Show that $W$ is not a subspace of $V$ where:

(i) $W = \{f : f(7) = 2 + f(1)\}$;

(ii) $W$ consists of all nonnegative functions, i.e. all functions $f$ for which $f(x) \geq 0$, $\forall x \in \mathbf{R}$.

(i) Suppose $f, g \in W$, i.e. $f(7) = 2 + f(1)$ and $g(7) = 2 + g(1)$. Then

$$\begin{aligned}(f + g)(7) &= f(7) + g(7) = 2 + f(1) + 2 + g(1) \\ &= 4 + f(1) + g(1) = 4 + (f + g)(1) \neq 2 + (f + g)(1)\end{aligned}$$

Hence $f + g \notin W$, and so $W$ is not a subspace of $V$.

(ii) Let $k = -2$ and let $f \in V$ be defined by $f(x) = x^2$. Then $f \in W$ since $f(x) = x^2 \geq 0$, $\forall x \in \mathbf{R}$. But $(kf)(5) = kf(5) = (-2)(5^2) = -50 < 0$. Hence $kf \notin W$, and so $W$ is not a subspace of $V$.

**4.15.** Let $V$ be the vector space of polynomials $a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ with real coefficients, i.e. $a_i \in \mathbf{R}$. Determine whether or not $W$ is a subspace of $V$ where:

(i) $W$ consists of all polynomials with integral coefficients;

(ii) $W$ consists of all polynomials with degree $\leq 3$;

(iii) $W$ consists of all polynomials $b_0 + b_1 t^2 + b_2 t^4 + \cdots + b_n t^{2n}$, i.e. polynomials with only even powers of $t$.

(i) No, since scalar multiples of vectors in $W$ do not always belong to $W$. For example, $v = 3 + 5t + 7t^2 \in W$ but $\frac{1}{2}v = \frac{3}{2} + \frac{5}{2}t + \frac{7}{2}t^2 \notin W$. (Observe that $W$ is "closed" under vector addition, i.e. sums of elements in $W$ belong to $W$.)

(ii) and (iii). Yes. For, in each case, $W$ is nonempty, the sum of elements in $W$ belong to $W$, and the scalar multiples of any element in $W$ belong to $W$.

**4.16.** Consider a homogeneous system of linear equations in $n$ unknowns $x_1, \ldots, x_n$ over a field $K$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Show that the solution set $W$ is a subspace of the vector space $K^n$.

$0 = (0, 0, \ldots, 0) \in W$ since, clearly,

$$a_{i1}0 + a_{i2}0 + \cdots + a_{in}0 = 0, \quad \text{for } i = 1, \ldots, m$$

Suppose $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ belong to $W$, i.e. for $i = 1, \ldots, m$

$$a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = 0$$
$$a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n = 0$$

Let $a$ and $b$ be scalars in $K$. Then

$$au + bv = (au_1 + bv_1, au_2 + bv_2, \ldots, au_n + bv_n)$$

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

$$\begin{aligned} a_{i1}(au_1 + bv_1) + a_{i2}(au_2 + bv_2) + \cdots + a_{in}(au_n + bv_n) &\\ = a(a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n) + b(a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n) &\\ = a0 + b0 = 0 & \end{aligned}$$

Hence $au + bv$ is a solution of the system, i.e. belongs to $W$. Accordingly, $W$ is a subspace of $K^n$.

## LINEAR COMBINATIONS

**4.17.** Write the vector $v = (1, -2, 5)$ as a linear combination of the vectors $e_1 = (1, 1, 1)$, $e_2 = (1, 2, 3)$ and $e_3 = (2, -1, 1)$.

We wish to express $v$ as $v = xe_1 + ye_2 + ze_3$, with $x$, $y$ and $z$ as yet unknown scalars. Thus we require

$$\begin{aligned} (1, -2, 5) &= x(1, 1, 1) + y(1, 2, 3) + z(2, -1, 1) \\ &= (x, x, x) + (y, 2y, 3y) + (2z, -z, z) \\ &= (x + y + 2z, x + 2y - z, x + 3y + z) \end{aligned}$$

Form the equivalent system of equations by setting corresponding components equal to each other, and then reduce to echelon form:

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned} \quad \text{or} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \quad \text{or} \quad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

Note that the above system is consistent and so has a solution. Solve for the unknowns to obtain $x = -6$, $y = 3$, $z = 2$. Hence $v = -6e_1 + 3e_2 + 2e_3$.

**4.18.** Write the vector $v = (2, -5, 3)$ in $\mathbf{R}^3$ as a linear combination of the vectors $e_1 = (1, -3, 2)$, $e_2 = (2, -4, -1)$ and $e_3 = (1, -5, 7)$.

Set $v$ as a linear combination of the $e_i$ using the unknowns $x$, $y$ and $z$: $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} (2, -5, 3) &= x(1, -3, 2) + y(2, -4, -1) + z(1, -5, 7) \\ &= (x + 2y + z, -3x - 4y - 5z, 2x - y + 7z) \end{aligned}$$

Form the equivalent system of equations and reduce to echelon form:

$$\begin{aligned} x + 2y + z &= 2 \\ -3x - 4y - 5z &= -5 \\ 2x - y + 7z &= 3 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ -5y + 5z &= -1 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + z &= 2 \\ 2y - 2z &= 1 \\ 0 &= 3 \end{aligned}$$

The system is inconsistent and so has no solution. Accordingly, $v$ cannot be written as a linear combination of the vectors $e_1$, $e_2$ and $e_3$.

v.v.v.I

**4.19.** For which value of $k$ will the vector $u = (1, -2, k)$ in $\mathbf{R}^3$ be a linear combination of the vectors $v = (3, 0, -2)$ and $w = (2, -1, -5)$?

Set $u = xv + yw$:

$$(1, -2, k) = x(3, 0, -2) + y(2, -1, -5) = (3x + 2y, -y, -2x - 5y)$$

Form the equivalent system of equations:

$$3x + 2y = 1, \quad -y = -2, \quad -2x - 5y = k$$

By the first two equations, $x = -1$, $y = 2$. Substitute into the last equation to obtain $k = -8$.

**4.20.** Write the polynomial $v = t^2 + 4t - 3$ over $\mathbf{R}$ as a linear combination of the polynomials $e_1 = t^2 - 2t + 5$, $e_2 = 2t^2 - 3t$ and $e_3 = t + 3$.

S

Set $v$ as a linear combination of the $e_i$ using the unknowns $x$, $y$ and $z$: $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} t^2 + 4t - 3 &= x(t^2 - 2t + 5) + y(2t^2 - 3t) + z(t+3) \\ &= xt^2 - 2xt + 5x + 2yt^2 - 3yt + zt + 3z \\ &= (x+2y)t^2 + (-2x - 3y + z)t + (5x + 3z) \end{aligned}$$

Set coefficients of the same powers of $t$ equal to each other, and reduce the system to echelon form:

$$\begin{aligned} x + 2y \quad &= 1 \\ -2x - 3y + z &= 4 \\ 5x \quad + 3z &= -3 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y \quad &= 1 \\ y + z &= 6 \\ -10y + 3z &= -8 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y \quad &= 1 \\ y + z &= 6 \\ 13z &= 52 \end{aligned}$$

Note that the system is consistent and so has a solution. Solve for the unknowns to obtain $x = -3$, $y = 2$, $z = 4$. Thus $v = -3e_1 + 2e_2 + 4e_3$.

v.v.S

S **4.21.** Write the matrix $E = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ as a linear combination of the matrices $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix}$.

Set $E$ as a linear combination of $A, B, C$ using the unknowns $x, y, z$: $E = xA + yB + zC$.

$$\begin{aligned} \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix} &= x\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + z\begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} x & x \\ x & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ y & y \end{pmatrix} + \begin{pmatrix} 0 & 2z \\ 0 & -z \end{pmatrix} = \begin{pmatrix} x & x + 2z \\ x + y & y - z \end{pmatrix} \end{aligned}$$

Form the equivalent system of equations by setting corresponding entries equal to each other:

$$x = 3, \quad x + y = 1, \quad x + 2z = 1, \quad y - z = -1$$

Substitute $x = 3$ in the second and third equations to obtain $y = -2$ and $z = -1$. Since these values also satisfy the last equation, they form a solution of the system. Hence $E = 3A - 2B - C$.

**4.22.** Suppose $u$ is a linear combination of the vectors $v_1, \ldots, v_m$ and suppose each $v_i$ is a linear combination of the vectors $w_1, \ldots, w_n$:

$$u = a_1v_1 + a_2v_2 + \cdots + a_mv_m \quad \text{and} \quad v_i = b_{i1}w_1 + b_{i2}w_2 + \cdots + b_{in}w_n$$

Show that $u$ is also a linear combination of the $w_i$. Thus if $S \subset L(T)$, then $L(S) \subset L(T)$.

$$\begin{aligned} u &= a_1v_1 + a_2v_2 + \cdots + a_mv_m \\ &= a_1(b_{11}w_1 + \cdots + b_{1n}w_n) + a_2(b_{21}w_1 + \cdots + b_{2n}w_n) + \cdots + a_m(b_{m1}w_1 + \cdots + b_{mn}w_n) \\ &= (a_1b_{11} + a_2b_{21} + \cdots + a_mb_{m1})w_1 + \cdots + (a_1b_{1n} + a_2b_{2n} + \cdots + a_mb_{mn})w_n \end{aligned}$$

or simply
$$u = \sum_{i=1}^{m} a_iv_i = \sum_{i=1}^{m} a_i\left(\sum_{j=1}^{n} b_{ij}w_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_ib_{ij}\right)w_j$$

## LINEAR SPANS, GENERATORS

**4.23.** Show that the vectors $u = (1,2,3)$, $v = (0,1,2)$ and $w = (0,0,1)$ generate $\mathbf{R}^3$.

We need to show that an arbitrary vector $(a,b,c) \in \mathbf{R}^3$ is a linear combination of $u$, $v$ and $w$.

Set $(a,b,c) = xu + yv + zw$:

$$(a, b, c) = x(1,2,3) + y(0,1,2) + z(0,0,1) = (x, 2x + y, 3x + 2y + z)$$

Then form the system of equations

$$\begin{aligned} x \qquad\quad &= a \\ 2x + y \quad &= b \\ 3x + 2y + z &= c \end{aligned} \qquad \text{or} \qquad \begin{aligned} z + 2y + 3x &= c \\ y + 2x &= b \\ x &= a \end{aligned}$$

The above system is in echelon form and is consistent; in fact $x = a$, $y = b - 2a$, $z = c - 2b + a$ is a solution. Thus $u$, $v$ and $w$ generate $\mathbf{R}^3$.

**4.24.** Find conditions on $a$, $b$ and $c$ so that $(a, b, c) \in \mathbf{R}^3$ belongs to the space generated by $u = (2, 1, 0)$, $v = (1, -1, 2)$ and $w = (0, 3, -4)$.

Set $(a, b, c)$ as a linear combination of $u$, $v$ and $w$ using unknowns $x$, $y$ and $z$: $(a, b, c) = xu + yv + zw$.

$$(a, b, c) \;=\; x(2, 1, 0) + y(1, -1, 2) + z(0, 3, -4) \;=\; (2x + y,\ x - y + 3z,\ 2y - 4z)$$

Form the equivalent system of linear equations and reduce it to echelon form:

$$\begin{aligned} 2x + y \qquad &= a \\ x - y + 3z &= b \\ 2y - 4z &= c \end{aligned} \quad \text{or} \quad \begin{aligned} 2x + y \qquad &= a \\ 3y - 6z &= a - 2b \\ 2y - 4z &= c \end{aligned} \quad \text{or} \quad \begin{aligned} 2x + y \qquad &= a \\ 3y - 6z &= a - 2b \\ 0 &= 2a - 4b - 3c \end{aligned}$$

The vector $(a, b, c)$ belongs to the space generated by $u$, $v$ and $w$ if and only if the above system is consistent, and it is consistent if and only if $2a - 4b - 3c = 0$. Note, in particular, that $u$, $v$ and $w$ do not generate the whole space $\mathbf{R}^3$.

**4.25.** Show that the $xy$ plane $W = \{(a, b, 0)\}$ in $\mathbf{R}^3$ is generated by $u$ and $v$ where: (i) $u = (1, 2, 0)$ and $v = (0, 1, 0)$; (ii) $u = (2, -1, 0)$ and $v = (1, 3, 0)$.

In each case show that an arbitrary vector $(a, b, 0) \in W$ is a linear combination of $u$ and $v$.

(i) Set $(a, b, 0) = xu + yv$:

$$(a, b, 0) \;=\; x(1, 2, 0) + y(0, 1, 0) \;=\; (x,\ 2x + y,\ 0)$$

Then form the system of equations

$$\begin{aligned} x \qquad &= a \\ 2x + y &= b \\ 0 &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} y + 2x &= b \\ x &= a \end{aligned}$$

The system is consistent; in fact $x = a$, $y = b - 2a$ is a solution. Hence $u$ and $v$ generate $W$.

(ii) Set $(a, b, 0) = xu + yv$:

$$(a, b, 0) \;=\; x(2, -1, 0) + y(1, 3, 0) \;=\; (2x + y,\ -x + 3y,\ 0)$$

Form the following system and reduce it to echelon form:

$$\begin{aligned} 2x + y &= a \\ -x + 3y &= b \\ 0 &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} 2x + y &= a \\ 7y &= a + 2b \end{aligned}$$

The system is consistent and so has a solution. Hence $W$ is generated by $u$ and $v$. (Observe that we do not need to solve for $x$ and $y$; it is only necessary to know that a solution exists.)

**4.26.** Show that the vector space $V$ of polynomials over any field $K$ cannot be generated by a finite number of vectors.

Any finite set $S$ of polynomials contains one of maximum degree, say $m$. Then the linear span $L(S)$ of $S$ cannot contain polynomials of degree greater than $m$. Accordingly, $V \neq L(S)$, for any finite set $S$.

**4.27.** 88, **Prove Theorem 4.5: Let $S$ be a nonempty subset of $V$. Then $L(S)$, the set of all linear combinations of vectors in $S$, is a subspace of $V$ containing $S$. Furthermore, if $W$ is any other subspace of $V$ containing $S$, then $L(S) \subset W$.**

Proof:- If $v \in S$, then $1v = v \in L(S)$; hence $S$ is a subset of $L(S)$. Also, $L(S)$ is nonempty since $S$ is nonempty. Now suppose $v, w \in L(S)$; say,

$$v = a_1v_1 + \cdots + a_mv_m \quad \text{and} \quad w = b_1w_1 + \cdots + b_nw_n$$

where $v_i, w_j \in S$ and $a_i, b_j$ are scalars. Then

$$v + w = a_1v_1 + \cdots + a_mv_m + b_1w_1 + \cdots + b_nw_n$$

and, for any scalar $k$,

$$kv = k(a_1v_1 + \cdots + a_mv_m) = ka_1v_1 + \cdots + ka_mv_m$$

belong to $L(S)$ since each is a linear combination of vectors in $S$. Accordingly, $L(S)$ is a subspace of $V$.

Now suppose $W$ is a subspace of $V$ containing $S$ and suppose $v_1, \ldots, v_m \in S \subset W$. Then all multiples $a_1v_1, \ldots, a_mv_m \in W$, where $a_i \in K$, and hence the sum $a_1v_1 + \cdots + a_mv_m \in W$. That is, $W$ contains all linear combinations of elements of $S$. Consequently, $L(S) \subset W$ as claimed.

## ROW SPACE OF A MATRIX

S **4.28. Determine whether the following matrices have the same row space:**

$$A = \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 13 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & -2 \\ 3 & -2 & -3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & -1 \\ 4 & -3 & -1 \\ 3 & -1 & 3 \end{pmatrix}$$

Row reduce each matrix to row canonical form:

$$A = \begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 13 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -1 & -2 \\ 3 & -2 & -3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & -1 & -1 \\ 4 & -3 & -1 \\ 3 & -1 & 3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 3 \\ 0 & 2 & 6 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Since the nonzero rows of the reduced form of $A$ and of the reduced form of $C$ are the same, $A$ and $C$ have the same row space. On the other hand, the nonzero rows of the reduced form of $B$ are not the same as the others, and so $B$ has a different row space.

**4.29. Consider an arbitrary matrix $A = (a_{ij})$. Suppose $u = (b_1, \ldots, b_n)$ is a linear combination of the rows $R_1, \ldots, R_m$ of $A$; say $u = k_1R_1 + \cdots + k_mR_m$. Show that, for each $i$, $b_i = k_1a_{1i} + k_2a_{2i} + \cdots + k_ma_{mi}$ where $a_{1i}, \ldots, a_{mi}$ are the entries of the $i$th column of $A$.**

We are given $u = k_1R_1 + \cdots + k_mR_m$; hence

$$\begin{aligned} (b_1, \ldots, b_n) &= k_1(a_{11}, \ldots, a_{1n}) + \cdots + k_m(a_{m1}, \ldots, a_{mn}) \\ &= (k_1a_{11} + \cdots + k_ma_{m1}, \ldots, k_1a_{m1} + \cdots + k_ma_{mn}) \end{aligned}$$

Setting corresponding components equal to each other, we obtain the desired result.

**4.30.** Prove: Let $A = (a_{ij})$ be an echelon matrix with distinguished entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$, and let $B = (b_{ij})$ be an echelon matrix with distinguished entries $b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$:

$$A = \begin{pmatrix} a_{1j_1} & * & * & * & * & * & * \\ & a_{2j_2} & * & * & * & * & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & a_{rj_r} & * & * & & \\ & & & & & & \end{pmatrix}, \qquad B = \begin{pmatrix} b_{1k_1} & * & * & * & * & * & * & * \\ & b_{2k_2} & * & * & * & * & * & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ & & b_{sk_s} & * & * & & & \\ & & & & & & & \end{pmatrix}$$

Suppose $A$ and $B$ have the same row space. Then the distinguished entries of $A$ and of $B$ are in the same position: $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$, and $r = s$.

Clearly $A = 0$ if and only if $B = 0$, and so we need only prove the theorem when $r \geq 1$ and $s \geq 1$. We first show that $j_1 = k_1$. Suppose $j_1 < k_1$. Then the $j_1$th column of $B$ is zero. Since the first row of $A$ is in the row space of $B$, we have by the preceding problem, $a_{1j_1} = c_1 0 + c_2 0 + \cdots + c_m 0 = 0$ for scalars $c_i$. But this contradicts the fact that the distinguished element $a_{1j_1} \neq 0$. Hence $j_1 \geq k_1$, and similarly $k_1 \geq j_1$. Thus $j_1 = k_1$.

Now let $A'$ be the submatrix of $A$ obtained by deleting the first row of $A$, and let $B'$ be the submatrix of $B$ obtained by deleting the first row of $B$. We prove that $A'$ and $B'$ have the same row space. The theorem will then follow by induction since $A'$ and $B'$ are also echelon matrices.

Let $R = (a_1, a_2, \ldots, a_n)$ be any row of $A'$ and let $R_1, \ldots, R_m$ be the rows of $B$. Since $R$ is in the row space of $B$, there exist scalars $d_1, \ldots, d_m$ such that $R = d_1R_1 + d_2R_2 + \cdots + d_mR_m$. Since $A$ is in echelon form and $R$ is not the first row of $A$, the $j_1$th entry of $R$ is zero: $a_i = 0$ for $i = j_1 = k_1$. Furthermore, since $B$ is in echelon form, all the entries in the $k_1$th column of $B$ are 0 except the first: $b_{1k_1} \neq 0$, but $b_{2k_1} = 0, \ldots, b_{mk_1} = 0$. Thus

$$0 = a_{k_1} = d_1 b_{1k_1} + d_2 0 + \cdots + d_m 0 = d_1 b_{1k_1}$$

Now $b_{1k_1} \neq 0$ and so $d_1 = 0$. Thus $R$ is a linear combination of $R_2, \ldots, R_m$ and so is in the row space of $B'$. Since $R$ was any row of $A'$, the row space of $A'$ is contained in the row space of $B'$. Similarly, the row space of $B'$ is contained in the row space of $A'$. Thus $A'$ and $B'$ have the same row space, and so the theorem is proved.

**4.31.** Prove Theorem 4.7: Let $A = (a_{ij})$ and $B = (b_{ij})$ be row reduced echelon matrices. Then $A$ and $B$ have the same row space if and only if they have the same nonzero rows.

Obviously, if $A$ and $B$ have the same nonzero rows then they have the same row space. Thus we only have to prove the converse.

Suppose $A$ and $B$ have the same row space, and suppose $R \neq 0$ is the $i$th row of $A$. Then there exist scalars $c_1, \ldots, c_s$ such that

$$R = c_1R_1 + c_2R_2 + \cdots + c_sR_s \tag{1}$$

where the $R_i$ are the nonzero rows of $B$. The theorem is proved if we show that $R = R_i$, or $c_i = 1$ but $c_k = 0$ for $k \neq i$.

Let $a_{ij_i}$ be the distinguished entry in $R$, i.e. the first nonzero entry of $R$. By (1) and Problem 4.29,

$$a_{ij_i} = c_1 b_{1j_i} + c_2 b_{2j_i} + \cdots + c_s b_{sj_i} \tag{2}$$

But by the preceding problem $b_{ij_i}$ is a distinguished entry of $B$ and, since $B$ is row reduced, it is the only nonzero entry in the $j_i$th column of $B$. Thus from (2) we obtain $a_{ij_i} = c_i b_{ij_i}$. However, $a_{ij_i} = 1$ and $b_{ij_i} = 1$ since $A$ and $B$ are row reduced; hence $c_i = 1$.

Now suppose $k \neq i$, and $b_{kj_k}$ is the distinguished entry in $R_k$. By (1) and Problem 4.29,

$$a_{ij_k} = c_1 b_{1j_k} + c_2 b_{2j_k} + \cdots + c_s b_{sj_k} \tag{3}$$

Since $B$ is row reduced, $b_{kj_k}$ is the only nonzero entry in the $j_k$th column of $B$; hence by (3), $a_{ij_k} = c_k b_{kj_k}$. Furthermore, by the preceding problem $a_{kj_k}$ is a distinguished entry of $A$ and, since $A$ is row reduced, $a_{ij_k} = 0$. Thus $c_k b_{kj_k} = 0$ and, since $b_{kj_k} = 1$, $c_k = 0$. Accordingly $R = R_i$ and the theorem is proved.

**4.32.** Determine whether the following matrices have the same column space:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 4 & 3 \\ 1 & 1 & 9 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 \\ -2 & -3 & -4 \\ 7 & 12 & 17 \end{pmatrix}$$

Observe that $A$ and $B$ have the same column space if and only if the transposes $A^t$ and $B^t$ have the same row space. Thus reduce $A^t$ and $B^t$ to row reduced echelon form:

$$A^t = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 4 & 1 \\ 5 & 3 & 9 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$B^t = \begin{pmatrix} 1 & -2 & 7 \\ 2 & -3 & 12 \\ 3 & -4 & 17 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 2 & -4 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Since $A^t$ and $B^t$ have the same row space, $A$ and $B$ have the same column space.

**4.33.** Let $R$ be a row vector and $B$ a matrix for which $RB$ is defined. Show that $RB$ is a linear combination of the rows of $B$. Furthermore, if $A$ is a matrix for which $AB$ is defined, show that the row space of $AB$ is contained in the row space of $B$.

Suppose $R = (a_1, a_2, \ldots, a_m)$ and $B = (b_{ij})$. Let $B_1, \ldots, B_m$ denote the rows of $B$ and $B^1, \ldots, B^n$ its columns. Then

$$\begin{aligned} RB &= (R \cdot B^1, R \cdot B^2, \ldots, R \cdot B^n) \\ &= (a_1 b_{11} + a_2 b_{21} + \cdots + a_m b_{m1}, a_1 b_{12} + a_2 b_{22} + \cdots + a_m b_{m2}, \ldots, a_1 b_{1n} + a_2 b_{2n} + \cdots + a_m b_{mn}) \\ &= a_1(b_{11}, b_{12}, \ldots, b_{1n}) + a_2(b_{21}, b_{22}, \ldots, b_{2n}) + \cdots + a_m(b_{m1}, b_{m2}, \ldots, b_{mn}) \\ &= a_1 B_1 + a_2 B_2 + \cdots + a_m B_m \end{aligned}$$

Thus $RB$ is a linear combination of the rows of $B$, as claimed.

By Problem 3.27, the rows of $AB$ are $R_i B$ where $R_i$ is the $i$th row of $A$. Hence by the above result each row of $AB$ is in the row space of $B$. Thus the row space of $AB$ is contained in the row space of $B$.

## SUMS AND DIRECT SUMS

**4.34.** Let $U$ and $W$ be subspaces of a vector space $V$. Show that:

(i) $U$ and $W$ are contained in $U + W$;

(ii) $U + W$ is the smallest subspace of $V$ containing $U$ and $W$, that is, $U + W$ is the linear span of $U$ and $W$: $U + W = L(U, W)$.

(i) Let $u \in U$. By hypothesis $W$ is a subspace of $V$ and so $0 \in W$. Hence $u = u + 0 \in U + W$. Accordingly, $U$ is contained in $U + W$. Similarly, $W$ is contained in $U + W$.

(ii) Since $U + W$ is a subspace of $V$ (Theorem 4.8) containing both $U$ and $W$, it must also contain the linear span of $U$ and $W$: $L(U, W) \subset U + W$.

On the other hand, if $v \in U + W$ then $v = u + w = 1u + 1w$ where $u \in U$ and $w \in W$; hence $v$ is a linear combination of elements in $U \cup W$ and so belongs to $L(U, W)$. Thus $U + W \subset L(U, W)$.

The two inclusion relations give us the required result.

**4.35.** Suppose $U$ and $W$ are subspaces of a vector space $V$, and that $\{u_i\}$ generates $U$ and $\{w_j\}$ generates $W$. Show that $\{u_i, w_j\}$, i.e. $\{u_i\} \cup \{w_j\}$, generates $U+W$.

Let $v \in U+W$. Then $v = u+w$ where $u \in U$ and $w \in W$. Since $\{u_i\}$ generates $U$, $u$ is a linear combination of $u_i$'s; and since $\{w_j\}$ generates $W$, $w$ is a linear combination of $w_j$'s:

$$u = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n}, \qquad a_j \in K$$
$$w = b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m}, \qquad b_j \in K$$

Thus
$$v = u + w = a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_n u_{i_n} + b_1 w_{j_1} + b_2 w_{j_2} + \cdots + b_m w_{j_m}$$
and so $\{u_i, w_j\}$ generates $U+W$.

**4.36.** Prove Theorem 4.9: The vector space $V$ is the direct sum of its subspaces $U$ and $W$ if and only if (i) $V = U+W$ and (ii) $U \cap W = \{0\}$.

Suppose $V = U \oplus W$. Then any $v \in V$ can be uniquely written in the form $v = u + w$ where $u \in U$ and $w \in W$. Thus, in particular, $V = U+W$. Now suppose $v \in U \cap W$. Then:

$$(1)\quad v = v + 0 \text{ where } v \in U,\ 0 \in W; \quad \text{and} \quad (2)\quad v = 0 + v \text{ where } 0 \in U,\ v \in W$$

Since such a sum for $v$ must be unique, $v = 0$. Accordingly, $U \cap W = \{0\}$.

On the other hand, suppose $V = U + W$ and $U \cap W = \{0\}$. Let $v \in V$. Since $V = U+W$, there exist $u \in U$ and $w \in W$ such that $v = u+w$. We need to show that such a sum is unique. Suppose also that $v = u' + w'$ where $u' \in U$ and $w' \in W$. Then

$$u + w = u' + w' \quad \text{and so} \quad u - u' = w' - w$$

But $u - u' \in U$ and $w' - w \in W$; hence by $U \cap W = \{0\}$,

$$u - u' = 0, \quad w' - w = 0 \quad \text{and so} \quad u = u', \quad w = w'$$

Thus such a sum for $v \in V$ is unique and $V = U \oplus W$.

**4.37.** Let $U$ and $W$ be the subspaces of $\mathbf{R}^3$ defined by

$$U = \{(a,b,c) : a = b = c\} \quad \text{and} \quad W = \{(0,b,c)\}$$

(Note that $W$ is the $yz$ plane.) Show that $\mathbf{R}^3 = U \oplus W$.

Note first that $U \cap W = \{0\}$, for $v = (a,b,c) \in U \cap W$ implies that

$$a = b = c \text{ and } a = 0 \quad \text{which implies} \quad a = 0,\ b = 0,\ c = 0$$

i.e. $v = (0,0,0)$.

We also claim that $\mathbf{R}^3 = U + W$. For if $v = (a,b,c) \in \mathbf{R}^3$, then $v = (a,a,a) + (0, b-a, c-a)$ where $(a,a,a) \in U$ and $(0, b-a, c-a) \in W$. Both conditions, $U \cap W = \{0\}$ and $\mathbf{R}^3 = U + W$, imply $\mathbf{R}^3 = U \oplus W$.

**4.38.** Let $V$ be the vector space of $n$-square matrices over a field $\mathbf{R}$. Let $U$ and $W$ be the subspaces of symmetric and antisymmetric matrices, respectively. Show that $V = U \oplus W$. (The matrix $M$ is symmetric iff $M = M^t$, and anti-symmetric iff $M^t = -M$.)

We first show that $V = U + W$. Let $A$ be any arbitrary $n$-square matrix. Note that

$$A = \tfrac{1}{2}(A + A^t) + \tfrac{1}{2}(A - A^t)$$

We claim that $\frac{1}{2}(A + A^t) \in U$ and that $\frac{1}{2}(A - A^t) \in W$. For

$$(\tfrac{1}{2}(A + A^t))^t = \tfrac{1}{2}(A + A^t)^t = \tfrac{1}{2}(A^t + A^{tt}) = \tfrac{1}{2}(A + A^t)$$

that is, $\frac{1}{2}(A + A^t)$ is symmetric. Furthermore,

$$(\tfrac{1}{2}(A - A^t))^t = \tfrac{1}{2}(A - A^t)^t = \tfrac{1}{2}(A^t - A) = -\tfrac{1}{2}(A - A^t)$$

that is, $\frac{1}{2}(A - A^t)$ is antisymmetric.

We next show that $U \cap W = \{0\}$. Suppose $M \in U \cap W$. Then $M = M^t$ and $M^t = -M$ which implies $M = -M$ or $M = 0$. Hence $U \cap W = \{0\}$. Accordingly, $V = U \oplus W$.

# Supplementary Problems

## VECTOR SPACES

**4.39.** Let $V$ be the set of infinite sequences $(a_1, a_2, \ldots)$ in a field $K$ with addition in $V$ and scalar multiplication on $V$ defined by

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$
$$k(a_1, a_2, \ldots) = (ka_1, ka_2, \ldots)$$

where $a_i, b_i, k \in K$. Show that $V$ is a vector space over $K$.

**4.40.** Let $V$ be the set of ordered pairs $(a, b)$ of real numbers with addition in $V$ and scalar multiplication on $V$ defined by

$$(a, b) + (c, d) = (a + c, b + d) \quad \text{and} \quad k(a, b) = (ka, 0)$$

Show that $V$ satisfies all of the axioms of a vector space except $[M_4]$: $1u = u$. Hence $[M_4]$ is not a consequence of the other axioms.

**4.41.** Let $V$ be the set of ordered pairs $(a, b)$ of real numbers. Show that $V$ is not a vector space over $\mathbf{R}$ with addition in $V$ and scalar multiplication on $V$ defined by:

(i) $(a, b) + (c, d) = (a + d, b + c)$ and $k(a, b) = (ka, kb)$;

(ii) $(a, b) + (c, d) = (a + c, b + d)$ and $k(a, b) = (a, b)$;

(iii) $(a, b) + (c, d) = (0, 0)$ and $k(a, b) = (ka, kb)$;

(iv) $(a, b) + (c, d) = (ac, bd)$ and $k(a, b) = (ka, kb)$.

**4.42** Let $V$ be the set of ordered pairs $(z_1, z_2)$ of complex numbers. Show that $V$ is a vector space over the real field $\mathbf{R}$ with addition in $V$ and scalar multiplication on $V$ defined by

$$(z_1, z_2) + (w_1, w_2) = (z_1 + w_1, z_2 + w_2) \quad \text{and} \quad k(z_1, z_2) = (kz_1, kz_2)$$

where $z_1, z_2, w_1, w_2 \in \mathbf{C}$ and $k \in \mathbf{R}$.

**4.43.** Let $V$ be a vector space over $K$, and let $F$ be a subfield of $K$. Show that $V$ is also a vector space over $F$ where vector addition with respect to $F$ is the same as that with respect to $K$, and where scalar multiplication by an element $k \in F$ is the same as multiplication by $k$ as an element of $K$.

**4.44.** Show that $[A_4]$, page 63, can be derived from the other axioms of a vector space.

**4.45.** Let $U$ and $W$ be vector spaces over a field $K$. Let $V$ be the set of ordered pairs $(u, w)$ where $u$ belongs to $U$ and $w$ to $W$: $V = \{(u, w) : u \in U, w \in W\}$. Show that $V$ is a vector space over $K$ with addition in $V$ and scalar multiplication on $V$ defined by

$$(u, w) + (u', w') = (u + u', w + w') \quad \text{and} \quad k(u, w) = (ku, kw)$$

where $u, u' \in U$, $w, w' \in W$ and $k \in K$. (This space $V$ is called the *external direct sum* of $U$ and $W$.)

## SUBSPACES

**4.46.** Consider the vector space $V$ in Problem 4.39, of infinite sequences $(a_1, a_2, \ldots)$ in a field $K$. Show that $W$ is a subspace of $V$ if:

(i) $W$ consists of all sequences with 0 as the first component;

(ii) $W$ consists of all sequences with only a finite number of nonzero components.

**4.47.** Determine whether or not $W$ is a subspace of $\mathbf{R}^3$ if $W$ consists of those vectors $(a, b, c) \in \mathbf{R}^3$ for which: (i) $a = 2b$; (ii) $a \leq b \leq c$; (iii) $ab = 0$; (iv) $a = b = c$; (v) $a = b^2$; (vi) $k_1a + k_2b + k_3c = 0$, where $k_i \in \mathbf{R}$.

**4.48.** Let $V$ be the vector space of $n$-square matrices over a field $K$. Show that $W$ is a subspace of $V$ if $W$ consists of all matrices which are (i) antisymmetric ($A^t = -A$), (ii) (upper) triangular, (iii) diagonal, (iv) scalar.

**4.49.** Let $AX = B$ be a nonhomogeneous system of linear equations in $n$ unknowns over a field $K$. Show that the solution set of the system is not a subspace of $K^n$.

**4.50.** Let $V$ be the vector space of all functions from the real field $\mathbf{R}$ into $\mathbf{R}$. Show that $W$ is a subspace of $V$ in each of the following cases.

(i) $W$ consists of all bounded functions. (Here $f: \mathbf{R} \to \mathbf{R}$ is bounded if there exists $M \in \mathbf{R}$ such that $|f(x)| \leq M$, $\forall x \in \mathbf{R}$.)

(ii) $W$ consists of all even functions. (Here $f: \mathbf{R} \to \mathbf{R}$ is even if $f(-x) = f(x)$, $\forall x \in \mathbf{R}$.)

(iii) $W$ consists of all continuous functions.

(iv) $W$ consists of all differentiable functions.

(v) $W$ consists of all integrable functions in, say, the interval $0 \leq x \leq 1$.

(The last three cases require some knowledge of analysis.)

**4.51.** Discuss whether or not $\mathbf{R}^2$ is a subspace of $\mathbf{R}^3$.

**4.52.** Prove Theorem 4.4: The intersection of any number of subspaces of a vector space $V$ is a subspace of $V$.

**4.53.** Suppose $U$ and $W$ are subspaces of $V$ for which $U \cup W$ is also a subspace. Show that either $U \subset W$ or $W \subset U$.

## LINEAR COMBINATIONS

**4.54.** Consider the vectors $u = (1, -3, 2)$ and $v = (2, -1, 1)$ in $\mathbf{R}^3$.

(i) Write $(1, 7, -4)$ as a linear combination of $u$ and $v$.

(ii) Write $(2, -5, 4)$ as a linear combination of $u$ and $v$.

(iii) For which value of $k$ is $(1, k, 5)$ a linear combination of $u$ and $v$?

(iv) Find a condition on $a$, $b$ and $c$ so that $(a, b, c)$ is a linear combination of $u$ and $v$.

**4.55.** Write $u$ as a linear combination of the polynomials $v = 2t^2 + 3t - 4$ and $w = t^2 - 2t - 3$ where (i) $u = 3t^2 + 8t - 5$, (ii) $u = 4t^2 - 6t - 1$.

**4.56.** Write $E$ as a linear combination of $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ and $C = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ where: (i) $E = \begin{pmatrix} 3 & -1 \\ 1 & -2 \end{pmatrix}$; (ii) $E = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$.

## LINEAR SPANS, GENERATORS

**4.57.** Show that $(1, 1, 1)$, $(0, 1, 1)$ and $(0, 1, -1)$ generate $\mathbf{R}^3$, i.e. that any vector $(a, b, c)$ is a linear combination of the given vectors.

**4.58.** Show that the $yz$ plane $W = \{(0, b, c)\}$ in $\mathbf{R}^3$ is generated by: (i) $(0, 1, 1)$ and $(0, 2, -1)$; (ii) $(0, 1, 2)$, $(0, 2, 3)$ and $(0, 3, 1)$.

**4.59.** Show that the complex numbers $w = 2 + 3i$ and $z = 1 - 2i$ generate the complex field $\mathbf{C}$ as a vector space over the real field $\mathbf{R}$.

**4.60.** Show that the polynomials $(1-t)^3$, $(1-t)^2$, $1-t$ and $1$ generate the space of polynomials of degree $\leq 3$.

**4.61.** Find one vector in $\mathbf{R}^3$ which generates the intersection of $U$ and $W$ where $U$ is the $xy$ plane: $U = \{(a, b, 0)\}$, and $W$ is the space generated by the vectors $(1, 2, 3)$ and $(1, -1, 1)$.

**4.62.** Prove: $L(S)$ is the intersection of all the subspaces of $V$ containing $S$.

**4.63.** Show that $L(S) = L(S \cup \{0\})$. That is, by joining or deleting the zero vector from a set, we do not change the space generated by the set.

**4.64.** Show that if $S \subset T$, then $L(S) \subset L(T)$.

**4.65.** Show that $L(L(S)) = L(S)$.

## ROW SPACE OF A MATRIX

**4.66.** Determine which of the following matrices have the same row space:

$$A = \begin{pmatrix} 1 & -2 & -1 \\ 3 & -4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -1 & 10 \\ 3 & -5 & 1 \end{pmatrix}$$

**4.67.** Let

$$u_1 = (1, 1, -1), \quad u_2 = (2, 3, -1), \quad u_3 = (3, 1, -5)$$
$$v_1 = (1, -1, -3), \quad v_2 = (3, -2, -8), \quad v_3 = (2, 1, -3)$$

Show that the subspace of $\mathbf{R}^3$ generated by the $u_i$ is the same as the subspace generated by the $v_i$.

**4.68.** Show that if any row of an echelon (row reduced echelon) matrix is deleted, then the resulting matrix is still in echelon (row reduced echelon) form.

**4.69.** Prove the converse of Theorem 4.6: Matrices with the same row space (and the same size) are row equivalent.

**4.70.** Show that $A$ and $B$ have the same column space iff $A^t$ and $B^t$ have the same row space.

**4.71.** Let $A$ and $B$ be matrices for which $AB$ is defined. Show that the column space of $AB$ is contained in the column space of $A$.

## SUMS AND DIRECT SUMS

**4.72.** We extend the notion of sum to arbitrary nonempty subsets (not necessarily subspaces) $S$ and $T$ of a vector space $V$ by defining $S + T = \{s + t : s \in S, t \in T\}$. Show that this operation satisfies:

(i) commutative law: $S + T = T + S$;

(ii) associative law: $(S_1 + S_2) + S_3 = S_1 + (S_2 + S_3)$;

(iii) $S + \{0\} = \{0\} + S = S$;

(iv) $S + V = V + S = V$.

**4.73.** Show that for any subspace $W$ of a vector space $V$, $W + W = W$.

**4.74.** Give an example of a subset $S$ of a vector space $V$ which is not a subspace of $V$ but for which (i) $S + S = S$, (ii) $S + S \subset S$ (properly contained).

**4.75.** We extend the notion of sum of subspaces to more than two summands as follows. If $W_1, W_2, \ldots, W_n$ are subspaces of $V$, then

$$W_1 + W_2 + \cdots + W_n = \{w_1 + w_2 + \cdots + w_n : w_i \in W_i\}$$

Show that:

(i) $L(W_1, W_2, \ldots, W_n) = W_1 + W_2 + \cdots + W_n$;

(ii) if $S_i$ generates $W_i$, $i = 1, \ldots, n$, then $S_1 \cup S_2 \cup \cdots \cup S_n$ generates $W_1 + W_2 + \cdots + W_n$.

**4.76.** Suppose $U$, $V$ and $W$ are subspaces of a vector space. Prove that

$$(U \cap V) + (U \cap W) \subset U \cap (V + W)$$

Find subspaces of $\mathbf{R}^2$ for which equality does not hold.

**4.77.** Let $U$, $V$ and $W$ be the following subspaces of $\mathbf{R}^3$:

$$U = \{(a,b,c): a+b+c=0\}, \quad V = \{(a,b,c): a=c\}, \quad W = \{(0,0,c): c \in \mathbf{R}\}$$

Show that (i) $\mathbf{R}^3 = U + V$, (ii) $\mathbf{R}^3 = U + W$, (iii) $\mathbf{R}^3 = V + W$. When is the sum direct?

**4.78.** Let $V$ be the vector space of all functions from the real field $\mathbf{R}$ into $\mathbf{R}$. Let $U$ be the subspace of even functions and $W$ the subspace of odd functions. Show that $V = U \oplus W$. (Recall that $f$ is even iff $f(-x) = f(x)$, and $f$ is odd iff $f(-x) = -f(x)$.)

**4.79.** Let $W_1, W_2, \ldots$ be subspaces of a vector space $V$ for which $W_1 \subset W_2 \subset \cdots$. Let $W = W_1 \cup W_2 \cup \cdots$. Show that $W$ is a subspace of $V$.

**4.80.** In the preceding problem, suppose $S_i$ generates $W_i$, $i = 1, 2, \ldots$. Show that $S = S_1 \cup S_2 \cup \cdots$ generates $W$.

85, 89 **4.81.** Let $V$ be the vector space of $n$-square matrices over a field $K$. Let $U$ be the subspace of upper triangular matrices and $W$ the subspace of lower triangular matrices. Find (i) $U + W$, (ii) $U \cap W$.

**4.82.** Let $V$ be the external direct sum of the vector spaces $U$ and $W$ over a field $K$. (See Problem 4.45.) Let

$$\hat{U} = \{(u, 0): u \in U\}, \quad \hat{W} = \{(0, w): w \in W\}$$

Show that (i) $\hat{U}$ and $\hat{W}$ are subspaces of $V$, (ii) $V = \hat{U} \oplus \hat{W}$.

## Answers to Supplementary Problems

**4.47.** (i) Yes.
(ii) No; e.g. $(1,2,3) \in W$ but $-2(1,2,3) \notin W$.
(iii) No; e.g. $(1,0,0), (0,1,0) \in W$, but not their sum.
(iv) Yes.
(v) No; e.g. $(9,3,0) \in W$ but $2(9,3,0) \notin W$.
(vi) Yes.

**4.50.** (i) Let $f, g \in W$ with $M_f$ and $M_g$ bounds for $f$ and $g$ respectively. Then for any scalars $a, b \in \mathbf{R}$,

$$|(af+bg)(x)| = |af(x) + bg(x)| \le |af(x)| + |bg(x)| = |a|\,|f(x)| + |b|\,|g(x)| \le |a|M_f + |b|M_g$$

That is, $|a|M_f + |b|M_g$ is a bound for the function $af + bg$.

(ii) $(af+bg)(-x) = af(-x) + bg(-x) = af(x) + bg(x) = (af+bg)(x)$

**4.51.** No. Although one may "identify" the vector $(a,b) \in \mathbf{R}^2$ with, say, $(a,b,0)$ in the $xy$ plane in $\mathbf{R}^3$, they are distinct elements belonging to distinct, disjoint sets.

**4.54.** (i) $-3u + 2v$. (ii) Impossible. (iii) $k = -8$. (iv) $a - 3b - 5c = 0$.

**4.55.** (i) $u = 2v - w$. (ii) Impossible.

**4.56.** (i) $E = 2A - B + 2C$. (ii) Impossible.

**4.61.** $(2, -5, 0)$.

**4.66.** $A$ and $C$.

**4.67.** Form the matrix $A$ whose rows are the $u_i$ and the matrix $B$ whose rows are the $v_i$, and then show that $A$ and $B$ have the same row canonical forms.

**4.74.** (i) In $\mathbf{R}^2$, let $S = \{(0,0), (0,1), (0,2), (0,3), \ldots\}$.
(ii) In $\mathbf{R}^2$, let $S = \{(0,5), (0,6), (0,7), \ldots\}$.

**4.77.** The sum is direct in (ii) and (iii).

**4.78.** *Hint.* $f(x) = \frac{1}{2}(f(x) + f(-x)) + \frac{1}{2}(f(x) - f(-x))$, where $\frac{1}{2}(f(x) + f(-x))$ is even and $\frac{1}{2}(f(x) - f(-x))$ is odd.

**4.81.** (i) $V = U + W$. (ii) $U \cap W$ is the space of diagonal matrices.

# Chapter 5

## Basis and Dimension

### INTRODUCTION

Some of the fundamental results proven in this chapter are:

(i) The "dimension" of a vector space is well defined (Theorem 5.3).

(ii) If $V$ has dimension $n$ over $K$, then $V$ is "isomorphic" to $K^n$ (Theorem 5.12).

(iii) A system of linear equations has a solution if and only if the coefficient and augmented matrices have the same "rank" (Theorem 5.10).

These concepts and results are nontrivial and answer certain questions raised and investigated by mathematicians of yesterday.

We will begin the chapter with the definition of linear dependence and independence. This concept plays an essential role in the theory of linear algebra and in mathematics in general.

### LINEAR DEPENDENCE

**Definition:** Let $V$ be a vector space over a field $K$. The vectors $v_1, \ldots, v_m \in V$ are said to be *linearly dependent over K*, or simply *dependent*, if there exist scalars $a_1, \ldots, a_m \in K$, not all of them 0, such that

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0 \qquad (*)$$

Otherwise, the vectors are said to be *linearly independent over K*, or simply *independent*.

Observe that the relation $(*)$ will always hold if the $a$'s are all 0. If this relation holds only in this case, that is,

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0 \quad \text{only if} \quad a_1 = 0, \ldots, a_m = 0$$

then the vectors are linearly independent. On the other hand, if the relation $(*)$ also holds when one of the $a$'s is not 0, then the vectors are linearly dependent.

Observe that if 0 is one of the vectors $v_1, \ldots, v_m$, say $v_1 = 0$, then the vectors must be dependent; for

$$1v_1 + 0v_2 + \cdots + 0v_m = 1\cdot 0 + 0 + \cdots + 0 = 0$$

and the coefficient of $v_1$ is not 0. On the other hand, any nonzero vector $v$ is, by itself, independent; for

$$kv = 0, \ v \neq 0 \quad \text{implies} \quad k = 0$$

Other examples of dependent and independent vectors follow.

**Example 5.1:** The vectors $u = (1, -1, 0)$, $v = (1, 3, -1)$ and $w = (5, 3, -2)$ are dependent since, for $3u + 2v - w = 0$,

$$3(1, -1, 0) + 2(1, 3, -1) - (5, 3, -2) = (0, 0, 0)$$

**Example 5.2:** We show that the vectors $u = (6, 2, 3, 4)$, $v = (0, 5, -3, 1)$ and $w = (0, 0, 7, -2)$ are independent. For suppose $xu + yv + zw = 0$ where $x$, $y$ and $z$ are unknown scalars. Then

$$\begin{aligned}(0, 0, 0, 0) &= x(6, 2, 3, 4) + y(0, 5, -3, 1) + z(0, 0, 7, -2) \\ &= (6x,\ 2x + 5y,\ 3x - 3y + 7z,\ 4x + y - 2z)\end{aligned}$$

and so, by the equality of the corresponding components,

$$\begin{aligned}6x \qquad\qquad &= 0 \\ 2x + 5y \qquad &= 0 \\ 3x - 3y + 7z &= 0 \\ 4x + y - 2z &= 0\end{aligned}$$

The first equation yields $x = 0$; the second equation with $x = 0$ yields $y = 0$, and the third equation with $x = 0$, $y = 0$ yields $z = 0$. Thus

$$xu + yv + zw = 0 \quad \text{implies} \quad x = 0,\ y = 0,\ z = 0$$

Accordingly $u$, $v$ and $w$ are independent.

Observe that the vectors in the preceding example form a matrix in echelon form:

$$\begin{pmatrix} 6 & 2 & 3 & 4 \\ 0 & 5 & -3 & 1 \\ 0 & 0 & 7 & -2 \end{pmatrix}$$

Thus we have shown that the (nonzero) rows of the above echelon matrix are independent. This result holds true in general; we state it formally as a theorem since it will be frequently used.

**Theorem 5.1:** The nonzero rows of a matrix in echelon form are linearly independent.

For more than one vector, the concept of dependence can be defined equivalently as follows:

The vectors $v_1, \ldots, v_m$ are linearly dependent if and only if one of them is a linear combination of the others.

For suppose, say, $v_i$ is a linear combination of the others:

$$v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_mv_m$$

Then by adding $-v_i$ to both sides, we obtain

$$a_1v_1 + \cdots + a_{i-1}v_{i-1} - v_i + a_{i+1}v_{i+1} + \cdots + a_mv_m = 0$$

where the coefficient of $v_i$ is not 0; hence the vectors are linearly dependent. Conversely, suppose the vectors are linearly dependent, say,

$$b_1v_1 + \cdots + b_jv_j + \cdots + b_mv_m = 0 \qquad \text{where} \quad b_j \neq 0$$

Then $\quad v_j = -b_j^{-1}b_1v_1 - \cdots - b_j^{-1}b_{j-1}v_{j-1} - b_j^{-1}b_{j+1}v_{j+1} - \cdots - b_j^{-1}b_mv_m$

and so $v_j$ is a linear combination of the other vectors.

We now make a slightly stronger statement than that above; this result has many important consequences.

**Lemma 5.2:** The nonzero vectors $v_1, \ldots, v_m$ are linearly dependent if and only if one of them, say $v_i$, is a linear combination of the preceding vectors:

$$v_i = k_1v_1 + k_2v_2 + \cdots + k_{i-1}v_{i-1}$$

**Remark 1.** The set $\{v_1, \ldots, v_m\}$ is called a *dependent* or *independent set* according as the vectors $v_1, \ldots, v_m$ are dependent or independent. We also define the empty set $\emptyset$ to be independent.

**Remark 2.** If two of the vectors $v_1, \ldots, v_m$ are equal, say $v_1 = v_2$, then the vectors are dependent. For

$$v_1 - v_2 + 0v_3 + \cdots + 0v_m = 0$$

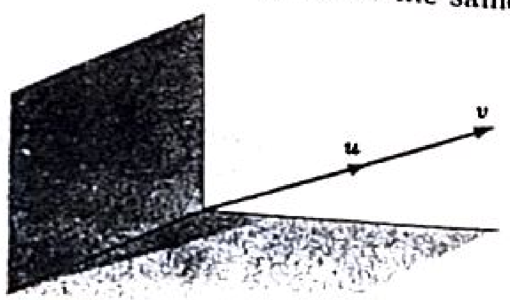
and the coefficient of $v_1$ is not 0.

**Remark 3.** Two vectors $v_1$ and $v_2$ are dependent if and only if one of them is a multiple of the other.

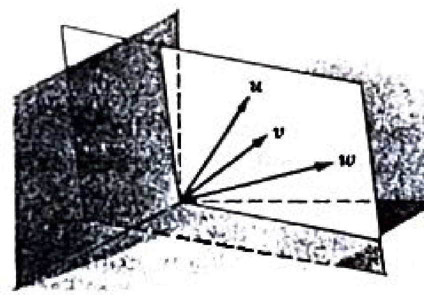
**Remark 4.** A set which contains a dependent subset is itself dependent. Hence any subset of an independent set is independent.

**Remark 5.** If the set $\{v_1, \ldots, v_m\}$ is independent, then any rearrangement of the vectors $\{v_{i_1}, v_{i_2}, \ldots, v_{i_m}\}$ is also independent.

**Remark 6.** In the real space $\mathbf{R}^3$, dependence of vectors can be described geometrically as follows: any two vectors $u$ and $v$ are dependent if and only if they lie on the same line through the origin; and any three vectors $u$, $v$ and $w$ are dependent if and only if they lie on the same plane through the origin:




$u$ and $v$ are dependent.




$u$, $v$ and $w$ are dependent.

## BASIS AND DIMENSION

We begin with a definition.

94, 86 **Definition:** A vector space $V$ is said to be of *finite dimension n* or to be *n-dimensional*, written $\dim V = n$, if there exists linearly independent vectors $e_1, e_2, \ldots, e_n$ which span $V$. The sequence $\{e_1, e_2, \ldots, e_n\}$ is then called a *basis* of $V$.

The above definition of dimension is well defined in view of the following theorem.

88, 84, 93 **Theorem 5.3:** Let $V$ be a finite dimensional vector space. Then every basis of $V$ has the same number of elements.

The vector space $\{0\}$ is defined to have dimension 0. (In a certain sense this agrees with the above definition since, by definition, $\emptyset$ is independent and generates $\{0\}$.) When a vector space is not of finite dimension, it is said to be of *infinite dimension*.

**Example 5.3:** Let $K$ be any field. Consider the vector space $K^n$ which consists of $n$-tuples of elements of $K$. The vectors

$$e_1 = (1, 0, 0, \ldots, 0, 0)$$
$$e_2 = (0, 1, 0, \ldots, 0, 0)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$e_n = (0, 0, 0, \ldots, 0, 1)$$

form a basis, called the *usual basis*, of $K^n$. Thus $K^n$ has dimension $n$.

**Example 5.4:** Let $U$ be the vector space of all $2 \times 3$ matrices over a field $K$. Then the matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

form a basis of $U$. Thus $\dim U = 6$. More generally, let $V$ be the vector space of all $m \times n$ matrices over $K$ and let $E_{ij} \in V$ be the matrix with $ij$-entry 1 and 0 elsewhere. Then the set $\{E_{ij}\}$ is a basis, called the *usual basis*, of $V$ (Problem 5.32); consequently $\dim V = mn$.

**Example 5.5:** Let $W$ be the vector space of polynomials (in $t$) of degree $\leq n$. The set $\{1, t, t^2, \ldots, t^n\}$ is linearly independent and generates $W$. Thus it is a basis of $W$ and so $\dim W = n + 1$.

We comment that the vector space $V$ of all polynomials is not finite dimensional since (Problem 4.26) no finite set of polynomials generates $V$.

The above fundamental theorem on dimension is a consequence of the following important "replacement lemma":

**Lemma 5.4:** Suppose the set $\{v_1, v_2, \ldots, v_n\}$ generates a vector space $V$. If $\{w_1, \ldots, w_m\}$ is linearly independent, then $m \leq n$ and $V$ is generated by a set of the form

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Thus, in particular, any $n+1$ or more vectors in $V$ are linearly dependent.

Observe in the above lemma that we have replaced $m$ of the vectors in the generating set by the $m$ independent vectors and still retained a generating set.

Now suppose $S$ is a subset of a vector space $V$. We call $\{v_1, \ldots, v_m\}$ a *maximal independent subset* of $S$ if:

(i) it is an independent subset of $S$; and

(ii) $\{v_1, \ldots, v_m, w\}$ is dependent for any $w \in S$.

The following theorem applies.

**Theorem 5.5:** Suppose $S$ generates $V$ and $\{v_1, \ldots, v_m\}$ is a maximal independent subset of $S$. Then $\{v_1, \ldots, v_m\}$ is a basis of $V$.

The main relationship between the dimension of a vector space and its independent subsets is contained in the next theorem.

**Theorem 5.6:** Let $V$ be of finite dimension $n$. Then:

(i) Any set of $n+1$ or more vectors is linearly dependent.

(ii) Any linearly independent set is part of a basis, i.e. can be extended to a basis.

(iii) A linearly independent set with $n$ elements is a basis.

**Example 5.6:** The four vectors in $K^4$

$$(1, 1, 1, 1), \quad (0, 1, 1, 1), \quad (0, 0, 1, 1), \quad (0, 0, 0, 1)$$

are linearly independent since they form a matrix in echelon form. Furthermore, since $\dim K^4 = 4$, they form a basis of $K^4$.

**Example 5.7:** The four vectors in $\mathbf{R}^3$,

$$(257, -132, 58), \quad (43, 0, -17), \quad (521, -317, 94), \quad (328, -512, -731)$$

must be linearly dependent since they come from a vector space of dimension 3.

## DIMENSION AND SUBSPACES

The following theorems give basic relationships between the dimension of a vector space and the dimension of a subspace.

**Theorem 5.7:** Let $W$ be a subspace of an $n$-dimensional vector space $V$. Then $\dim W \leq n$. In particular if $\dim W = n$, then $W = V$.

**Example 5.8:** Let $W$ be a subspace of the real space $\mathbf{R}^3$. Now $\dim \mathbf{R}^3 = 3$, hence by the preceding theorem the dimension of $W$ can only be 0, 1, 2 or 3. The following cases apply:

(i) $\dim W = 0$, then $W = \{0\}$, a point;

(ii) $\dim W = 1$, then $W$ is a line through the origin;

(iii) $\dim W = 2$, then $W$ is a plane through the origin;

(iv) $\dim W = 3$, then $W$ is the entire space $\mathbf{R}^3$.

**Theorem 5.8:** Let $U$ and $W$ be finite-dimensional subspaces of a vector space $V$. Then $U + W$ has finite dimension and

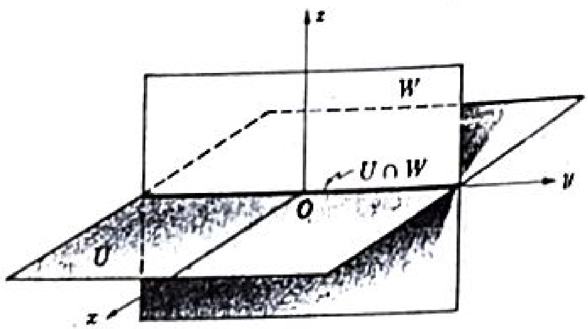
$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W)$$

Note that if $V$ is the direct sum of $U$ and $W$, i.e. $V = U \oplus W$, then $\dim V = \dim U + \dim W$ (Problem 5.48).

**Example 5.9:** Suppose $U$ and $W$ are the $xy$ plane and $yz$ plane, respectively, in $\mathbf{R}^3$: $U = \{(a, b, 0)\}$, $W = \{(0, b, c)\}$. Since $\mathbf{R}^3 = U + W$, $\dim(U + W) = 3$. Also, $\dim U = 2$ and $\dim W = 2$. By the above theorem,

$$3 = 2 + 2 - \dim(U \cap W) \quad \text{or} \quad \dim(U \cap W) = 1$$

Observe that this agrees with the fact that $U \cap W$ is the $y$ axis, i.e. $U \cap W = \{(0, b, 0)\}$, and so has dimension 1.




## 86,92 RANK OF A MATRIX

Let $A$ be an arbitrary $m \times n$ matrix over a field $K$. Recall that the row space of $A$ is the subspace of $K^n$ generated by its rows, and the column space of $A$ is the subspace of $K^m$ generated by its columns. The dimensions of the row space and of the column space of $A$ are called, respectively, the *row rank* and the *column rank* of $A$.

88,84 **Theorem 5.9:** The row rank and the column rank of the matrix $A$ are equal.

**Definition:** The *rank* of the matrix $A$, written rank$(A)$, is the common value of its row rank and column rank.
88

Thus the rank of a matrix gives the maximum number of independent rows, and also the maximum number of independent columns. We can obtain the rank of a matrix as follows.

Suppose $A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{pmatrix}$. We reduce $A$ to echelon form using the elementary row operations:

$$A \quad \text{to} \quad \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & -6 & -2 \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Recall that row equivalent matrices have the same row space. Thus the nonzero rows of the echelon matrix, which are independent by Theorem 5.1, form a basis of the row space of $A$. Hence the rank of $A$ is 2.

## APPLICATIONS TO LINEAR EQUATIONS

Consider a system of $m$ linear equations in $n$ unknowns $x_1, \ldots, x_n$ over a field $K$:

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

or the equivalent matrix equation

$$AX = B$$

where $A = (a_{ij})$ is the coefficient matrix, and $X = (x_i)$ and $B = (b_i)$ are the column vectors consisting of the unknowns and of the constants, respectively. Recall that the *augmented matrix* of the system is defined to be the matrix

$$(A, B) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

**Remark 1.** The above linear equations are said to be dependent or independent according as the corresponding vectors, i.e. the rows of the augmented matrix, are dependent or independent.

**Remark 2.** Two systems of linear equations are equivalent if and only if the corresponding augmented matrices are row equivalent, i.e. have the same row space.

**Remark 3.** We can always replace a system of equations by a system of independent equations, such as a system in echelon form. The number of independent equations will always be equal to the rank of the augmented matrix.

Observe that the above system is also equivalent to the vector equation

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot\cdot \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot\cdot \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot\cdot \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot\cdot \\ b_m \end{pmatrix}$$

Thus the system $AX = B$ has a solution if and only if the column vector $B$ is a linear combination of the columns of the matrix $A$, i.e. belongs to the column space of $A$. This gives us the following basic existence theorem.

**Theorem 5.10:** The system of linear equations $AX = B$ has a solution if and only if the coefficient matrix $A$ and the augmented matrix $(A, B)$ have the same rank.

Recall (Theorem 2.1) that if the system $AX = B$ does have a solution, say $v$, then its general solution is of the form $v + W = \{v + w : w \in W\}$ where $W$ is the general solution of the associated homogeneous system $AX = 0$. Now $W$ is a subspace of $K^n$ and so has a dimension. The next theorem, whose proof is postponed until the next chapter (page 127), applies.

**Theorem 5.11:** The dimension of the solution space $W$ of the homogeneous system of linear equations $AX = 0$ is $n - r$ where $n$ is the number of unknowns and $r$ is the rank of the coefficient matrix $A$.

In case the system $AX = 0$ is in echelon form, then it has precisely $n - r$ free variables (see page 21), say, $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-r}}$. Let $v_j$ be the solution obtained by setting $x_{i_j} = 1$, and all other free variables $= 0$. Then the solutions $v_1, \ldots, v_{n-r}$ are linearly independent (Problem 5.43) and so form a basis for the solution space.

**Example 5.10:** Find the dimension and a basis of the solution space $W$ of the system of linear equations

$$\begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ x + 2y - 2z + 2r + s &= 0 \\ 2x + 4y - 2z + 3r + 4s &= 0 \end{aligned}$$

Reduce the system to echelon form:

$$\begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ 2z - r + 2s &= 0 \\ 6z - 3r + 6s &= 0 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x + 2y - 4z + 3r - s &= 0 \\ 2z - r + 2s &= 0 \end{aligned}$$

There are 5 unknowns and 2 (nonzero) equations in echelon form; hence $\dim W = 5 - 2 = 3$. Note that the free variables are $y, r$ and $s$. Set:

(i) $y = 1, r = 0, s = 0$, (ii) $y = 0, r = 1, s = 0$, (iii) $y = 0, r = 0, s = 1$

to obtain the following respective solutions:

$$v_1 = (-2, 1, 0, 0, 0), \quad v_2 = (-1, 0, \tfrac{1}{2}, 1, 0), \quad v_3 = (-3, 0, -1, 0, 1)$$

The set $\{v_1, v_2, v_3\}$ is a basis of the solution space $W$.

## COORDINATES

Let $\{e_1, \ldots, e_n\}$ be a basis of an $n$-dimensional vector space $V$ over a field $K$, and let $v$ be any vector in $V$. Since $\{e_i\}$ generates $V$, $v$ is a linear combination of the $e_i$:

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n, \qquad a_i \in K$$

Since the $e_i$ are independent, such a representation is unique (Problem 5.7), i.e. the $n$ scalars $a_1, \ldots, a_n$ are completely determined by the vector $v$ and the basis $\{e_i\}$. We call these scalars the *coordinates* of $v$ in $\{e_i\}$, and we call the $n$-tuple $(a_1, \ldots, a_n)$ the *coordinate vector* of $v$ relative to $\{e_i\}$ and denote it by $[v]_e$ or simply $[v]$:

$$[v]_e = (a_1, a_2, \ldots, a_n)$$

**Example 5.11:** Let $V$ be the vector space of polynomials with degree $\leq 2$:

$$V = \{at^2 + bt + c : a, b, c \in \mathbf{R}\}$$

The polynomials

$$e_1 = 1, \quad e_2 = t - 1 \quad \text{and} \quad e_3 = (t-1)^2 = t^2 - 2t + 1$$

form a basis for $V$. Let $v = 2t^2 - 5t + 6$. Find $[v]_e$, the coordinate vector of $v$ relative to the basis $\{e_1, e_2, e_3\}$.

Set $v$ as a linear combination of the $e_i$ using the unknowns $x$, $y$ and $z$: $v = xe_1 + ye_2 + ze_3$.

$$\begin{aligned} 2t^2 - 5t + 6 &= x(1) + y(t-1) + z(t^2 - 2t + 1) \\ &= x + yt - y + zt^2 - 2zt + z \\ &= zt^2 + (y - 2z)t + (x - y + z) \end{aligned}$$

Then set the coefficients of the same powers of $t$ equal to each other:

$$\begin{aligned} x - y + z &= 6 \\ y - 2z &= -5 \\ z &= 2 \end{aligned}$$

The solution of the above system is $x = 3$, $y = -1$, $z = 2$. Thus

$$v = 3e_1 - e_2 + 2e_3, \quad \text{and so} \quad [v]_e = (3, -1, 2)$$

**Example 5.12:** Consider the real space $\mathbf{R}^3$. Find the coordinate vector of $v = (3, 1, -4)$ relative to the basis $f_1 = (1, 1, 1)$, $f_2 = (0, 1, 1)$, $f_3 = (0, 0, 1)$.

Set $v$ as a linear combination of the $f_i$ using the unknowns $x$, $y$ and $z$: $v = xf_1 + yf_2 + zf_3$.

$$\begin{aligned} (3, 1, -4) &= x(1, 1, 1) + y(0, 1, 1) + z(0, 0, 1) \\ &= (x, x, x) + (0, y, y) + (0, 0, z) \\ &= (x, x + y, x + y + z) \end{aligned}$$

Then set the corresponding components equal to each other to obtain the equivalent system of equations

$$\begin{aligned} x \qquad &= 3 \\ x + y \quad &= 1 \\ x + y + z &= -4 \end{aligned}$$

having solution $x = 3$, $y = -2$, $z = -5$. Thus $[v]_f = (3, -2, -5)$.

We remark that relative to the usual basis $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, the coordinate vector of $v$ is identical to $v$ itself: $[v]_e = (3, 1, -4) = v$.

We have shown above that to each vector $v \in V$ there corresponds, relative to a given basis $\{e_1, \ldots, e_n\}$, an $n$-tuple $[v]_e$ in $K^n$. On the other hand, if $(a_1, \ldots, a_n) \in K^n$, then there exists a vector in $V$ of the form $a_1e_1 + \cdots + a_ne_n$. Thus the basis $\{e_i\}$ determines a one-to-one correspondence between the vectors in $V$ and the $n$-tuples in $K^n$. Observe also that if

$$v = a_1e_1 + \cdots + a_ne_n \quad \text{corresponds to} \quad (a_1, \ldots, a_n)$$

and

$$w = b_1e_1 + \cdots + b_ne_n \quad \text{corresponds to} \quad (b_1, \ldots, b_n)$$

then

$$v + w = (a_1 + b_1)e_1 + \cdots + (a_n + b_n)e_n \quad \text{corresponds to} \quad (a_1, \ldots, a_n) + (b_1, \ldots, b_n)$$

and, for any scalar $k \in K$,

$$kv = (ka_1)e_1 + \cdots + (ka_n)e_n \quad \text{corresponds to} \quad k(a_1, \ldots, a_n)$$

That is,

$$[v + w]_e = [v]_e + [w]_e \quad \text{and} \quad [kv]_e = k[v]_e$$

Thus the above one-to-one correspondence between $V$ and $K^n$ preserves the vector space operations of vector addition and scalar multiplication; we then say that $V$ and $K^n$ are *isomorphic*, written $V \cong K^n$. We state this result formally.

**Theorem 5.12:** Let $V$ be an $n$-dimensional vector space over a field $K$. Then $V$ and $K^n$ are isomorphic.

The next example gives a practical application of the above result.

**Example 5.13:** Determine whether the following matrices are dependent or independent

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & -4 \\ 6 & 5 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 8 & -11 \\ 16 & 10 & 9 \end{pmatrix}$$

The coordinate vectors of the above matrices relative to the basis in Example 5.4, page 89, are

$$[A] = (1, 2, -3, 4, 0, 1), \quad [B] = (1, 3, -4, 6, 5, 4), \quad [C] = (3, 8, -11, 16, 10, 9)$$

Form the matrix $M$ whose rows are the above coordinate vectors

$$M = \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 1 & 3 & -4 & 6 & 5 & 4 \\ 3 & 8 & -11 & 16 & 10 & 9 \end{pmatrix}$$

Row reduce $M$ to echelon form:

$$M \text{ to } \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 2 & -2 & 4 & 10 & 6 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -3 & 4 & 0 & 1 \\ 0 & 1 & -1 & 2 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the echelon matrix has only two nonzero rows, the coordinate vectors $[A]$, $[B]$ and $[C]$ generate a space of dimension 2 and so are dependent. Accordingly, the original matrices $A$, $B$ and $C$ are dependent.

# Solved Problems

## LINEAR DEPENDENCE

**5.1.** Determine whether or not $u$ and $v$ are linearly dependent if:

(i) $u = (3, 4)$, $v = (1, -3)$

(ii) $u = (2, -3)$, $v = (6, -9)$

(iii) $u = (4, 3, -2)$, $v = (2, -6, 7)$

(iv) $u = (-4, 6, -2)$, $v = (2, -3, 1)$

(v) $u = \begin{pmatrix} 1 & -2 & 4 \\ 3 & 0 & -1 \end{pmatrix}$, $v = \begin{pmatrix} 2 & -4 & 8 \\ 6 & 0 & -2 \end{pmatrix}$

(vi) $u = \begin{pmatrix} 1 & 2 & -3 \\ 6 & -5 & 4 \end{pmatrix}$, $v = \begin{pmatrix} 6 & -5 & 4 \\ 1 & 2 & -3 \end{pmatrix}$

(vii) $u = 2 - 5t + 6t^2 - t^3$, $v = 3 + 2t - 4t^2 + 5t^3$

(viii) $u = 1 - 3t + 2t^2 - 3t^3$, $v = -3 + 9t - 6t^2 + 9t^3$

Two vectors $u$ and $v$ are dependent if and only if one is a multiple of the other.

(i) No. (ii) Yes, for $v = 3u$. (iii) No. (iv) Yes; for $u = -2v$. (v) Yes; for $v = 2u$. (vi) No. (vii) No. (viii) Yes; for $v = -3u$.

**5.2.** Determine whether or not the following vectors in $\mathbf{R}^3$ are linearly dependent:

(i) $(1, -2, 1), (2, 1, -1), (7, -4, 1)$

(ii) $(1, -3, 7), (2, 0, -6), (3, -1, -1), (2, 4, -5)$

(iii) $(1, 2, -3), (1, -3, 2), (2, -1, 5)$

(iv) $(2, -3, 7), (0, 0, 0), (3, -1, -4)$

(i) **Method 1.** Set a linear combination of the vectors equal to the zero vector using unknown scalars $x$, $y$ and $z$:

$$x(1, -2, 1) + y(2, 1, -1) + z(7, -4, 1) = (0, 0, 0)$$

Then
$$(x, -2x, x) + (2y, y, -y) + (7z, -4z, z) = (0, 0, 0)$$
or
$$(x + 2y + 7z, -2x + y - 4z, x - y + z) = (0, 0, 0)$$

Set corresponding components equal to each other to obtain the equivalent homogeneous system, and reduce to echelon form:

$$\begin{aligned} x + 2y + 7z &= 0 \\ -2x + y - 4z &= 0 \\ x - y + z &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + 7z &= 0 \\ 5y + 10z &= 0 \\ -3y - 6z &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y + 7z &= 0 \\ y + 2z &= 0 \end{aligned}$$

The system, in echelon form, has only two nonzero equations in the three unknowns; hence the system has a nonzero solution. Thus the original vectors are linearly dependent.

Method 2. Form the matrix whose rows are the given vectors, and reduce to echelon form using the elementary row operations:

$$\begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 7 & -4 & 1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 10 & -6 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 1 \\ 0 & 5 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Since the echelon matrix has a zero row, the vectors are dependent. (The three given vectors generate a space of dimension 2.)

(ii) Yes, since any four (or more) vectors in $\mathbf{R}^3$ are dependent.

(iii) Form the matrix whose rows are the given vectors, and row reduce the matrix to echelon form:

$$\begin{pmatrix} 1 & 2 & -3 \\ 1 & -3 & 2 \\ 2 & -1 & 5 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & -5 & 11 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

Since the echelon matrix has no zero rows, the vectors are independent. (The three given vectors generate a space of dimension 3.)

(iv) Since $0 = (0, 0, 0)$ is one of the vectors, the vectors are dependent.

S **5.3.** Let $V$ be the vector space of $2 \times 2$ matrices over $\mathbf{R}$. Determine whether the matrices $A, B, C \in V$ are dependent where:

(i) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 1 & -5 \\ -4 & 0 \end{pmatrix}$

(i) Set a linear combination of the matrices $A$, $B$ and $C$ equal to the zero matrix using unknown scalars $x$, $y$ and $z$, that is, set $xA + yB + zC = 0$. Thus:

$$x\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + y\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + z\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

or
$$\begin{pmatrix} x & x \\ x & x \end{pmatrix} + \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} + \begin{pmatrix} z & z \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

or
$$\begin{pmatrix} x + y + z & x + z \\ x & x + y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Set corresponding entries equal to each other to obtain the equivalent homogeneous system of equations:

$$\begin{aligned} x + y + z &= 0 \\ x + \phantom{y +} z &= 0 \\ x \phantom{+ y + z} &= 0 \\ x + y \phantom{+ z} &= 0 \end{aligned}$$

Solving the above system we obtain only the zero solution, $x = 0,\ y = 0,\ z = 0$. We have shown that $xA + yB + zC$ implies $x = 0,\ y = 0,\ z = 0$; hence the matrices $A, B$ and $C$ are linearly independent.

(ii) Set a linear combination of the matrices $A, B$ and $C$ equal to the zero vector using unknown scalars $x, y$ and $z$; that is, set $xA + yB + zC = 0$. Thus:

$$x\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} + y\begin{pmatrix} 3 & -1 \\ 2 & 2 \end{pmatrix} + z\begin{pmatrix} 1 & -5 \\ -4 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

or

$$\begin{pmatrix} x & 2x \\ 3x & x \end{pmatrix} + \begin{pmatrix} 3y & -y \\ 2y & 2y \end{pmatrix} + \begin{pmatrix} z & -5z \\ -4z & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

or

$$\begin{pmatrix} x + 3y + z & 2x - y - 5z \\ 3x + 2y - 4z & x + 2y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Set corresponding entries equal to each other to obtain the equivalent homogeneous system of linear equations and reduce to echelon form:

$$\begin{aligned} x + 3y + z &= 0 \\ 2x - y - 5z &= 0 \\ 3x + 2y - 4z &= 0 \\ x + 2y &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 3y + z &= 0 \\ -7y - 7z &= 0 \\ -7y - 7z &= 0 \\ -y - z &= 0 \end{aligned}$$

or finally

$$\begin{aligned} x + 3y + z &= 0 \\ y + z &= 0 \end{aligned}$$

The system in echelon form has a free variable and hence a nonzero solution, for example, $x = 2,\ y = -1,\ z = 1$. We have shown that $xA + yB + zC = 0$ does not imply that $x = 0$, $y = 0,\ z = 0$; hence the matrices are linearly dependent.

**5.4.** Let $V$ be the vector space of polynomials of degree $\leq 3$ over $\mathbf{R}$. Determine whether $u, v, w \in V$ are independent or dependent where:

(i) $u = t^3 - 3t^2 + 5t + 1,\quad v = t^3 - t^2 + 8t + 2,\quad w = 2t^3 - 4t^2 + 9t + 5$

(ii) $u = t^3 + 4t^2 - 2t + 3,\quad v = t^3 + 6t^2 - t + 4,\quad w = 3t^3 + 8t^2 - 8t + 7$

(i) Set a linear combination of the polynomials $u, v$ and $w$ equal to the zero polynomial using unknown scalars $x, y$ and $z$; that is, set $xu + yv + zw = 0$. Thus:

$$x(t^3 - 3t^2 + 5t + 1) + y(t^3 - t^2 + 8t + 2) + z(2t^3 - 4t^2 + 9t + 5) = 0$$

or

$$xt^3 - 3xt^2 + 5xt + x + yt^3 - yt^2 + 8yt + 2y + 2zt^3 - 4zt^2 + 9zt + 5z = 0$$

or

$$(x + y + 2z)t^3 + (-3x - y - 4z)t^2 + (5x + 8y + 9z)t + (x + 2y + 5z) = 0$$

The coefficients of the powers of $t$ must each be 0:

$$\begin{aligned} x + y + 2z &= 0 \\ -3x - y - 4z &= 0 \\ 5x + 8y + 9z &= 0 \\ x + 2y + 5z &= 0 \end{aligned}$$

Solving the above homogeneous system, we obtain only the zero solution: $x = 0,\ y = 0,\ z = 0$; hence $u, v$ and $w$ are independent.

(ii) Set a linear combination of the polynomials $u, v$ and $w$ equal to the zero polynomial using unknown scalars $x, y$ and $z$; that is, set $xu + yv + zw = 0$. Thus

$$x(t^3 + 4t^2 - 2t + 3) + y(t^3 + 6t^2 - t + 4) + z(3t^3 + 8t^2 - 8t + 7) = 0$$

or $$xt^3 + 4xt^2 - 2xt + 3x + yt^3 + 6yt^2 - yt + 4y + 3zt^3 + 8zt^2 - 8zt + 7z = 0$$

or $$(x + y + 3z)t^3 + (4x + 6y + 8z)t^2 + (-2x - y - 8z)t + (3x + 4y + 7z) = 0$$

Set the coefficients of the powers of $t$ each equal to 0 and reduce the system to echelon form:

$$\begin{aligned} x + y + 3z &= 0 \\ 4x + 6y + 8z &= 0 \\ -2x - y - 8z &= 0 \\ 3x + 4y + 7z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + y + 3z &= 0 \\ 2y - 4z &= 0 \\ y - 2z &= 0 \\ y - 2z &= 0 \end{aligned}$$

or finally
$$\begin{aligned} x + y + 3z &= 0 \\ y - 2z &= 0 \end{aligned}$$

The system in echelon form has a free variable and hence a nonzero solution. We have shown that $xu + yv + zw = 0$ does not imply that $x = 0, y = 0, z = 0$; hence the polynomials are linearly dependent.

**5.5.** Let $V$ be the vector space of functions from $\mathbf{R}$ into $\mathbf{R}$. Show that $f, g, h \in V$ are independent where: (i) $f(t) = e^{2t}$, $g(t) = t^2$, $h(t) = t$; (ii) $f(t) = \sin t$, $g(t) = \cos t$, $h(t) = t$.

In each case set a linear combination of the functions equal to the zero function $\mathbf{0}$ using unknown scalars $x, y$ and $z$: $xf + yg + zh = \mathbf{0}$; and then show that $x = 0, y = 0, z = 0$. We emphasize that $xf + yg + zh = \mathbf{0}$ means that, for every value of $t$, $xf(t) + yg(t) + zh(t) = 0$.

(i) In the equation $xe^{2t} + yt^2 + zt = 0$, substitute

$$\begin{aligned} t &= 0 \quad \text{to obtain} \quad xe^0 + y0 + z0 = 0 \quad \text{or} \quad x = 0 \\ t &= 1 \quad \text{to obtain} \quad xe^2 + y + z = 0 \\ t &= 2 \quad \text{to obtain} \quad xe^4 + 4y + 2z = 0 \end{aligned}$$

Solve the system $\begin{cases} x = 0 \\ xe^2 + y + z = 0 \\ xe^4 + 4y + 2z = 0 \end{cases}$ to obtain only the zero solution: $x = 0, y = 0, z = 0$.

Hence $f, g$ and $h$ are independent.

(ii) Method 1. In the equation $x \sin t + y \cos t + zt = 0$, substitute

$$\begin{aligned} t &= 0 \quad \text{to obtain} \quad x \cdot 0 + y \cdot 1 + z \cdot 0 = 0 \quad \text{or} \quad y = 0 \\ t &= \pi/2 \quad \text{to obtain} \quad x \cdot 1 + y \cdot 0 + z\pi/2 = 0 \quad \text{or} \quad x + \pi z/2 = 0 \\ t &= \pi \quad \text{to obtain} \quad x \cdot 0 + y(-1) + z \cdot \pi = 0 \quad \text{or} \quad -y + \pi z = 0 \end{aligned}$$

Solve the system $\begin{cases} y = 0 \\ x + \pi z/2 = 0 \\ -y + \pi z = 0 \end{cases}$ to obtain only the zero solution: $x = 0, y = 0, z = 0$. Hence $f, g$ and $h$ are independent.

Method 2. Take the first, second and third derivatives of $x \sin t + y \cos t + zt = 0$ with respect to $t$ to get

$$x \cos t - y \sin t + z = 0 \qquad (1)$$

$$-x \sin t - y \cos t = 0 \qquad (2)$$

$$-x \cos t + y \sin t = 0 \qquad (3)$$

Add (*1*) and (*3*) to obtain $z = 0$. Multiply (*2*) by $\sin t$ and (*3*) by $\cos t$, and then add:

$$\begin{array}{lrl} \sin t \times (2): & -x\sin^2 t - y\sin t\cos t & = 0 \\ \cos t \times (3): & -x\cos^2 t + y\sin t\cos t & = 0 \\ \hline & -x(\sin^2 t + \cos^2 t) & = 0 \quad \text{or} \quad x = 0 \end{array}$$

Lastly, multiply (*2*) by $-\cos t$ and (*3*) by $\sin t$; and then add to obtain

$$y(\cos^2 t + \sin^2 t) = 0 \quad \text{or} \quad y = 0$$

Since $\quad x\sin t + y\cos t + zt = 0 \quad$ implies $\quad x = 0,\ y = 0,\ z = 0$

$f$, $g$ and $h$ are independent.

**5.6.** **Let $u$, $v$ and $w$ be independent vectors. Show that $u+v$, $u-v$ and $u-2v+w$ are also independent.**

Suppose $x(u+v) + y(u-v) + z(u-2v+w) = 0$ where $x, y$ and $z$ are scalars. Then $xu + xv + yu - yv + zu - 2zv + zw = 0$ or

$$(x+y+z)u + (x-y-2z)v + zw = 0$$

But $u$, $v$ and $w$ are linearly independent; hence the coefficients in the above relation are each 0:

$$\begin{aligned} x + y + z &= 0 \\ x - y - 2z &= 0 \\ z &= 0 \end{aligned}$$

The only solution to the above system is $x = 0$, $y = 0$, $z = 0$. Hence $u+v$, $u-v$ and $u-2v+w$ are independent.

**5.7.** **Let $v_1, v_2, \ldots, v_m$ be independent vectors, and suppose $u$ is a linear combination of the $v_i$, say $u = a_1v_1 + a_2v_2 + \cdots + a_mv_m$ where the $a_i$ are scalars. Show that the above representation of $u$ is unique.**

Suppose $u = b_1v_1 + b_2v_2 + \cdots + b_mv_m$ where the $b_i$ are scalars. Subtracting,

$$0 = u - u = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_m - b_m)v_m$$

But the $v_i$ are linearly independent; hence the coefficients in the above relation are each 0:

$$a_1 - b_1 = 0, \quad a_2 - b_2 = 0, \quad \ldots, \quad a_m - b_m = 0$$

Hence $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_m = b_m$ and so the above representation of $u$ as a linear combination of the $v_i$ is unique.

**5.8.** **Show that the vectors $v = (1+i, 2i)$ and $w = (1, 1+i)$ in $\mathbf{C}^2$ are linearly dependent over the complex field $\mathbf{C}$ but are linearly independent over the real field $\mathbf{R}$.**

Recall that 2 vectors are dependent iff one is a multiple of the other. Since the first coordinate of $w$ is 1, $v$ can be a multiple of $w$ iff $v = (1+i)w$. But $1 + i \notin \mathbf{R}$; hence $v$ and $w$ are independent over $\mathbf{R}$. Since

$$(1+i)w = (1+i)(1, 1+i) = (1+i, 2i) = v$$

and $1 + i \in \mathbf{C}$, they are dependent over $\mathbf{C}$.

**5.9.** **Suppose $S = \{v_1, \ldots, v_m\}$ contains a dependent subset, say $\{v_1, \ldots, v_r\}$. Show that $S$ is also dependent. Hence every subset of an independent set is independent.**

Since $\{v_1, \ldots, v_r\}$ is dependent, there exist scalars $a_1, \ldots, a_r$, not all 0, such that

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = 0$$

Hence there exist scalars $a_1, \ldots, a_r, 0, \ldots, 0$, not all 0, such that

$$a_1v_1 + \cdots + a_rv_r + 0v_{r+1} + \cdots + 0v_m = 0$$

Accordingly, $S$ is dependent.

92 **5.10.** Suppose $\{v_1, \ldots, v_m\}$ is independent, but $\{v_1, \ldots, v_m, w\}$ is dependent. Show that $w$ is a linear combination of the $v_i$.

**Method 1.** Since $\{v_1, \ldots, v_m, w\}$ is dependent, there exist scalars $a_1, \ldots, a_m, b$, not all 0, such that $a_1v_1 + \cdots + a_mv_m + bw = 0$. If $b = 0$, then one of the $a_i$ is not zero and $a_1v_1 + \cdots + a_mv_m = 0$. But this contradicts the hypothesis that $\{v_1, \ldots, v_m\}$ is independent. Accordingly, $b \neq 0$ and so

$$w = b^{-1}(-a_1v_1 - \cdots - a_mv_m) = -b^{-1}a_1v_1 - \cdots - b^{-1}a_mv_m$$

That is, $w$ is a linear combination of the $v_i$.

**Method 2.** If $w = 0$, then $w = 0v_1 + \cdots + 0v_m$. On the other hand, if $w \neq 0$ then, by Lemma 5.2, one of the vectors in $\{v_1, \ldots, v_m, w\}$ is a linear combination of the preceding vectors. This vector cannot be one of the $v$'s since $\{v_1, \ldots, v_m\}$ is independent. Hence $w$ is a linear combination of the $v_i$.

## PROOFS OF THEOREMS

A **5.11.** Prove Lemma 5.2: The nonzero vectors $v_1, \ldots, v_m$ are linearly dependent if and only if one of them, say $v_i$, is a linear combination of the preceding vectors: $v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1}$.

Suppose $v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1}$. Then

$$a_1v_1 + \cdots + a_{i-1}v_{i-1} - v_i + 0v_{i+1} + \cdots + 0v_m = 0$$

and the coefficient of $v_i$ is not 0. Hence the $v_i$ are linearly dependent.

Conversely, suppose the $v_i$ are linearly dependent. Then there exist scalars $a_1, \ldots, a_m$, not all 0, such that $a_1v_1 + \cdots + a_mv_m = 0$. Let $k$ be the largest integer such that $a_k \neq 0$. Then

$$a_1v_1 + \cdots + a_kv_k + 0v_{k+1} + \cdots + 0v_m = 0 \quad \text{or} \quad a_1v_1 + \cdots + a_kv_k = 0$$

Suppose $k = 1$; then $a_1v_1 = 0$, $a_1 \neq 0$ and so $v_1 = 0$. But the $v_i$ are nonzero vectors; hence $k > 1$ and

$$v_k = -a_k^{-1}a_1v_1 - \cdots - a_k^{-1}a_{k-1}v_{k-1}$$

That is, $v_k$ is a linear combination of the preceding vectors.

**5.12.** Prove Theorem 5.1: The nonzero rows $R_1, \ldots, R_n$ of a matrix in echelon form are linearly independent.

Suppose $\{R_n, R_{n-1}, \ldots, R_1\}$ is dependent. Then one of the rows, say $R_m$, is a linear combination of the preceding rows:

$$R_m = a_{m+1}R_{m+1} + a_{m+2}R_{m+2} + \cdots + a_nR_n \tag{*}$$

Now suppose the $k$th component of $R_m$ is its first nonzero entry. Then, since the matrix is in echelon form, the $k$th components of $R_{m+1}, \ldots, R_n$ are all 0, and so the $k$th component of (*) is $a_{m+1} \cdot 0 + a_{m+2} \cdot 0 + \cdots + a_n \cdot 0 = 0$. But this contradicts the assumption that the $k$th component of $R_m$ is not 0. Thus $R_1, \ldots, R_n$ are independent.

A **5.13.** Suppose $\{v_1, \ldots, v_m\}$ generates a vector space $V$. Prove:

(i) If $w \in V$, then $\{w, v_1, \ldots, v_m\}$ is linearly dependent and generates $V$.

(ii) If $v_i$ is a linear combination of the preceding vectors, then $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ generates $V$.

(i) If $w \in V$, then $w$ is a linear combination of the $v_i$ since $\{v_i\}$ generates $V$. Accordingly, $\{w, v_1, \ldots, v_m\}$ is linearly dependent. Clearly, $w$ with the $v_i$ generate $V$ since the $v_i$ by themselves generate $V$. That is, $\{w, v_1, \ldots, v_m\}$ generates $V$.

(ii) Suppose $v_i = k_1v_1 + \cdots + k_{i-1}v_{i-1}$. Let $u \in V$. Since $\{v_i\}$ generates $V$, $u$ is a linear combination of the $v_i$, say, $u = a_1v_1 + \cdots + a_mv_m$. Substituting for $v_i$, we obtain

$$u = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_i(k_1v_1 + \cdots + k_{i-1}v_{i-1}) + a_{i+1}v_{i+1} + \cdots + a_mv_m$$
$$= (a_1 + a_ik_1)v_1 + \cdots + (a_{i-1} + a_ik_{i-1})v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_mv_m$$

Thus $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ generates $V$. In other words, we can delete $v_i$ from the generating set and still retain a generating set.

**5.14.** Prove Lemma 5.4: Suppose $\{v_1, \ldots, v_n\}$ generates a vector space $V$. If $\{w_1, \ldots, w_m\}$ is linearly independent, then $m \leq n$ and $V$ is generated by a set of the form $\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$. Thus, in particular, any $n+1$ or more vectors in $V$ are linearly dependent.

90

It suffices to prove the theorem in the case that the $v_i$ are all not 0. (Prove!) Since the $\{v_i\}$ generates $V$, we have by the preceding problem that (1)

$$\{w_1, v_1, \ldots, v_n\}$$

is linearly dependent and also generates $V$. By Lemma 5.2, one of the vectors in (1) is a linear combination of the preceding vectors. This vector cannot be $w_1$, so it must be one of the $v$'s, say $v_j$. Thus by the preceding problem we can delete $v_j$ from the generating set (1) and obtain the generating set (2)

$$\{w_1, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\}$$

Now we repeat the argument with the vector $w_2$. That is, since (2) generates $V$, the set (3)

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\}$$

is linearly dependent and also generates $V$. Again by Lemma 5.2, one of the vectors in (3) is a linear combination of the preceding vectors. We emphasize that this vector cannot be $w_1$ or $w_2$ since $\{w_1, \ldots, w_m\}$ is independent; hence it must be one of the $v$'s, say $v_k$. Thus by the preceding problem we can delete $v_k$ from the generating set (3) and obtain the generating set

$$\{w_1, w_2, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\}$$

We repeat the argument with $w_3$ and so forth. At each step we are able to add one of the $w$'s and delete one of the $v$'s in the generating set. If $m \leq n$, then we finally obtain a generating set of the required form:

$$\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\}$$

Lastly, we show that $m > n$ is not possible. Otherwise, after $n$ of the above steps, we obtain the generating set $\{w_1, \ldots, w_n\}$. This implies that $w_{n+1}$ is a linear combination of $w_1, \ldots, w_n$ which contradicts the hypothesis that $\{w_i\}$ is linearly independent.

**5.15.** Prove Theorem 5.3: Let $V$ be a finite dimensional vector space. Then every basis of $V$ has the same number of vectors.

91

Suppose $\{e_1, e_2, \ldots, e_n\}$ is a basis of $V$, and suppose $\{f_1, f_2, \ldots\}$ is another basis of $V$. Since $\{e_i\}$ generates $V$, the basis $\{f_1, f_2, \ldots\}$ must contain $n$ or less vectors, or else it is dependent by the preceding problem. On the other hand, if the basis $\{f_1, f_2, \ldots\}$ contains less than $n$ vectors, then $\{e_1, \ldots, e_n\}$ is dependent by the preceding problem. Thus the basis $\{f_1, f_2, \ldots\}$ contains exactly $n$ vectors, and so the theorem is true.

**5.16.** Prove Theorem 5.5: Suppose $\{v_1, \ldots, v_m\}$ is a maximal independent subset of a set $S$ which generates a vector space $V$. Then $\{v_1, \ldots, v_m\}$ is a basis of $V$.

93

Suppose $w \in S$. Then, since $\{v_i\}$ is a maximal independent subset of $S$, $\{v_1, \ldots, v_m, w\}$ is linearly dependent. By Problem 5.10, $w$ is a linear combination of the $v_i$, that is, $w \in L(v_i)$. Hence $S \subset L(v_i)$. This leads to $V = L(S) \subset L(v_i) \subset V$. Accordingly, $\{v_i\}$ generates $V$ and, since it is independent, it is a basis of $V$.

A 5.17. Suppose $V$ is generated by a finite set $S$. Show that $V$ is of finite dimension and, in particular, a subset of $S$ is a basis of $V$.

**Method 1.** Of all the independent subsets of $S$, and there is a finite number of them since $S$ is finite, one of them is maximal. By the preceding problem this subset of $S$ is a basis of $V$.

**Method 2.** If $S$ is independent, it is a basis of $V$. If $S$ is dependent, one of the vectors is a linear combination of the preceding vectors. We may delete this vector and still retain a generating set. We continue this process until we obtain a subset which is independent and generates $V$, i.e. is a basis of $V$.

**5.18.** Prove Theorem 5.6: Let $V$ be of finite dimension $n$. Then:

(i) Any set of $n+1$ or more vectors is linearly dependent.

(ii) Any linearly independent set is part of a basis.

(iii) A linearly independent set with $n$ elements is a basis.

Suppose $\{e_1, \ldots, e_n\}$ is a basis of $V$.

(i) Since $\{e_1, \ldots, e_n\}$ generates $V$, any $n+1$ or more vectors is dependent by Lemma 5.4.

(ii) Suppose $\{v_1, \ldots, v_r\}$ is independent. By Lemma 5.4, $V$ is generated by a set of the form

$$S = \{v_1, \ldots, v_r, e_{i_1}, \ldots, e_{i_{n-r}}\}$$

By the preceding problem, a subset of $S$ is a basis. But $S$ contains $n$ elements and every basis of $V$ contains $n$ elements. Thus $S$ is a basis of $V$ and contains $\{v_1, \ldots, v_r\}$ as a subset.

(iii) By (ii), an independent set $T$ with $n$ elements is part of a basis. But every basis of $V$ contains $n$ elements. Thus, $T$ is a basis.

**5.19.** Prove Theorem 5.7: Let $W$ be a subspace of an $n$-dimensional vector space $V$. Then $\dim W \leq n$. In particular, if $\dim W = n$, then $W = V$.

Since $V$ is of dimension $n$, any $n+1$ or more vectors are linearly dependent. Furthermore, since a basis of $W$ consists of linearly independent vectors, it cannot contain more than $n$ elements. Accordingly, $\dim W \leq n$.

In particular, if $\{w_1, \ldots, w_n\}$ is a basis of $W$, then since it is an independent set with $n$ elements it is also a basis of $V$. Thus $W = V$ when $\dim W = n$.

9/1/92 v.v.I

**5.20.** Prove Theorem 5.8: $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

Observe that $U \cap W$ is a subspace of both $U$ and $W$. Suppose $\dim U = m$, $\dim W = n$ and $\dim(U \cap W) = r$. Suppose $\{v_1, \ldots, v_r\}$ is a basis of $U \cap W$. By Theorem 5.6(ii), we can extend $\{v_i\}$ to a basis of $U$ and to a basis of $W$; say,

$$\{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}\} \quad \text{and} \quad \{v_1, \ldots, v_r, w_1, \ldots, w_{n-r}\}$$

are bases of $U$ and $W$ respectively. Let

$$B = \{v_1, \ldots, v_r, u_1, \ldots, u_{m-r}, w_1, \ldots, w_{n-r}\}$$

Note that $B$ has exactly $m+n-r$ elements. Thus the theorem is proved if we can show that $B$ is a basis of $U+W$. Since $\{v_i, u_j\}$ generates $U$ and $\{v_i, w_k\}$ generates $W$, the union $B = \{v_i, u_j, w_k\}$ generates $U+W$. Thus it suffices to show that $B$ is independent.

Suppose

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0 \qquad (1)$$

where $a_i, b_j, c_k$ are scalars. Let

$$v = a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} \qquad (2)$$

By (1), we also have that

$$v = -c_1w_1 - \cdots - c_{n-r}w_{n-r} \tag{3}$$

Since $\{v_i, u_j\} \subset U$, $v \in U$ by (2); and since $\{w_k\} \subset W$, $v \in W$ by (3). Accordingly, $v \in U \cap W$. Now $\{v_i\}$ is a basis of $U \cap W$ and so there exist scalars $d_1, \ldots, d_r$ for which $v = d_1v_1 + \cdots + d_rv_r$. Thus by (3) we have

$$d_1v_1 + \cdots + d_rv_r + c_1w_1 + \cdots + c_{n-r}w_{n-r} = 0$$

But $\{v_i, w_k\}$ is a basis of $W$ and so is independent. Hence the above equation forces $c_1 = 0, \ldots, c_{n-r} = 0$. Substituting this into (1), we obtain

$$a_1v_1 + \cdots + a_rv_r + b_1u_1 + \cdots + b_{m-r}u_{m-r} = 0$$

But $\{v_i, u_j\}$ is a basis of $U$ and so is independent. Hence the above equation forces $a_1 = 0, \ldots, a_r = 0$, $b_1 = 0, \ldots, b_{m-r} = 0$.

Since the equation (1) implies that the $a_i$, $b_j$ and $c_k$ are all 0, $B = \{v_i, u_j, w_k\}$ is independent and the theorem is proved.

93, 92, 88, 86

**5.21.** Prove Theorem 5.9: The row rank and the column rank of any matrix are equal.

Let $A$ be an arbitrary $m \times n$ matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Let $R_1, R_2, \ldots, R_m$ denote its rows:

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \quad \ldots, \quad R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

Suppose the row rank is $r$ and that the following $r$ vectors form a basis for the row space:

$$S_1 = (b_{11}, b_{12}, \ldots, b_{1n}), \quad S_2 = (b_{21}, b_{22}, \ldots, b_{2n}), \quad \ldots, \quad S_r = (b_{r1}, b_{r2}, \ldots, b_{rn})$$

Then each of the row vectors is a linear combination of the $S_i$:

$$\begin{aligned} R_1 &= k_{11}S_1 + k_{12}S_2 + \cdots + k_{1r}S_r \\ R_2 &= k_{21}S_1 + k_{22}S_2 + \cdots + k_{2r}S_r \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ R_m &= k_{m1}S_1 + k_{m2}S_2 + \cdots + k_{mr}S_r \end{aligned}$$

where the $k_{ij}$ are scalars. Setting the $i$th components of each of the above vector equations equal to each other, we obtain the following system of equations, each valid for $i = 1, \ldots, n$:

$$\begin{aligned} a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots + k_{1r}b_{ri} \\ a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots + k_{2r}b_{ri} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots + k_{mr}b_{ri} \end{aligned}$$

Thus for $i = 1, \ldots, n$:

$$\begin{pmatrix} a_{1i} \\ a_{2i} \\ \cdots \\ a_{mi} \end{pmatrix} = b_{1i}\begin{pmatrix} k_{11} \\ k_{21} \\ \cdots \\ k_{m1} \end{pmatrix} + b_{2i}\begin{pmatrix} k_{12} \\ k_{22} \\ \cdots \\ k_{m2} \end{pmatrix} + \cdots + b_{ri}\begin{pmatrix} k_{1r} \\ k_{2r} \\ \cdots \\ k_{mr} \end{pmatrix}$$

In other words, each of the columns of $A$ is a linear combination of the $r$ vectors

$$\begin{pmatrix} k_{11} \\ k_{21} \\ \cdots \\ k_{m1} \end{pmatrix}, \begin{pmatrix} k_{12} \\ k_{22} \\ \cdots \\ k_{m2} \end{pmatrix}, \ldots, \begin{pmatrix} k_{1r} \\ k_{2r} \\ \cdots \\ k_{mr} \end{pmatrix}$$

Thus the column space of the matrix $A$ has dimension at most $r$, i.e. column rank $\leq r$. Hence, column rank $\leq$ row rank.

Similarly (or considering the transpose matrix $A^t$) we obtain row rank $\leq$ column rank. Thus the row rank and column rank are equal.

## BASIS AND DIMENSION

**5.22.** Determine whether or not the following form a basis for the vector space $\mathbf{R}^3$:

(i) $(1, 1, 1)$ and $(1, -1, 5)$

(ii) $(1, 2, 3)$, $(1, 0, -1)$, $(3, -1, 0)$ and $(2, 1, -2)$

(iii) $(1, 1, 1)$, $(1, 2, 3)$ and $(2, -1, 1)$

(iv) $(1, 1, 2)$, $(1, 2, 5)$ and $(5, 3, 4)$

(i) and (ii). No; for a basis of $\mathbf{R}^3$ must contain exactly 3 elements, since $\mathbf{R}^3$ is of dimension 3.

(iii) The vectors form a basis if and only if they are independent. Thus form the matrix whose rows are the given vectors, and row reduce to echelon form:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & -1 & 1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -3 & -1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{pmatrix}$$

The echelon matrix has no zero rows; hence the three vectors are independent and so form a basis for $\mathbf{R}^3$.

(iv) Form the matrix whose rows are the given vectors, and row reduce to echelon form:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 5 & 3 & 4 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & -2 & -6 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

The echelon matrix has a zero row, i.e. only two nonzero rows; hence the three vectors are dependent and so do not form a basis for $\mathbf{R}^3$.

**5.23.** Let $W$ be the subspace of $\mathbf{R}^4$ generated by the vectors $(1, -2, 5, -3)$, $(2, 3, 1, -4)$ and $(3, 8, -3, -5)$. (i) Find a basis and the dimension of $W$. (ii) Extend the basis of $W$ to a basis of the whole space $\mathbf{R}^4$.

(i) Form the matrix whose rows are the given vectors, and row reduce to echelon form:

$$\begin{pmatrix} 1 & -2 & 5 & -3 \\ 2 & 3 & 1 & -4 \\ 3 & 8 & -3 & -5 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 5 & 3 \\ 0 & 7 & -9 & 2 \\ 0 & 14 & -18 & 4 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 5 & -3 \\ 0 & 7 & -9 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The nonzero rows $(1, -2, 5, -3)$ and $(0, 7, -9, 2)$ of the echelon matrix form a basis of the row space, that is, of $W$. Thus, in particular, $\dim W = 2$.

(ii) We seek four independent vectors which include the above two vectors. The vectors $(1, -2, 5, -3)$, $(0, 7, -9, 2)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$ are independent (since they form an echelon matrix), and so they form a basis of $\mathbf{R}^4$ which is an extension of the basis of $W$.

**5.24.** Let $W$ be the space generated by the polynomials

$$v_1 = t^3 - 2t^2 + 4t + 1 \qquad v_3 = t^3 + 6t - 5$$
$$v_2 = 2t^3 - 3t^2 + 9t - 1 \qquad v_4 = 2t^3 - 5t^2 + 7t + 5$$

Find a basis and the dimension of $W$.

The coordinate vectors of the given polynomials relative to the basis $\{t^3, t^2, t, 1\}$ are respectively

$$[v_1] = (1, -2, 4, 1) \qquad [v_3] = (1, 0, 6, -5)$$
$$[v_2] = (2, -3, 9, -1) \qquad [v_4] = (2, -5, 7, 5)$$

Form the matrix whose rows are the above coordinate vectors, and row reduce to echelon form:

$$\begin{pmatrix} 1 & -2 & 4 & 1 \\ 2 & -3 & 9 & -1 \\ 1 & 0 & 6 & -5 \\ 2 & -5 & 7 & 5 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 2 & 2 & -6 \\ 0 & -1 & -1 & 3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 4 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The nonzero rows $(1,-2,4,1)$ and $(0,1,1,-3)$ of the echelon matrix form a basis of the space generated by the coordinate vectors, and so the corresponding polynomials

$$t^3 - 2t^2 + 4t + 1 \quad \text{and} \quad t^2 + t - 3$$

form a basis of $W$. Thus $\dim W = 2$.

**5.25.** Find the dimension and a basis of the solution space $W$ of the system

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ x + 2y + 3z + s + t &= 0 \\ 3x + 6y + 8z + s + 5t &= 0 \end{aligned}$$

Reduce the system to echelon form:

$$\begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \\ 2z + 4s - 4t &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y + 2z - s + 3t &= 0 \\ z + 2s - 2t &= 0 \end{aligned}$$

The system in echelon form has 2 (nonzero) equations in 5 unknowns; hence the dimension of the solution space $W$ is $5 - 2 = 3$. The free variables are $y$, $s$ and $t$. Set

(i) $y = 1, s = 0, t = 0$, (ii) $y = 0, s = 1, t = 0$, (iii) $y = 0, s = 0, t = 1$

to obtain the respective solutions

$$v_1 = (-2, 1, 0, 0, 0), \quad v_2 = (5, 0, -2, 1, 0), \quad v_3 = (-7, 0, 2, 0, 1)$$

The set $\{v_1, v_2, v_3\}$ is a basis of the solution space $W$.

**5.26.** Find a homogeneous system whose solution set $W$ is generated by

$$\{(1, -2, 0, 3), (1, -1, -1, 4), (1, 0, -2, 5)\}$$

Method 1. Let $v = (x, y, z, w)$. Form the matrix $M$ whose first rows are the given vectors and whose last row is $v$; and then row reduce to echelon form:

$$M = \begin{pmatrix} 1 & -2 & 0 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 0 & -2 & 5 \\ x & y & z & w \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & -2 & 2 \\ 0 & 2x + y & z & -3x + w \end{pmatrix} \text{ to } \begin{pmatrix} 1 & -2 & 0 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2x + y + z & -5x - y + w \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The original first three rows show that $W$ has dimension 2. Thus $v \in W$ if and only if the additional row does not increase the dimension of the row space. Hence we set the last two entries in the third row on the right equal to 0 to obtain the required homogeneous system

$$\begin{aligned} 2x + y + z \phantom{- w} &= 0 \\ 5x + y \phantom{+ z} - w &= 0 \end{aligned}$$

Method 2. We know that $v = (x, y, z, w) \in W$ if and only if $v$ is a linear combination of the generators of $W$:

$$(x, y, z, w) = r(1, -2, 0, 3) + s(1, -1, -1, 4) + t(1, 0, -2, 5)$$

The above vector equation in unknowns $r$, $s$ and $t$ is equivalent to the following system:

$$\begin{aligned} r + s + t &= x \\ -2r - s \phantom{+ 2t} &= y \\ -s - 2t &= z \\ 3r + 4s + 5t &= w \end{aligned} \quad \text{or} \quad \begin{aligned} r + s + t &= x \\ s + 2t &= 2x + y \\ -s - 2t &= z \\ s + 2t &= w - 3x \end{aligned} \quad \text{or} \quad \begin{aligned} r + s + t &= x \\ s + 2t &= 2x + y \\ 0 &= 2x + y + z \\ 0 &= 5x + y - w \end{aligned} \qquad (1)$$

Thus $v \in W$ if and only if the above system has a solution, i.e. if

$$\begin{aligned} 2x + y + z &= 0 \\ 5x + y - w &= 0 \end{aligned}$$

The above is the required homogeneous system.

**Remark:** Observe that the augmented matrix of the system (1) is the transpose of the matrix $M$ used in the first method.

A

**5.27.** Let $U$ and $W$ be the following subspaces of $\mathbf{R}^4$:

91 v.v.9

$$U = \{(a,b,c,d) : b+c+d = 0\}, \qquad W = \{(a,b,c,d) : a+b = 0,\ c = 2d\}$$

Find the dimension and a basis of (i) $U$, (ii) $W$, (iii) $U \cap W$.

(i) We seek a basis of the set of solutions $(a,b,c,d)$ of the equation

$$b + c + d = 0 \quad \text{or} \quad 0 \cdot a + b + c + d = 0$$

The free variables are $a$, $c$ and $d$. Set

$$(1)\ a = 1,\ c = 0,\ d = 0, \qquad (2)\ a = 0,\ c = 1,\ d = 0, \qquad (3)\ a = 0,\ c = 0,\ d = 1$$

to obtain the respective solutions

$$v_1 = (1,0,0,0), \qquad v_2 = (0,-1,1,0), \qquad v_3 = (0,-1,0,1)$$

The set $\{v_1, v_2, v_3\}$ is a basis of $U$, and $\dim U = 3$.

(ii) We seek a basis of the set of solutions $(a,b,c,d)$ of the system

$$\begin{aligned} a + b &= 0 \\ c &= 2d \end{aligned} \quad \text{or} \quad \begin{aligned} a + b &= 0 \\ c - 2d &= 0 \end{aligned}$$

The free variables are $b$ and $d$. Set

$$(1)\ b = 1,\ d = 0, \qquad (2)\ b = 0,\ d = 1$$

to obtain the respective solutions

$$v_1 = (-1,1,0,0), \qquad v_2 = (0,0,2,1)$$

The set $\{v_1, v_2\}$ is a basis of $W$, and $\dim W = 2$.

(iii) $U \cap W$ consists of those vectors $(a,b,c,d)$ which satisfy the conditions defining $U$ and the conditions defining $W$, i.e. the three equations

$$\begin{aligned} b + c + d &= 0 \\ a + b \phantom{+ c + d} &= 0 \\ c &= 2d \end{aligned} \quad \text{or} \quad \begin{aligned} a + b \phantom{+ c + d} &= 0 \\ b + c + d &= 0 \\ c - 2d &= 0 \end{aligned}$$

The free variable is $d$. Set $d = 1$ to obtain the solution $v = (3,-3,2,1)$. Thus $\{v\}$ is a basis of $U \cap W$, and $\dim(U \cap W) = 1$.

**5.28.** Find the dimension of the vector space spanned by:

(i) $(1,-2,3,-1)$ and $(1,1,-2,3)$

(ii) $(3,-6,3,-9)$ and $(-2,4,-2,6)$

(iii) $t^3 + 2t^2 + 3t + 1$ and $2t^3 + 4t^2 + 6t + 2$

(iv) $t^3 - 2t^2 + 5$ and $t^2 + 3t - 4$

(v) $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$

(vi) $\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ and $\begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix}$

(vii) 3 and $-3$

Two nonzero vectors span a space $W$ of dimension 2 if they are independent, and of dimension 1 if they are dependent. Recall that two vectors are dependent if and only if one is a multiple of the other. Hence: (i) 2, (ii) 1, (iii) 1, (iv) 2, (v) 2, (vi) 1, (vii) 1.

58 **5.29.** Let $V$ be the vector space of 2 by 2 symmetric matrices over $K$. Show that $\dim V = 3$. (Recall that $A = (a_{ij})$ is symmetric iff $A = A^t$ or, equivalently, $a_{ij} = a_{ji}$.)

An arbitrary 2 by 2 symmetric matrix is of the form $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ where $a, b, c \in K$. (Note that there are three "variables".) Setting

(i) $a = 1,\ b = 0,\ c = 0$, (ii) $a = 0,\ b = 1,\ c = 0$, (iii) $a = 0,\ b = 0,\ c = 1$

we obtain the respective matrices

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

We show that $\{E_1, E_2, E_3\}$ is a basis of $V$, that is, that it (1) generates $V$ and (2) is independent.

(1) For the above arbitrary matrix $A$ in $V$, we have

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} = aE_1 + bE_2 + cE_3$$

Thus $\{E_1, E_2, E_3\}$ generates $V$.

(2) Suppose $xE_1 + yE_2 + zE_3 = 0$, where $x, y, z$ are unknown scalars. That is, suppose

$$x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + z\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} x & y \\ y & z \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Setting corresponding entries equal to each other, we obtain $x = 0,\ y = 0,\ z = 0$. In other words,

$$xE_1 + yE_2 + zE_3 = 0 \quad \text{implies} \quad x = 0,\ y = 0,\ z = 0$$

Accordingly, $\{E_1, E_2, E_3\}$ is independent.

Thus $\{E_1, E_2, E_3\}$ is a basis of $V$ and so the dimension of $V$ is 3.

**5.30.** Let $V$ be the space of polynomials in $t$ of degree $\leq n$. Show that each of the following is a basis of $V$:

(i) $\{1, t, t^2, \ldots, t^{n-1}, t^n\}$, (ii) $\{1, 1-t, (1-t)^2, \ldots, (1-t)^{n-1}, (1-t)^n\}$.

Thus $\dim V = n + 1$.

(i) Clearly each polynomial in $V$ is a linear combination of $1, t, \ldots, t^{n-1}$ and $t^n$. Furthermore, $1, t, \ldots, t^{n-1}$ and $t^n$ are independent since none is a linear combination of the preceding polynomials. Thus $\{1, t, \ldots, t^n\}$ is a basis of $V$.

(ii) (Note that by (i), $\dim V = n+1$; and so any $n+1$ independent polynomials form a basis of $V$.) Now each polynomial in the sequence $1, 1-t, \ldots, (1-t)^n$ is of degree higher than the preceding ones and so is not a linear combination of the preceding ones. Thus the $n+1$ polynomials $1, 1-t, \ldots, (1-t)^n$ are independent and so form a basis of $V$.

**5.31.** Let $V$ be the vector space of ordered pairs of complex numbers over the real field $\mathbf{R}$ (see Problem 4.42). Show that $V$ is of dimension 4.

We claim that the following is a basis of $V$:

$$B = \{(1, 0), (i, 0), (0, 1), (0, i)\}$$

Suppose $v \in V$. Then $v = (z, w)$ where $z, w$ are complex numbers, and so $v = (a + bi, c + di)$ where $a, b, c, d$ are real numbers. Then

$$v = a(1, 0) + b(i, 0) + c(0, 1) + d(0, i)$$

Thus $B$ generates $V$.

The proof is complete if we show that $B$ is independent. Suppose

$$x_1(1, 0) + x_2(i, 0) + x_3(0, 1) + x_4(0, i) = 0$$

where $x_1, x_2, x_3, x_4 \in R$. Then

$$(x_1 + x_2 i, x_3 + x_4 i) = (0, 0) \quad \text{and so} \quad \begin{cases} x_1 + x_2 i = 0 \\ x_3 + x_4 i = 0 \end{cases}$$

Accordingly $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$ and so $B$ is independent.

**5.32.** Let $V$ be the vector space of $m \times n$ matrices over a field $K$. Let $E_{ij} \in V$ be the matrix with 1 as the $ij$-entry and 0 elsewhere. Show that $\{E_{ij}\}$ is a basis of $V$. Thus $\dim V = mn$.

We need to show that $\{E_{ij}\}$ generates $V$ and is independent.

Let $A = (a_{ij})$ be any matrix in $V$. Then $A = \sum_{i,j} a_{ij}E_{ij}$. Hence $\{E_{ij}\}$ generates $V$.

Now suppose that $\sum_{i,j} x_{ij}E_{ij} = 0$ where the $x_{ij}$ are scalars. The $ij$-entry of $\sum_{i,j} x_{ij}E_{ij}$ is $x_{ij}$, and the $ij$-entry of $0$ is $0$. Thus $x_{ij} = 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Accordingly the matrices $E_{ij}$ are independent.

Thus $\{E_{ij}\}$ is a basis of $V$.

**Remark:** Viewing a vector in $K^n$ as a $1 \times n$ matrix, we have shown by the above result that the usual basis defined in Example 5.3, page 88, is a basis of $K^n$ and that $\dim K^n = n$.

## SUMS AND INTERSECTIONS

**5.33.** Suppose $U$ and $W$ are distinct 4-dimensional subspaces of a vector space $V$ of dimension 6. Find the possible dimensions of $U \cap W$.

Since $U$ and $W$ are distinct, $U + W$ properly contains $U$ and $W$; hence $\dim(U + W) > 4$. But $\dim(U + W)$ cannot be greater than 6, since $\dim V = 6$. Hence we have two possibilities: (i) $\dim(U + W) = 5$, or (ii) $\dim(U + W) = 6$. Using Theorem 5.8 that $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, we obtain

$$\text{(i)} \quad 5 = 4 + 4 - \dim(U \cap W) \quad \text{or} \quad \dim(U \cap W) = 3$$

$$\text{(ii)} \quad 6 = 4 + 4 - \dim(U \cap W) \quad \text{or} \quad \dim(U \cap W) = 2$$

That is, the dimension of $U \cap W$ must be either 2 or 3.

AM 89

**5.34.** Let $U$ and $W$ be the subspaces of $\mathbf{R}^4$ generated by

$$\{(1, 1, 0, -1), (1, 2, 3, 0), (2, 3, 3, -1)\} \quad \text{and} \quad \{(1, 2, 2, -2), (2, 3, 2, -3), (1, 3, 4, -3)\}$$

respectively. Find (i) $\dim(U + W)$, (ii) $\dim(U \cap W)$.

(i) $U + W$ is the space spanned by all six vectors. Hence form the matrix whose rows are the given six vectors, and then row reduce to echelon form:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \\ 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 2 & 4 & -2 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the echelon matrix has three nonzero rows, $\dim(U + W) = 3$.

(ii) First find $\dim U$ and $\dim W$. Form the two matrices whose rows are the generators of $U$ and $W$ respectively and then row reduce each to echelon form:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 3 & -1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 & 2 & -2 \\ 2 & 3 & 2 & -3 \\ 1 & 3 & 4 & -3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 1 & 2 & -1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & 2 & -2 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since each of the echelon matrices has two nonzero rows, $\dim U = 2$ and $\dim W = 2$. Using Theorem 5.8 that $\dim(U+W) = \dim U + \dim W - \dim(U \cap W)$, we have

$$3 = 2 + 2 - \dim(U \cap W) \quad \text{or} \quad \dim(U \cap W) = 1$$

A M

**5.35.** Let $U$ be the subspace of $\mathbf{R}^5$ generated by

$$\{(1, 3, -2, 2, 3), (1, 4, -3, 4, 2), (2, 3, -1, -2, 9)\}$$

and let $W$ be the subspace generated by

$$\{(1, 3, 0, 2, 1), (1, 5, -6, 6, 3), (2, 5, 3, 2, 1)\}$$

Find a basis and the dimension of (i) $U + W$, (ii) $U \cap W$.

(i) $U + W$ is the space generated by all six vectors. Hence form the matrix whose rows are the six vectors and then row reduce to echelon form:

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & -4 & 4 & 0 \\ 0 & -1 & 7 & -2 & -5 \end{pmatrix}$$

$$\text{to } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 6 & 0 & -6 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The set of nonzero rows of the echelon matrix,

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

is a basis of $U + W$; thus $\dim(U + W) = 3$.

(ii) First find homogeneous systems whose solution sets are $U$ and $W$ respectively. Form the matrix whose first rows are the generators of $U$ and whose last row is $(x, y, z, s, t)$ and then row reduce to echelon form:

$$\begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ x & y & z & s & t \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & -3x+y & 2x+z & -2x+s & -3x+t \end{pmatrix}$$

$$\text{to } \begin{pmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & -x+y+z & 4x-2y+s & -6x+y+t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Set the entries of the third row equal to 0 to obtain the homogeneous system whose solution set is $U$: $\quad -x + y + z = 0, \quad 4x - 2y + s = 0, \quad -6x + y + t = 0$

Now form the matrix whose first rows are the generators of $W$ and whose last row is $(x, y, z, s, t)$ and then row reduce to echelon form:

$$\begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \\ x & y & z & s & t \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 2 & -6 & 4 & 2 \\ 0 & -1 & 3 & -2 & -1 \\ 0 & -3x+y & z & -2x+s & -x+t \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 1 & -3 & 2 & 1 \\ 0 & 0 & -9x+3y+z & 4x-2y+s & 2x-y+t \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Set the entries of the third row equal to 0 to obtain the homogeneous system whose solution set is $W$: $\quad -9x + 3y + z = 0, \quad 4x - 2y + s = 0, \quad 2x - y + t = 0$

Combining both systems, we obtain the homogeneous system whose solution set is $U \cap W$:

$$\begin{cases} -x + y + z = 0 \\ 4x - 2y + s = 0 \\ -6x + y + t = 0 \\ -9x + 3y + z = 0 \\ 4x - 2y + s = 0 \\ 2x - y + t = 0 \end{cases} \quad \text{or} \quad \begin{cases} -x + y + z = 0 \\ 2y + 4z + s = 0 \\ -5y - 6z + t = 0 \\ -6y - 8z = 0 \\ 2y + 4z + s = 0 \\ y + 2z + t = 0 \end{cases}$$

$$\begin{cases} -x + y + z = 0 \\ 2y + 4z + s = 0 \\ 8z + 5s + 2t = 0 \\ 4z + 3s = 0 \\ s - 2t = 0 \end{cases} \quad \text{or} \quad \begin{cases} -x + y + z = 0 \\ 2y + 4z + s = 0 \\ 8z + 5s + 2t = 0 \\ s - 2t = 0 \end{cases}$$

There is one free variable, which is $t$; hence $\dim(U \cap W) = 1$. Setting $t = 2$, we obtain the solution $x = 1$, $y = 4$, $z = -3$, $s = 4$, $t = 2$. Thus $\{(1, 4, -3, 4, 2)\}$ is a basis of $U \cap W$.

## COORDINATE VECTORS

**5.36.** Find the coordinate vector of $v$ relative to the basis $\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ of $\mathbf{R}^3$ where (i) $v = (4, -3, 2)$, (ii) $v = (a, b, c)$.

In each case set $v$ as a linear combination of the basis vectors using unknown scalars $x$, $y$ and $z$:

$$v = x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0)$$

and then solve for the solution vector $(x, y, z)$. (The solution is unique since the basis vectors are linearly independent.)

(i)
$$\begin{aligned} (4, -3, 2) &= x(1, 1, 1) + y(1, 1, 0) + z(1, 0, 0) \\ &= (x, x, x) + (y, y, 0) + (z, 0, 0) \\ &= (x + y + z, x + y, x) \end{aligned}$$

Set corresponding components equal to each other to obtain the system

$$x + y + z = 4, \quad x + y = -3, \quad x = 2$$

Substitute $x = 2$ into the second equation to obtain $y = -5$; then put $x = 2$, $y = -5$ into the first equation to obtain $z = 7$. Thus $x = 2$, $y = -5$, $z = 7$ is the unique solution to the system and so the coordinate vector of $v$ relative to the given basis is $[v] = (2, -5, 7)$.

(ii) $(a, b, c) = x(1,1,1) + y(1,1,0) + z(1,0,0) = (x+y+z, x+y, x)$

Then $x+y+z = c, \quad x+y = b, \quad x = c$

from which $x = c$, $y = b-c$, $z = a-b$. Thus $[v] = (c, b-c, a-b)$, that is, $[(a,b,c)] = (c, b-c, a-b)$.

**5.37.** Let $V$ be the vector space of $2\times 2$ matrices over $\mathbf{R}$. Find the coordinate vector of the matrix $A \in V$ relative to the basis

$$\left\{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right\} \quad \text{where} \quad A = \begin{pmatrix}2 & 3\\ 4 & -7\end{pmatrix}$$

Set $A$ as a linear combination of the matrices in the basis using unknown scalars $x, y, z, w$:

$$A = \begin{pmatrix}2 & 3\\ 4 & -7\end{pmatrix} = x\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} + y\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} + z\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix} + w\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$$

$$= \begin{pmatrix}x & x\\ x & x\end{pmatrix} + \begin{pmatrix}0 & -y\\ y & 0\end{pmatrix} + \begin{pmatrix}z & -z\\ 0 & 0\end{pmatrix} + \begin{pmatrix}w & 0\\ 0 & 0\end{pmatrix}$$

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

Set corresponding entries equal to each other to obtain the system

$$x+z+w = 2, \quad x-y-z = 3, \quad x+y = 4, \quad x = -7$$

from which $x = -7$, $y = 11$, $z = -21$, $w = 30$. Thus $[A] = (-7, 11, -21, 30)$. (Note that the coordinate vector of $A$ must be a vector in $\mathbf{R}^4$ since $\dim V = 4$.)

**5.38.** Let $W$ be the vector space of $2\times 2$ symmetric matrices over $\mathbf{R}$. (See Problem 5.29.) Find the coordinate vector of the matrix $A = \begin{pmatrix}4 & -11\\ -11 & -7\end{pmatrix}$ relative to the basis $\left\{\begin{pmatrix}1 & -2\\ -2 & 1\end{pmatrix}, \begin{pmatrix}2 & 1\\ 1 & 3\end{pmatrix}, \begin{pmatrix}4 & -1\\ -1 & -5\end{pmatrix}\right\}$.

Set $A$ as a linear combination of the matrices in the basis using unknown scalars $x, y$ and $z$:

$$A = \begin{pmatrix}4 & -11\\ -11 & -7\end{pmatrix} = x\begin{pmatrix}1 & -2\\ -2 & 1\end{pmatrix} + y\begin{pmatrix}2 & 1\\ 1 & 3\end{pmatrix} + z\begin{pmatrix}4 & -1\\ -1 & -5\end{pmatrix} = \begin{pmatrix}x+2y+4z & -2x+y-z\\ -2x+y-z & x+3y-5z\end{pmatrix}$$

Set corresponding entries equal to each other to obtain the equivalent system of linear equations and reduce to echelon form:

$$\begin{aligned} x+2y+4z &= 4\\ -2x+y-z &= -11\\ -2x+y-z &= -11\\ x+3y-5z &= -7 \end{aligned} \quad \text{or} \quad \begin{aligned} x+2y+4z &= 4\\ 5y+7z &= -3\\ y-9z &= -11 \end{aligned} \quad \text{or} \quad \begin{aligned} x+2y+4z &= 4\\ 5y+7z &= -3\\ 52z &= 52 \end{aligned}$$

We obtain $z = 1$ from the third equation, then $y = -2$ from the second equation, and then $x = 4$ from the first equation. Thus the solution of the system is $x = 4$, $y = -2$, $z = 1$; hence $[A] = (4, -2, 1)$. (Since $\dim W = 3$ by Problem 5.29, the coordinate vector of $A$ must be a vector in $\mathbf{R}^3$.)

**5.39.** Let $\{e_1, e_2, e_3\}$ and $\{f_1, f_2, f_3\}$ be bases of a vector space $V$ (of dimension 3). Suppose

$$\begin{aligned} e_1 &= a_1f_1 + a_2f_2 + a_3f_3\\ e_2 &= b_1f_1 + b_2f_2 + b_3f_3\\ e_3 &= c_1f_1 + c_2f_2 + c_3f_3 \end{aligned} \tag{1}$$

Let $P$ be the matrix whose rows are the coordinate vectors of $e_1$, $e_2$ and $e_3$ respectively, relative to the basis $\{f_i\}$:

$$P = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Show that, for any vector $v \in V$, $[v]_e P = [v]_f$. That is, multiplying the coordinate vector of $v$ relative to the basis $\{e_i\}$ by the matrix $P$, we obtain the coordinate vector of $v$ relative to the basis $\{f_i\}$. (The matrix $P$ is frequently called the change of basis matrix.)

Suppose $v = re_1 + se_2 + te_3$; then $[v]_e = (r, s, t)$. Using (1), we have

$$\begin{aligned} v &= r(a_1f_1 + a_2f_2 + a_3f_3) + s(b_1f_1 + b_2f_2 + b_3f_3) + t(c_1f_1 + c_2f_2 + c_3f_3) \\ &= (ra_1 + sb_1 + tc_1)f_1 + (ra_2 + sb_2 + tc_2)f_2 + (ra_3 + sb_3 + tc_3)f_3 \end{aligned}$$

Hence $$[v]_f = (ra_1 + sb_1 + tc_1,\ ra_2 + sb_2 + tc_2,\ ra_3 + sb_3 + tc_3)$$

On the other hand,

$$\begin{aligned} [v]_e P &= (r, s, t)\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \\ &= (ra_1 + sb_1 + tc_1,\ ra_2 + sb_2 + tc_2,\ ra_3 + sb_3 + tc_3) \end{aligned}$$

Accordingly, $[v]_e P = [v]_f$.

**Remark:** In Chapter 8 we shall write coordinate vectors as column vectors rather than row vectors. Then, by above,

$$Q[v]_e = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 + tc_1 \\ ra_2 + sb_2 + tc_2 \\ ra_3 + sb_3 + tc_3 \end{pmatrix} = [v]_f$$

where $Q$ is the matrix whose columns are the coordinate vectors of $e_1$, $e_2$ and $e_3$ respectively, relative to the basis $\{f_i\}$. Note that $Q$ is the transpose of $P$ and that $Q$ appears on the left of the column vector $[v]_e$ whereas $P$ appears on the right of the row vector $[v]_e$.

## RANK OF A MATRIX

**5.40.** Find the rank of the matrix $A$ where:

(i) $A = \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{pmatrix}$ (iii) $A = \begin{pmatrix} 1 & 3 \\ 0 & -2 \\ 5 & -1 \\ -2 & 3 \end{pmatrix}$

(i) Row reduce to echelon form:

$$\begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & -3 & -6 & -3 & 3 \\ 0 & -1 & -2 & -1 & 1 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the echelon matrix has two nonzero rows, $\operatorname{rank}(A) = 2$.

(ii) Since row rank equals column rank, it is easier to form the transpose of $A$ and then row reduce to echelon form:

$$\begin{pmatrix} 1 & 2 & -2 & -1 \\ 2 & 1 & -1 & 4 \\ -3 & 0 & 3 & -2 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 6 & -3 & -5 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 0 & 3 & 7 \end{pmatrix}$$

Thus $\operatorname{rank}(A) = 3$.

(iii) The two columns are linearly independent since one is not a multiple of the other. Hence $\operatorname{rank}(A) = 2$.

**5.41.** Let $A$ and $B$ be arbitrary matrices for which the product $AB$ is defined. Show that $\operatorname{rank}(AB) \le \operatorname{rank}(B)$ and $\operatorname{rank}(AB) \le \operatorname{rank}(A)$.

By Problem 4.33, page 80, the row space of $AB$ is contained in the row space of $B$; hence $\operatorname{rank}(AB) \le \operatorname{rank}(B)$. Furthermore, by Problem 4.71, page 84, the column space of $AB$ is contained in the column space of $A$; hence $\operatorname{rank}(AB) \le \operatorname{rank}(A)$.

**5.42.** Let $A$ be an $n$-square matrix. Show that $A$ is invertible if and only if $\operatorname{rank}(A) = n$.

Note that the rows of the $n$-square identity matrix $I_n$ are linearly independent since $I_n$ is in echelon form; hence $\operatorname{rank}(I_n) = n$. Now if $A$ is invertible then, by Problem 3.36, page 57, $A$ is row equivalent to $I_n$; hence $\operatorname{rank}(A) = n$. But if $A$ is not invertible then $A$ is row equivalent to a matrix with a zero row; hence $\operatorname{rank}(A) < n$. That is, $A$ is invertible if and only if $\operatorname{rank}(A) = n$.

**5.43.** Let $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ be the free variables of a homogeneous system of linear equations with $n$ unknowns. Let $v_j$ be the solution for which $x_{i_j} = 1$, and all other free variables $= 0$. Show that the solutions $v_1, v_2, \ldots, v_k$ are linearly independent.

Let $A$ be the matrix whose rows are the $v_i$ respectively. We interchange column 1 and column $i_1$, then column 2 and column $i_2$, $\ldots$, and then column $k$ and column $i_k$; and obtain the $k \times n$ matrix

$$B = (I, C) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & c_{1,k+1} & \cdots & c_{1n} \\ 0 & 1 & 0 & \ldots & 0 & 0 & c_{2,k+1} & \cdots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & c_{k,k+1} & \cdots & c_{kn} \end{pmatrix}$$

The above matrix $B$ is in echelon form and so its rows are independent; hence $\operatorname{rank}(B) = k$. Since $A$ and $B$ are column equivalent, they have the same rank, i.e. $\operatorname{rank}(A) = k$. But $A$ has $k$ rows; hence these rows, i.e. the $v_i$, are linearly independent as claimed.

## MISCELLANEOUS PROBLEMS

**5.44.** The concept of linear dependence is extended to every set of vectors, finite or infinite, as follows: the set of vectors $A = \{v_i\}$ is linearly dependent iff there exist vectors $v_{i_1}, \ldots, v_{i_n} \in A$ and scalars $a_1, \ldots, a_n \in K$, not all of them 0, such that

$$a_1 v_{i_1} + a_2 v_{i_2} + \cdots + a_n v_{i_n} = 0$$

Otherwise $A$ is said to be linearly independent. Suppose that $A_1, A_2, \ldots$ are linearly independent sets of vectors, and that $A_1 \subset A_2 \subset \cdots$. Show that the union $A = A_1 \cup A_2 \cup \cdots$ is also linearly independent.

Suppose $A$ is linearly dependent. Then there exist vectors $v_1, \ldots, v_n \in A$ and scalars $a_1, \ldots, a_n \in K$, not all of them 0, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0 \tag{1}$$

Since $A = \cup A_i$ and the $v_i \in A$, there exist sets $A_{i_1}, \ldots, A_{i_n}$ such that

$$v_1 \in A_{i_1}, \quad v_2 \in A_{i_2}, \quad \ldots, \quad v_n \in A_{i_n}$$

Let $k$ be the maximum index of the sets $A_{i_j}$: $k = \max(i_1, \ldots, i_n)$. It follows then, since $A_1 \subset A_2 \subset \cdots$, that each $A_{i_j}$ is contained in $A_k$. Hence $v_1, v_2, \ldots, v_n \in A_k$ and so, by (1), $A_k$ is linearly dependent, which contradicts our hypothesis. Thus $A$ is linearly independent.

**5.45.** Consider a finite sequence of vectors $S = (v_1, v_2, \ldots, v_n)$. Let $T$ be the sequence of vectors obtained from $S$ by one of the following "elementary operations": (i) interchange two vectors, (ii) multiply a vector by a nonzero scalar, (iii) add a multiple of one vector to another. Show that $S$ and $T$ generate the same space $W$. Also show that $T$ is independent if and only if $S$ is independent.

Observe that, for each operation, the vectors in $T$ are linear combinations of vectors in $S$. On the other hand, each operation has an inverse of the same type (Prove!); hence the vectors in $S$ are linear combinations of vectors in $T$. Thus $S$ and $T$ generate the same space $W$. Also, $T$ is independent if and only if $\dim W = n$, and this is true iff $S$ is also independent.

**5.46.** Let $A = (a_{ij})$ and $B = (b_{ij})$ be row equivalent $m \times n$ matrices over a field $K$, and let $v_1, \ldots, v_n$ be any vectors in a vector space $V$ over $K$. Let

$$\begin{array}{lll} u_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n & \qquad & w_1 = b_{11}v_1 + b_{12}v_2 + \cdots + b_{1n}v_n \\ u_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n & & w_2 = b_{21}v_1 + b_{22}v_2 + \cdots + b_{2n}v_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_m = a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n & & w_m = b_{m1}v_1 + b_{m2}v_2 + \cdots + b_{mn}v_n \end{array}$$

Show that $\{u_i\}$ and $\{w_i\}$ generate the same space.

Applying an "elementary operation" of the preceding problem to $\{u_i\}$ is equivalent to applying an elementary row operation to the matrix $A$. Since $A$ and $B$ are row equivalent, $B$ can be obtained from $A$ by a sequence of elementary row operations; hence $\{w_i\}$ can be obtained from $\{u_i\}$ by the corresponding sequence of operations. Accordingly, $\{u_i\}$ and $\{w_i\}$ generate the same space.

**5.47.** Let $v_1, \ldots, v_n$ belong to a vector space $V$ over a field $K$. Let

$$\begin{aligned} w_1 &= a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ w_2 &= a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_n &= a_{n1}v_1 + a_{n2}v_2 + \cdots + a_{nn}v_n \end{aligned}$$

where $a_{ij} \in K$. Let $P$ be the $n$-square matrix of coefficients, i.e. let $P = (a_{ij})$.

(i) Suppose $P$ is invertible. Show that $\{w_i\}$ and $\{v_i\}$ generate the same space; hence $\{w_i\}$ is independent if and only if $\{v_i\}$ is independent.

(ii) Suppose $P$ is not invertible. Show that $\{w_i\}$ is dependent.

(iii) Suppose $\{w_i\}$ is independent. Show that $P$ is invertible.

(i) Since $P$ is invertible, it is row equivalent to the identity matrix $I$. Hence by the preceding problem $\{w_i\}$ and $\{v_i\}$ generate the same space. Thus one is independent if and only if the other is.

(ii) Since $P$ is not invertible, it is row equivalent to a matrix with a zero row. This means that $\{w_i\}$ generates a space which has a generating set of less than $n$ elements. Thus $\{w_i\}$ is dependent.

(iii) This is the contrapositive of the statement of (ii), and so it follows from (ii).

**5.48.** Suppose $V$ is the direct sum of its subspaces $U$ and $W$, i.e. $V = U \oplus W$. Show that: (i) if $\{u_1, \ldots, u_m\} \subset U$ and $\{w_1, \ldots, w_n\} \subset W$ are independent, then $\{u_i, w_j\}$ is also independent; (ii) $\dim V = \dim U + \dim W$.

(i) Suppose $a_1u_1 + \cdots + a_mu_m + b_1w_1 + \cdots + b_nw_n = 0$, where $a_i, b_j$ are scalars. Then

$$0 = (a_1u_1 + \cdots + a_mu_m) + (b_1w_1 + \cdots + b_nw_n) = 0 + 0$$

where $0, a_1u_1 + \cdots + a_mu_m \in U$ and $0, b_1w_1 + \cdots + b_nw_n \in W$. Since such a sum for 0 is unique, this leads to

$$a_1u_1 + \cdots + a_mu_m = 0, \qquad b_1w_1 + \cdots + b_nw_n = 0$$

The independence of the $u_i$ implies that the $a_i$ are all 0, and the independence of the $w_j$ implies that the $b_j$ are all 0. Consequently, $\{u_i, w_j\}$ is independent.

(ii) Method 1. Since $V = U \oplus W$, we have $V = U + W$ and $U \cap W = \{0\}$. Thus, by Theorem 5.8, page 90,

$$\dim V = \dim U + \dim W - \dim(U \cap W) = \dim U + \dim W - 0 = \dim U + \dim W$$

Method 2. Suppose $\{u_1, \ldots, u_r\}$ and $\{w_1, \ldots, w_s\}$ are bases of $U$ and $W$ respectively. Since they generate $U$ and $W$ respectively, $\{u_i, w_j\}$ generates $V = U + W$. On the other hand, by (i), $\{u_i, w_j\}$ is independent. Thus $\{u_i, w_j\}$ is a basis of $V$; hence $\dim V = \dim U + \dim W$.

**5.49.** Let $U$ be a subspace of a vector space $V$ of finite dimension. Show that there exists a subspace $W$ of $V$ such that $V = U \oplus W$.

Let $\{u_1, \ldots, u_r\}$ be a basis of $U$. Since $\{u_i\}$ is linearly independent, it can be extended to a basis of $V$, say, $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$. Let $W$ be the space generated by $\{w_1, \ldots, w_s\}$. Since $\{u_i, w_j\}$ generates $V$, $V = U + W$. On the other hand, $U \cap W = \{0\}$ (Problem 5.62). Accordingly, $V = U \oplus W$.

**5.50.** Recall (page 65) that if $K$ is a subfield of a field $E$ (or: $E$ is an extension of $K$), then $E$ may be viewed as a vector space over $K$. (i) Show that the complex field $\mathbf{C}$ is a vector space of dimension 2 over the real field $\mathbf{R}$. (ii) Show that the real field $\mathbf{R}$ is a vector space of infinite dimension over the rational field $\mathbf{Q}$.

(i) We claim that $\{1, i\}$ is a basis of $\mathbf{C}$ over $\mathbf{R}$. For if $v \in \mathbf{C}$, then $v = a + bi = a \cdot 1 + b \cdot i$ where $a, b \in \mathbf{R}$; that is, $\{1, i\}$ generates $\mathbf{C}$ over $\mathbf{R}$. Furthermore, if $x \cdot 1 + y \cdot i = 0$ or $x + yi = 0$, where $x, y \in \mathbf{R}$, then $x = 0$ and $y = 0$; that is, $\{1, i\}$ is linearly independent over $\mathbf{R}$. Thus $\{1, i\}$ is a basis of $\mathbf{C}$ over $\mathbf{R}$, and so $\mathbf{C}$ is of dimension 2 over $\mathbf{R}$.

(ii) We claim that, for any $n$, $\{1, \pi, \pi^2, \ldots, \pi^n\}$ is linearly independent over $\mathbf{Q}$. For suppose $a_0 1 + a_1\pi + a_2\pi^2 + \cdots + a_n\pi^n = 0$, where the $a_i \in \mathbf{Q}$, and not all the $a_i$ are 0. Then $\pi$ is a root of the following nonzero polynomial over $\mathbf{Q}$: $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. But it can be shown that $\pi$ is a transcendental number, i.e. that $\pi$ is not a root of any nonzero polynomial over $\mathbf{Q}$. Accordingly, the $n+1$ real numbers $1, \pi, \pi^2, \ldots, \pi^n$ are linearly independent over $\mathbf{Q}$. Thus for any finite $n$, $\mathbf{R}$ cannot be of dimension $n$ over $\mathbf{Q}$, i.e. $\mathbf{R}$ is of infinite dimension over $\mathbf{Q}$.

**5.51.** Let $K$ be a subfield of a field $L$ and $L$ a subfield of a field $E$: $K \subset L \subset E$. (Hence $K$ is a subfield of $E$.) Suppose that $E$ is of dimension $n$ over $L$ and $L$ is of dimension $m$ over $K$. Show that $E$ is of dimension $mn$ over $K$.

Suppose $\{v_1, \ldots, v_n\}$ is a basis of $E$ over $L$ and $\{a_1, \ldots, a_m\}$ is a basis of $L$ over $K$. We claim that $\{a_iv_j : i = 1, \ldots, m, j = 1, \ldots, n\}$ is a basis of $E$ over $K$. Note that $\{a_iv_j\}$ contains $mn$ elements.

Let $w$ be any arbitrary element in $E$. Since $\{v_1, \ldots, v_n\}$ generates $E$ over $L$, $w$ is a linear combination of the $v_i$ with coefficients in $L$:

$$w = b_1v_1 + b_2v_2 + \cdots + b_nv_n, \qquad b_i \in L \tag{1}$$

Since $\{a_1, \ldots, a_m\}$ generates $L$ over $K$, each $b_i \in L$ is a linear combination of the $a_j$ with coefficients in $K$:

$$\begin{aligned} b_1 &= k_{11}a_1 + k_{12}a_2 + \cdots + k_{1m}a_m \\ b_2 &= k_{21}a_1 + k_{22}a_2 + \cdots + k_{2m}a_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_n &= k_{n1}a_1 + k_{n2}a_2 + \cdots + k_{nm}a_m \end{aligned}$$

where $k_{ij} \in K$. Substituting in (1), we obtain

$$\begin{aligned} w &= (k_{11}a_1 + \cdots + k_{1m}a_m)v_1 + (k_{21}a_1 + \cdots + k_{2m}a_m)v_2 + \cdots + (k_{n1}a_1 + \cdots + k_{nm}a_m)v_n \\ &= k_{11}a_1v_1 + \cdots + k_{1m}a_mv_1 + k_{21}a_1v_2 + \cdots + k_{2m}a_mv_2 + \cdots + k_{n1}a_1v_n + \cdots + k_{nm}a_mv_n \\ &= \sum_{i,j} k_{ji}(a_iv_j) \end{aligned}$$

where $k_{ji} \in K$. Thus $w$ is a linear combination of the $a_iv_j$ with coefficients in $K$; hence $\{a_iv_j\}$ generates $E$ over $K$.

The proof is complete if we show that $\{a_iv_j\}$ is linearly independent over $K$. Suppose, for scalars $x_{ji} \in K$, $\sum_{i,j} x_{ji}(a_iv_j) = 0$; that is,

$$(x_{11}a_1v_1 + x_{12}a_2v_1 + \cdots + x_{1m}a_mv_1) + \cdots + (x_{n1}a_1v_n + x_{n2}a_2v_n + \cdots + x_{nm}a_mv_n) = 0$$

or
$$(x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m)v_1 + \cdots + (x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m)v_n = 0$$

Since $\{v_1, \ldots, v_n\}$ is linearly independent over $L$ and since the above coefficients of the $v_i$ belong to $L$, each coefficient must be 0:

$$x_{11}a_1 + x_{12}a_2 + \cdots + x_{1m}a_m = 0, \quad \ldots, \quad x_{n1}a_1 + x_{n2}a_2 + \cdots + x_{nm}a_m = 0$$

But $\{a_1, \ldots, a_m\}$ is linearly independent over $K$; hence since the $x_{ji} \in K$,

$$x_{11} = 0, \; x_{12} = 0, \; \ldots, \; x_{1m} = 0, \; \ldots, \; x_{n1} = 0, \; x_{n2} = 0, \; \ldots, \; x_{nm} = 0$$

Accordingly, $\{a_iv_j\}$ is linearly independent over $K$ and the theorem is proved.

## Supplementary Problems

### LINEAR DEPENDENCE

**5.52.** Determine whether $u$ and $v$ are linearly dependent where:

(i) $u = (1, 2, 3, 4)$, $v = (4, 3, 2, 1)$

(ii) $u = (-1, 6, -12)$, $v = (\frac{1}{2}, -3, 6)$

(iii) $u = (0, 1)$, $v = (0, -3)$

(iv) $u = (1, 0, 0)$, $v = (0, 0, -3)$

(v) $u = \begin{pmatrix} 4 & -2 \\ 0 & -1 \end{pmatrix}$, $v = \begin{pmatrix} -2 & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$

(vi) $u = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $v = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

(vii) $u = -t^3 + \frac{1}{2}t^2 - 16$, $v = \frac{1}{2}t^3 - \frac{1}{4}t^2 + 8$

(viii) $u = t^3 + 3t + 4$, $v = t^3 + 4t + 3$

**5.53.** Determine whether the following vectors in $\mathbf{R}^4$ are linearly dependent or independent: (i) $(1, 3, -1, 4)$, $(3, 8, -5, 7)$, $(2, 9, 4, 23)$; (ii) $(1, -2, 4, 1)$, $(2, 1, 0, -3)$, $(3, -6, 1, 4)$.

**5.54.** Let $V$ be the vector space of $2 \times 3$ matrices over $\mathbf{R}$. Determine whether the matrices $A, B, C \in V$ are linearly dependent or independent where:

(i) $A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 4 \\ 4 & 5 & -2 \end{pmatrix}$, $C = \begin{pmatrix} 3 & -8 & 7 \\ 2 & 10 & -1 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & -2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & -3 \\ -2 & 0 & 5 \end{pmatrix}$, $C = \begin{pmatrix} 4 & -1 & 2 \\ 1 & -2 & -3 \end{pmatrix}$

**5.55.** Let $V$ be the vector space of polynomials of degree $\leq 3$ over $\mathbf{R}$. Determine whether $u, v, w \in V$ are linearly dependent or independent where:

(i) $u = t^3 - 4t^2 + 2t + 3$, $v = t^3 + 2t^2 + 4t - 1$, $w = 2t^3 - t^2 - 3t + 5$

(ii) $u = t^3 - 5t^2 - 2t + 3$, $v = t^3 - 4t^2 - 3t + 4$, $w = 2t^3 - 7t^2 - 7t + 9$

**5.56.** Let $V$ be the vector space of functions from $\mathbf{R}$ into $\mathbf{R}$. Show that $f, g, h \in V$ are linearly independent where: (i) $f(t) = e^t$, $g(t) = \sin t$, $h(t) = t^2$; (ii) $f(t) = e^t$, $g(t) = e^{2t}$, $h(t) = t$; (iii) $f(t) = e^t$, $g(t) = \sin t$, $h(t) = \cos t$.

**5.57.** Show that: (i) the vectors $(1-i, i)$ and $(2, -1+i)$ in $\mathbf{C}^2$ are linearly dependent over the complex field $\mathbf{C}$ but are linearly independent over the real field $\mathbf{R}$; (ii) the vectors $(3+\sqrt{2}, 1+\sqrt{2})$ and $(7, 1+2\sqrt{2})$ in $\mathbf{R}^2$ are linearly dependent over the real field $\mathbf{R}$ but are linearly independent over the rational field $\mathbf{Q}$.

**5.58.** Suppose $u$, $v$ and $w$ are linearly independent vectors. Show that:

(i) $u+v-2w$, $u-v-w$ and $u+w$ are linearly independent;

(ii) $u+v-3w$, $u+3v-w$ and $v+w$ are linearly dependent.

**5.59.** Prove or show a counterexample: If the nonzero vectors $u$, $v$ and $w$ are linearly dependent, then $w$ is a linear combination of $u$ and $v$.

**5.60.** Suppose $v_1, v_2, \ldots, v_n$ are linearly independent vectors. Prove the following:

(i) $\{a_1v_1, a_2v_2, \ldots, a_nv_n\}$ is linearly independent where each $a_i \neq 0$.

(ii) $\{v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n\}$ is linearly independent where $w = b_1v_1 + \cdots + b_iv_i + \cdots + b_nv_n$ and $b_i \neq 0$.

**5.61.** Let $v = (a, b)$ and $w = (c, d)$ belong to $K^2$. Show that $\{v, w\}$ is linearly dependent if and only if $ad - bc = 0$.

**5.62.** Suppose $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ is a linearly independent subset of a vector space $V$. Show that $L(u_i) \cap L(w_j) = \{0\}$. (Recall that $L(u_i)$ is the linear span, i.e. the space generated by the $u_i$.)

**5.63.** Suppose $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn})$ are linearly independent vectors in $K^n$, and suppose $v_1, \ldots, v_n$ are linearly independent vectors in a vector space $V$ over $K$. Show that the vectors

$$w_1 = a_{11}v_1 + \cdots + a_{1n}v_n, \quad \ldots, \quad w_m = a_{m1}v_1 + \cdots + a_{mn}v_n$$

are also linearly independent.

## BASIS AND DIMENSION

**5.64.** Determine whether or not each of the following forms a basis of $\mathbf{R}^2$:

(i) $(1, 1)$ and $(3, 1)$ (iii) $(0, 1)$ and $(0, -3)$

(ii) $(2, 1)$, $(1, -1)$ and $(0, 2)$ (iv) $(2, 1)$ and $(-3, 87)$

**5.65.** Determine whether or not each of the following forms a basis of $\mathbf{R}^3$:

(i) $(1, 2, -1)$ and $(0, 3, 1)$

(ii) $(2, 4, -3)$, $(0, 1, 1)$ and $(0, 1, -1)$

(iii) $(1, 5, -6)$, $(2, 1, 8)$, $(3, -1, 4)$ and $(2, 1, 1)$

(iv) $(1, 3, -4)$, $(1, 4, -3)$ and $(2, 3, -11)$

**5.66.** Find a basis and the dimension of the subspace $W$ of $\mathbf{R}^4$ generated by:

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

(ii) $(1, -4, -2, 1)$, $(1, -3, -1, 2)$ and $(3, -8, -2, 7)$

A 92/

**5.67.** Let $V$ be the space of $2 \times 2$ matrices over $\mathbf{R}$ and let $W$ be the subspace generated by

$$\begin{pmatrix} 1 & -5 \\ -4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ -1 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & -4 \\ -5 & 7 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & -7 \\ -5 & 1 \end{pmatrix}$$

Find a basis and the dimension of $W$.

**5.68.** Let $W$ be the space generated by the polynomials

$$u = t^3 + 2t^2 - 2t + 1, \quad v = t^3 + 3t^2 - t + 4 \quad \text{and} \quad w = 2t^3 + t^2 - 7t - 7$$

Find a basis and the dimension of $W$.

5.69. Find a basis and the dimension of the solution space $W$ of each homogeneous system:

$$\begin{array}{ccc} x+3y+2z=0 & x-2y+7z=0 & \\ x+5y+z=0 & 2x+3y-2z=0 & x+4y+2z=0 \\ 3x+5y+8z=0 & 2x-y+z=0 & 2x+y+5z=0 \\ \text{(i)} & \text{(ii)} & \text{(iii)} \end{array}$$

5.70. Find a basis and the dimension of the solution space $W$ of each homogeneous system:

$$\begin{array}{cc} x+2y-2z+2s-t=0 & x+2y-z+3s-4t=0 \\ x+2y-z+3s-2t=0 & 2x+4y-2z-s+5t=0 \\ 2x+4y-7z+s+t=0 & 2x+4y-2z+4s-2t=0 \\ \text{(i)} & \text{(ii)} \end{array}$$

5.71. Find a homogeneous system whose solution set $W$ is generated by

$$\{(1,-2,0,3,-1),\ (2,-3,2,5,-3),\ (1,-2,1,2,-2)\}$$

5.27 90 v.v.9 5.72. Let $V$ and $W$ be the following subspaces of $\mathbf{R}^4$:

$$V = \{(a,b,c,d):\ b-2c+d=0\}, \qquad W = \{(a,b,c,d):\ a=d,\ b=2c\}$$

Find a basis and the dimension of (i) $V$, (ii) $W$, (iii) $V \cap W$.

5.73. Let $V$ be the vector space of polynomials in $t$ of degree $\le n$. Determine whether or not each of the following is a basis of $V$:

(i) $\{1,\ 1+t,\ 1+t+t^2,\ 1+t+t^2+t^3,\ \ldots,\ 1+t+t^2+\cdots+t^{n-1}+t^n\}$

(ii) $\{1+t,\ t+t^2,\ t^2+t^3,\ \ldots,\ t^{n-2}+t^{n-1},\ t^{n-1}+t^n\}$.

## SUMS AND INTERSECTIONS

5.74. Suppose $U$ and $W$ are 2-dimensional subspaces of $\mathbf{R}^3$. Show that $U \cap W \neq \{0\}$.

5.75. Suppose $U$ and $W$ are subspaces of $V$ and that $\dim U = 4$, $\dim W = 5$ and $\dim V = 7$. Find the possible dimensions of $U \cap W$.

5.76. Let $U$ and $W$ be subspaces of $\mathbf{R}^3$ for which $\dim U = 1$, $\dim W = 2$ and $U \not\subset W$. Show that $\mathbf{R}^3 = U \oplus W$.

5.77. Let $U$ be the subspace of $\mathbf{R}^5$ generated by

$$\{(1,3,-3,-1,-4),\ (1,4,-1,-2,-2),\ (2,9,0,-5,-2)\}$$

and let $W$ be the subspace generated by

$$\{(1,6,2,-2,3),\ (2,8,-1,-6,-5),\ (1,3,-1,-5,-6)\}$$

Find (i) $\dim(U+W)$, (ii) $\dim(U \cap W)$.

5.24 91 5.78. Let $V$ be the vector space of polynomials over $\mathbf{R}$. Let $U$ and $W$ be the subspaces generated by $\{t^3+4t^2-t+3,\ t^3+5t^2+5,\ 3t^3+10t^2-5t+5\}$ and $\{t^3+4t^2+6,\ t^3+2t^2-t+5,\ 2t^3+2t^2-3t+9\}$ respectively. Find (i) $\dim(U+W)$, (ii) $\dim(U \cap W)$.

5.79. Let $U$ be the subspace of $\mathbf{R}^5$ generated by

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

and let $W$ be the subspace generated by

$$\{(1,-2,-3,0,-2),\ (1,-1,-3,2,-4),\ (1,-1,-2,2,-5)\}$$

(i) Find two homogeneous systems whose solution spaces are $U$ and $W$, respectively.

(ii) Find a basis and the dimension of $U \cap W$.

### COORDINATE VECTORS

**5.80.** Consider the following basis of $\mathbf{R}^2$: $\{(2,1), (1,-1)\}$. Find the coordinate vector of $v \in \mathbf{R}^2$ relative to the above basis where: (i) $v = (2,3)$; (ii) $v = (4,-1)$, (iii) $(3,-3)$; (iv) $v = (a,b)$.

**5.81.** In the vector space $V$ of polynomials in $t$ of degree $\leq 3$, consider the following basis: $\{1, 1-t, (1-t)^2, (1-t)^3\}$. Find the coordinate vector of $v \in V$ relative to the above basis if: (i) $v = 2 - 3t + t^2 + 2t^3$; (ii) $v = 3 - 2t - t^2$; (iii) $v = a + bt + ct^2 + dt^3$.

**5.82.** In the vector space $W$ of $2 \times 2$ symmetric matrices over $\mathbf{R}$, consider the following basis:

$$\left\{\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix}\right\}$$

Find the coordinate vector of the matrix $A \in W$ relative to the above basis if:

(i) $A = \begin{pmatrix} 1 & -5 \\ -5 & 5 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

**5.83.** Consider the following two bases of $\mathbf{R}^3$:

$$\{e_1 = (1,1,1),\ e_2 = (0,2,3),\ e_3 = (0,2,-1)\} \quad \text{and} \quad \{f_1 = (1,1,0),\ f_2 = (1,-1,0),\ f_3 = (0,0,1)\}$$

(i) Find the coordinate vector of $v = (3,5,-2)$ relative to each basis: $[v]_e$ and $[v]_f$.

(ii) Find the matrix $P$ whose rows are respectively the coordinate vectors of the $e_i$ relative to the basis $\{f_1, f_2, f_3\}$.

(iii) Verify that $[v]_e P = [v]_f$.

**5.84.** Suppose $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_n\}$ are bases of a vector space $V$ (of dimension $n$). Let $P$ be the matrix whose rows are respectively the coordinate vectors of the $e$'s relative to the basis $\{f_i\}$. Prove that for any vector $v \in V$, $[v]_e P = [v]_f$. (This result is proved in Problem 5.39 in the case $n = 3$.)

**5.85.** Show that the coordinate vector of $0 \in V$ relative to any basis of $V$ is always the zero $n$-tuple $(0, 0, \ldots, 0)$.

### RANK OF A MATRIX

**5.86.** Find the rank of each matrix:

$$\underset{\text{(i)}}{\begin{pmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{pmatrix}} \quad \underset{\text{(ii)}}{\begin{pmatrix} 1 & 2 & -3 & -2 & -3 \\ 1 & 3 & -2 & 0 & -4 \\ 3 & 8 & -7 & -2 & -11 \\ 2 & 1 & -9 & -10 & -3 \end{pmatrix}} \quad \underset{\text{(iii)}}{\begin{pmatrix} 1 & 1 & 2 \\ 4 & 5 & 5 \\ 5 & 8 & 1 \\ -1 & -2 & 2 \end{pmatrix}} \quad \underset{\text{(iv)}}{\begin{pmatrix} 2 & 1 \\ 3 & -7 \\ -6 & 1 \\ 5 & -8 \end{pmatrix}}$$

**5.87.** Let $A$ and $B$ be arbitrary $m \times n$ matrices. Show that $\operatorname{rank}(A+B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$.

**5.88.** Give examples of $2 \times 2$ matrices $A$ and $B$ such that:

(i) $\operatorname{rank}(A+B) < \operatorname{rank}(A), \operatorname{rank}(B)$ (ii) $\operatorname{rank}(A+B) = \operatorname{rank}(A) = \operatorname{rank}(B)$

(iii) $\operatorname{rank}(A+B) > \operatorname{rank}(A), \operatorname{rank}(B)$

### MISCELLANEOUS PROBLEMS

**5.89.** Let $W$ be the vector space of $3 \times 3$ symmetric matrices over $K$. Show that $\dim W = 6$ by exhibiting a basis of $W$. (Recall that $A = (a_{ij})$ is symmetric iff $a_{ij} = a_{ji}$.)

**5.90.** Let $W$ be the vector space of $3 \times 3$ antisymmetric matrices over $K$. Show that $\dim W = 3$ by exhibiting a basis of $W$. (Recall that $A = (a_{ij})$ is antisymmetric iff $a_{ij} = -a_{ji}$.)

**5.91.** Suppose $\dim V = n$. Show that a generating set with $n$ elements is a basis. (Compare with Theorem 5.6(iii), page 89).

**5.92.** Let $t_1, t_2, \ldots, t_n$ be symbols, and let $K$ be any field. Let $V$ be the set of expressions $a_1t_1 + a_2t_2 + \cdots + a_nt_n$ where $a_i \in K$. Define addition in $V$ by

$$(a_1t_1 + a_2t_2 + \cdots + a_nt_n) + (b_1t_1 + b_2t_2 + \cdots + b_nt_n)$$
$$= (a_1 + b_1)t_1 + (a_2 + b_2)t_2 + \cdots + (a_n + b_n)t_n$$

Define scalar multiplication on $V$ by

$$k(a_1t_1 + a_2t_2 + \cdots + a_nt_n) = ka_1t_1 + ka_2t_2 + \cdots + ka_nt_n$$

Show that $V$ is a vector space over $K$ with the above operations. Also show that $\{t_1, \ldots, t_n\}$ is a basis of $V$ where, for $i = 1, \ldots, n$,

$$t_i = 0t_1 + \cdots + 0t_{i-1} + 1t_i + 0t_{i+1} + \cdots + 0t_n$$

**5.93.** Let $V$ be a vector space of dimension $n$ over a field $K$, and let $K$ be a vector space of dimension $m$ over a subfield $F$. (Hence $V$ may also be viewed as a vector space over the subfield $F$.) Prove that the dimension of $V$ over $F$ is $mn$.

**5.94.** Let $U$ and $W$ be vector spaces over the same field $K$, and let $V$ be the external direct sum of $U$ and $W$ (see Problem 4.45). Let $\hat{U}$ and $\hat{W}$ be the subspaces of $V$ defined by $\hat{U} = \{(u, 0) : u \in U\}$ and $\hat{W} = \{(0, w) : w \in W\}$.

(i) Show that $U$ is isomorphic to $\hat{U}$ under the correspondence $u \leftrightarrow (u, 0)$, and that $W$ is isomorphic to $\hat{W}$ under the correspondence $w \leftrightarrow (0, w)$.

(ii) Show that $\dim V = \dim U + \dim W$.

**5.95.** Suppose $V = U \oplus W$. Let $\hat{V}$ be the external direct product of $U$ and $W$. Show that $V$ is isomorphic to $\hat{V}$ under the correspondence $v = u + w \leftrightarrow (u, w)$.

## Answers to Supplementary Problems

**5.52.** (i) no, (ii) yes, (iii) yes, (iv) no, (v) yes, (vi) no, (vii) yes, (viii) no.

**5.53.** (i) dependent, (ii) independent.

**5.54.** (i) dependent, (ii) independent.

**5.55.** (i) independent, (ii) dependent.

**5.57.** (i) $(2, -1+i) = (1+i)(1-i, i)$; (ii) $(7, 1+2\sqrt{2}) = (3-\sqrt{2})(3+\sqrt{2}, 1+\sqrt{2})$.

**5.59.** The statement is false. Counterexample: $u = (1, 0)$, $v = (2, 0)$ and $w = (1, 1)$ in $\mathbf{R}^2$. Lemma 5.2 requires that one of the nonzero vectors $u, v, w$ is a linear combination of the preceding ones. In this case, $v = 2u$.

**5.64.** (i) yes, (ii) no, (iii) no, (iv) yes.

**5.65.** (i) no, (ii) yes, (iii) no, (iv) no.

**5.66.** (i) $\dim W = 3$, (ii) $\dim W = 2$.

**5.67.** $\dim W = 2$.

**5.68.** $\dim W = 2$.

**5.69.** (i) basis, $\{(7, -1, -2)\}$; $\dim W = 1$. (ii) $\dim W = 0$. (iii) basis, $\{(18, -1, -7)\}$; $\dim W = 1$.

**5.70.** (i) basis, $\{(2, -1, 0, 0, 0), (4, 0, 1, -1, 0), (3, 0, 1, 0, 1)\}$; $\dim W = 3$.

(ii) basis, $\{(2, -1, 0, 0, 0), (1, 0, 1, 0, 0)\}$; $\dim W = 2$.

5.71. $\begin{cases} 5x + y - z - s = 0 \\ x + y - z - t = 0 \end{cases}$

5.72. (i) basis, $\{(1,0,0,0), (0,2,1,0), (0,-1,0,1)\}$; $\dim V = 3$.

(ii) basis, $\{(1,0,0,1), (0,2,1,0)\}$; $\dim W = 2$.

(iii) basis, $\{(0,2,1,0)\}$; $\dim (V \cap W) = 1$. *Hint.* $V \cap W$ must satisfy all three conditions on $a, b, c$ and $d$.

5.73. (i) yes, (ii) no. For $\dim V = n+1$, but the set contains only $n$ elements.

5.75. $\dim (U \cap W) = 2$, 3 or 4.

5.77. $\dim (U + W) = 3$, $\dim (U \cap W) = 2$.

5.78. $\dim (U + W) = 3$, $\dim (U \cap W) = 1$.

5.79. (i) $\begin{cases} 3x + 4y - z \phantom{+ s} - t = 0 \\ 4x + 2y \phantom{- z} + s \phantom{- t} = 0 \end{cases}$, $\begin{cases} 4x + 2y \phantom{+ z} - s \phantom{+ t} = 0 \\ 9x + 2y + z \phantom{- s} + t = 0 \end{cases}$

(ii) $\{(1, -2, -5, 0, 0), (0, 0, 1, 0, -1)\}$. $\dim (U \cap W) = 2$.

5.80. (i) $[v] = (5/3, -4/3)$, (ii) $[v] = (1, 2)$, (iii) $[v] = (0, 3)$, (iv) $[v] = ((a+b)/3, (a-2b)/3)$.

5.81. (i) $[v] =$ ([illegible], $-5, 7, -2)$, (ii) $[v] = (0, 4, -1, 0)$, (iii) $[v] = (a+b+c+d, -b-2c-3d, c+3d, -d)$.

5.82. (i) $[A] = (2, -1, 1)$, (ii) $[A] = (3, 1, -2)$.

5.83. (i) $[v]_e = (3, -1, 2)$, $[v]_f = (4, -1, -2)$; (ii) $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 3 \\ 1 & -1 & -1 \end{pmatrix}$.

5.86. (i) 3, (ii) 2, (iii) 3, (iv) 2.

5.88. (i) $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}$ (iii) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

(ii) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$

5.89. $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$

5.90. $\left\{ \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \right\}$

5.93. *Hint.* The proof is identical to that given in Problem 5.48, page 113, for a special case (when $V$ is an extension field of $K$).

# Chapter 6

## Linear Mappings

### MAPPINGS

Let $A$ and $B$ be arbitrary sets. Suppose to each $a \in A$ there is assigned a unique element of $B$; the collection, $f$, of such assignments is called a *function* or *mapping* (or: *map*) from $A$ into $B$, and is written

$$f: A \to B \qquad \text{or} \qquad A \xrightarrow{f} B$$

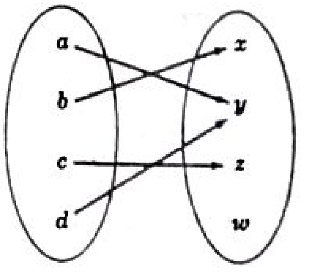
We write $f(a)$, read "$f$ of $a$", for the element of $B$ that $f$ assigns to $a \in A$; it is called the *value* of $f$ at $a$ or the *image of $a$ under $f$*. If $A'$ is any subset of $A$, then $f(A')$ denotes the set of images of elements of $A'$; and if $B'$ is any subset of $B$, then $f^{-1}(B')$ denotes the set of elements of $A$ each of whose image lies in $B'$:

$$f(A') = \{f(a) : a \in A'\} \qquad \text{and} \qquad f^{-1}(B') = \{a \in A : f(a) \in B'\}$$

We call $f(A')$ the *image* of $A'$ and $f^{-1}(B')$ the *inverse image* or *preimage* of $B'$. In particular, the set of all images, i.e. $f(A)$, is called the *image* (or: *range*) of $f$. Furthermore, $A$ is called the *domain* of the mapping $f: A \to B$, and $B$ is called its *co-domain*.

To each mapping $f: A \to B$ there corresponds the subset of $A \times B$ given by $\{(a, f(a)) : a \in A\}$. We call this set the *graph* of $f$. Two mappings $f: A \to B$ and $g: A \to B$ are defined to be *equal*, written $f = g$, if $f(a) = g(a)$ for every $a \in A$, that is, if they have the same graph. Thus we do not distinguish between a function and its graph. The negation of $f = g$ is written $f \neq g$ and is the statement: there exists an $a \in A$ for which $f(a) \neq g(a)$.

**Example 6.1:** Let $A = \{a, b, c, d\}$ and $B = \{x, y, z, w\}$. The following diagram defines a mapping $f$ from $A$ into $B$:




Here $f(a) = y$, $f(b) = x$, $f(c) = z$, and $f(d) = y$. Also,

$$f(\{a, b, d\}) = \{f(a), f(b), f(d)\} = \{y, x, y\} = \{x, y\}$$

The image (or: range) of $f$ is the set $\{x, y, z\}$: $f(A) = \{x, y, z\}$.

**Example 6.2:** Let $f: \mathbf{R} \to \mathbf{R}$ be the mapping which assigns to each real number $x$ its square $x^2$:

$$x \mapsto x^2 \qquad \text{or} \qquad f(x) = x^2$$

Here the image of $-3$ is $9$ so we may write $f(-3) = 9$.

We use the arrow $\mapsto$ to denote the image of an arbitrary element $x \in A$ under a mapping $f: A \to B$ by writing
$$x \mapsto f(x)$$
as illustrated in the preceding example.

**Example 6.3:** Consider the $2 \times 3$ matrix $A = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix}$. If we write the vectors in $\mathbf{R}^3$ and $\mathbf{R}^2$ as column vectors, then $A$ determines the mapping $T: \mathbf{R}^3 \to \mathbf{R}^2$ defined by
$$v \mapsto Av, \quad \text{that is,} \quad T(v) = Av, \quad v \in \mathbf{R}^3$$
Thus if $v = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}$, then $T(v) = Av = \begin{pmatrix} 1 & -3 & 5 \\ 2 & 4 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -10 \\ 12 \end{pmatrix}$.

**Remark:** Every $m \times n$ matrix $A$ over a field $K$ determines the mapping $T: K^n \to K^m$ defined by
$$v \mapsto Av$$
where the vectors in $K^n$ and $K^m$ are written as column vectors. For convenience we shall usually denote the above mapping by $A$, the same symbol used for the matrix.

**Example 6.4:** Let $V$ be the vector space of polynomials in the variable $t$ over the real field $\mathbf{R}$. Then the derivative defines a mapping $D: V \to V$ where, for any polynomial $f \in V$, we let $D(f) = df/dt$. For example, $D(3t^2 - 5t + 2) = 6t - 5$.

**Example 6.5:** Let $V$ be the vector space of polynomials in $t$ over $\mathbf{R}$ (as in the preceding example). Then the integral from, say, 0 to 1 defines a mapping $\mathcal{J}: V \to \mathbf{R}$ where, for any polynomial $f \in V$, we let $\mathcal{J}(f) = \int_0^1 f(t)\,dt$. For example,
$$\mathcal{J}(3t^2 - 5t + 2) = \int_0^1 (3t^2 - 5t + 2)\,dt = \tfrac{1}{2}$$
Note that this map is from the vector space $V$ into the scalar field $\mathbf{R}$, whereas the map in the preceding example is from $V$ into itself.

**Example 6.6:** Consider two mappings $f: A \to B$ and $g: B \to C$ illustrated below:
$$A \xrightarrow{f} B \xrightarrow{g} C$$
Let $a \in A$; then $f(a) \in B$, the domain of $g$. Hence we can obtain the image of $f(a)$ under the mapping $g$, that is, $g(f(a))$. This map
$$a \mapsto g(f(a))$$
from $A$ into $C$ is called the *composition* or *product* of $f$ and $g$, and is denoted by $g \circ f$. In other words, $(g \circ f): A \to C$ is the mapping defined by
$$(g \circ f)(a) = g(f(a))$$

Our first theorem tells us that composition of mappings satisfies the associative law.

**Theorem 6.1:** Let $f: A \to B$, $g: B \to C$ and $h: C \to D$. Then $h \circ (g \circ f) = (h \circ g) \circ f$.

We prove this theorem now. If $a \in A$, then
$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a)))$$
and
$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a)))$$
Thus $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$ for every $a \in A$, and so $h \circ (g \circ f) = (h \circ g) \circ f$.

**Remark:** Let $F: A \to B$. Some texts write $aF$ instead of $F(a)$ for the image of $a \in A$ under $F$. With this notation, the composition of functions $F: A \to B$ and $G: B \to C$ is denoted by $F \circ G$ and not by $G \circ F$ as used in this text.

We next introduce some special types of mappings.

**Definition:** A mapping $f: A \to B$ is said to be *one-to-one* (or one-one or 1-1) or *injective* if different elements of $A$ have distinct images; that is,

$$\text{if } a \neq a' \quad \text{implies} \quad f(a) \neq f(a')$$

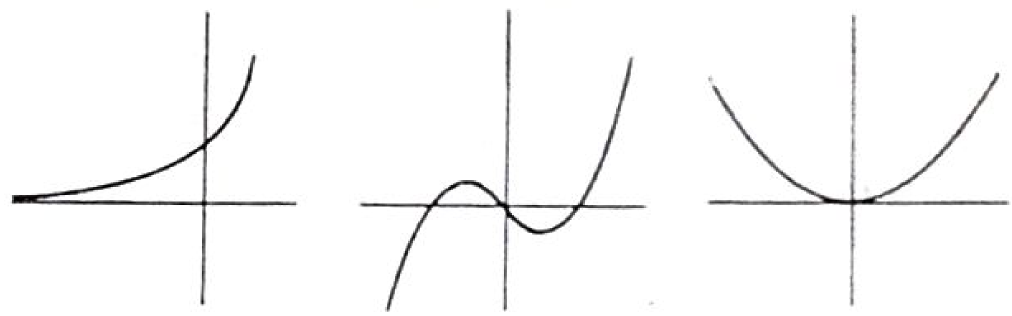
or, equivalently,

$$\text{if } f(a) = f(a') \quad \text{implies} \quad a = a'$$

**Definition:** A mapping $f: A \to B$ is said to be *onto* (or: $f$ maps $A$ *onto* $B$) or *surjective* if every $b \in B$ is the image of at least one $a \in A$.

A mapping which is both one-one and onto is said to be *bijective*.

**Example 6.7:** Let $f: \mathbf{R} \to \mathbf{R}$, $g: \mathbf{R} \to \mathbf{R}$ and $h: \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = 2^x$, $g(x) = x^3 - x$ and $h(x) = x^2$. The graphs of these mappings follow:



$f(x) = 2^x$ $\qquad$ $g(x) = x^3 - x$ $\qquad$ $h(x) = x^2$

The mapping $f$ is one-one; geometrically, this means that each horizontal line does not contain more than one point of $f$. The mapping $g$ is onto; geometrically, this means that each horizontal line contains at least one point of $g$. The mapping $h$ is neither one-one nor onto; for example, 2 and $-2$ have the same image 4, and $-16$ is not the image of any element of $\mathbf{R}$.

**Example 6.8:** Let $A$ be any set. The mapping $f: A \to A$ defined by $f(a) = a$, i.e. which assigns to each element in $A$ itself, is called the *identity mapping* on $A$ and is denoted by $1_A$ or 1 or $I$.

**Example 6.9:** Let $f: A \to B$. We call $g: B \to A$ the *inverse* of $f$, written $f^{-1}$, if

$$f \circ g = 1_B \quad \text{and} \quad g \circ f = 1_A$$

We emphasize that $f$ has an inverse if and only if $f$ is both one-to-one and onto (Problem 6.9). Also, if $b \in B$ then $f^{-1}(b) = a$ where $a$ is the unique element of $A$ for which $f(a) = b$.

## LINEAR MAPPINGS

Let $V$ and $U$ be vector spaces over the same field $K$. A mapping $F: V \to U$ is called a *linear mapping* (or *linear transformation* or *vector space homomorphism*) if it satisfies the following two conditions:

(1) For any $v, w \in V$, $F(v + w) = F(v) + F(w)$.

(2) For any $k \in K$ and any $v \in V$, $F(kv) = kF(v)$.

In other words, $F: V \to U$ is linear if it "preserves" the two basic operations of a vector space, that of vector addition and that of scalar multiplication.

Substituting $k = 0$ into (2) we obtain $F(0) = 0$. That is, every linear mapping takes the zero vector into the zero vector.

Now for any scalars $a, b \in K$ and any vectors $v, w \in V$ we obtain, by applying both conditions of linearity,

$$F(av + bw) = F(av) + F(bw) = aF(v) + bF(w)$$

More generally, for any scalars $a_i \in K$ and any vectors $v_i \in V$ we obtain the basic property of linear mappings:

$$F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n)$$

We remark that the condition $F(av + bw) = aF(v) + bF(w)$ completely characterizes linear mappings and is sometimes used as its definition.

**Example 6.10:** Let $A$ be any $m \times n$ matrix over a field $K$. As noted previously, $A$ determines a mapping $T: K^n \to K^m$ by the assignment $v \mapsto Av$. (Here the vectors in $K^n$ and $K^m$ are written as columns.) We claim that $T$ is linear. For, by properties of matrices,

$$T(v + w) = A(v + w) = Av + Aw = T(v) + T(w)$$

and

$$T(kv) = A(kv) = kAv = kT(v)$$

where $v, w \in K^n$ and $k \in K$.

We comment that the above type of linear mapping shall occur again and again. In fact, in the next chapter we show that every linear mapping from one finite-dimensional vector space into another can be represented as a linear mapping of the above type.

**Example 6.11:** Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the "projection" mapping into the $xy$ plane: $F(x, y, z) = (x, y, 0)$. We show that $F$ is linear. Let $v = (a, b, c)$ and $w = (a', b', c')$. Then

$$F(v + w) = F(a + a', b + b', c + c') = (a + a', b + b', 0)$$
$$= (a, b, 0) + (a', b', 0) = F(v) + F(w)$$

and, for any $k \in \mathbf{R}$,

$$F(kv) = F(ka, kb, kc) = (ka, kb, 0) = k(a, b, 0) = kF(v)$$

That is, $F$ is linear.

**Example 6.12:** Let $F: \mathbf{R}^2 \to \mathbf{R}^2$ be the "translation" mapping defined by $F(x, y) = (x + 1, y + 2)$. Observe that $F(0) = F(0, 0) = (1, 2) \neq 0$. That is, the zero vector is not mapped onto the zero vector. Hence $F$ is not linear.

**Example 6.13:** Let $F: V \to U$ be the mapping which assigns $0 \in U$ to every $v \in V$. Then, for any $v, w \in V$ and any $k \in K$, we have

$$F(v + w) = 0 = 0 + 0 = F(v) + F(w) \quad \text{and} \quad F(kv) = 0 = k0 = kF(v)$$

Thus $F$ is linear. We call $F$ the *zero mapping* and shall usually denote it by 0.

**Example 6.14:** Consider the identity mapping $I: V \to V$ which maps each $v \in V$ into itself. Then, for any $v, w \in V$ and any $a, b \in K$, we have

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Thus $I$ is linear.

**Example 6.15:** Let $V$ be the vector space of polynomials in the variable $t$ over the real field $\mathbf{R}$. Then the differential mapping $D: V \to V$ and the integral mapping $\mathcal{J}: V \to \mathbf{R}$ defined in Examples 6.4 and 6.5 are linear. For it is proven in calculus that for any $u, v \in V$ and $k \in \mathbf{R}$,

$$\frac{d(u + v)}{dt} = \frac{du}{dt} + \frac{dv}{dt} \quad \text{and} \quad \frac{d(ku)}{dt} = k\frac{du}{dt}$$

that is, $D(u + v) = D(u) + D(v)$ and $D(ku) = k\,D(u)$; and also,

$$\int_0^1 (u(t) + v(t))\,dt = \int_0^1 u(t)\,dt + \int_0^1 v(t)\,dt$$

and

$$\int_0^1 k\,u(t)\,dt = k\int_0^1 u(t)\,dt$$

that is, $\mathcal{J}(u + v) = \mathcal{J}(u) + \mathcal{J}(v)$ and $\mathcal{J}(ku) = k\,\mathcal{J}(u)$.

**Example 6.16:** Let $F: V \to U$ be a linear mapping which is both one-one and onto. Then an inverse mapping $F^{-1}: U \to V$ exists. We will show (Problem 6.17) that this inverse mapping is also linear.

When we investigated the coordinates of a vector relative to a basis, we also introduced the notion of two spaces being isomorphic. We now give a formal definition.

**Definition:** A linear mapping $F: V \to U$ is called an *isomorphism* if it is one-to-one. The vector spaces $V$, $U$ are said to be *isomorphic* if there is an isomorphism of $V$ onto $U$.

**Example 6.17:** Let $V$ be a vector space over $K$ of dimension $n$ and let $\{e_1, \ldots, e_n\}$ be a basis of $V$. Then as noted previously the mapping $v \mapsto [v]_e$, i.e. which maps each $v \in V$ into its coordinate vector relative to the basis $\{e_i\}$, is an isomorphism of $V$ onto $K^n$.

Our next theorem gives us an abundance of examples of linear mappings; in particular, it tells us that a linear mapping is completely determined by its values on the elements of a basis.

**Theorem 6.2:** Let $V$ and $U$ be vector spaces over a field $K$. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of $V$ and let $u_1, u_2, \ldots, u_n$ be any vectors in $U$. Then there exists a unique linear mapping $F: V \to U$ such that $F(v_1) = u_1$, $F(v_2) = u_2$, $\ldots$, $F(v_n) = u_n$.

We emphasize that the vectors $u_1, \ldots, u_n$ in the preceding theorem are completely arbitrary; they may be linearly dependent or they may even be equal to each other.

## KERNEL AND IMAGE OF A LINEAR MAPPING

We begin by defining two concepts.

**Definition:** Let $F: V \to U$ be a linear mapping. The *image* of $F$, written $\operatorname{Im} F$, is the set of image points in $U$:

$$\operatorname{Im} F = \{u \in U : F(v) = u \text{ for some } v \in V\}$$

The *kernel* of $F$, written $\operatorname{Ker} F$, is the set of elements in $V$ which map into $0 \in U$:

$$\operatorname{Ker} F = \{v \in V : F(v) = 0\}$$

The following theorem is easily proven (Problem 6.22).

**Theorem 6.3:** Let $F: V \to U$ be a linear mapping. Then the image of $F$ is a subspace of $U$ and the kernel of $F$ is a subspace of $V$.

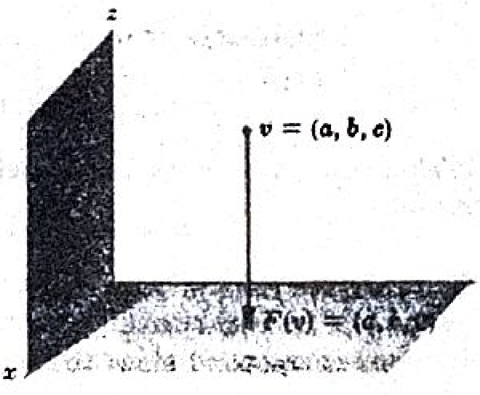
**Example 6.18:** Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the projection mapping into the $xy$ plane: $F(x, y, z) = (x, y, 0)$. Clearly the image of $F$ is the entire $xy$ plane:

$$\operatorname{Im} F = \{(a, b, 0) : a, b \in \mathbf{R}\}$$

Note that the kernel of $F$ is the $z$ axis:

$$\operatorname{Ker} F = \{(0, 0, c) : c \in \mathbf{R}\}$$

since these points and only these points map into the zero vector $0 = (0, 0, 0)$.



Now for any scalars $a, b \in K$ and any vectors $v, w \in V$ we obtain, by applying both conditions of linearity,

$$F(av + bw) = F(av) + F(bw) = aF(v) + bF(w)$$

More generally, for any scalars $a_i \in K$ and any vectors $v_i \in V$ we obtain the basic property of linear mappings:

$$F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n)$$

We remark that the condition $F(av + bw) = aF(v) + bF(w)$ completely characterizes linear mapping and is sometimes used as its definition.

**Example 6.10:** Let $A$ be any $m \times n$ matrix over a field $K$. As noted previously, $A$ determines a mapping $T: K^n \to K^m$ by the assignment $v \mapsto Av$. (Here the vectors in $K^n$ and $K^m$ are written as columns.) We claim that $T$ is linear. For, by properties of matrices,

$$T(v + w) = A(v + w) = Av + Aw = T(v) + T(w)$$

and

$$T(kv) = A(kv) = kAv = kT(v)$$

where $v, w \in K^n$ and $k \in K$.

We comment that the above type of linear mapping shall occur again and again. In fact, in the next chapter we show that every linear mapping from one finite-dimensional vector space into another can be represented as a linear mapping of the above type.

**Example 6.11:** Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the "projection" mapping into the $xy$ plane: $F(x, y, z) = (x, y, 0)$. We show that $F$ is linear. Let $v = (a, b, c)$ and $w = (a', b', c')$. Then

$$F(v + w) = F(a + a', b + b', c + c') = (a + a', b + b', 0)$$
$$= (a, b, 0) + (a', b', 0) = F(v) + F(w)$$

and, for any $k \in \mathbf{R}$,

$$F(kv) = F(ka, kb, kc) = (ka, kb, 0) = k(a, b, 0) = kF(v)$$

That is, $F$ is linear.

**Example 6.12:** Let $F: \mathbf{R}^2 \to \mathbf{R}^2$ be the "translation" mapping defined by $F(x, y) = (x + 1, y + 2)$. Observe that $F(0) = F(0, 0) = (1, 2) \neq 0$. That is, the zero vector is not mapped onto the zero vector. Hence $F$ is not linear.

**Example 6.13:** Let $F: V \to U$ be the mapping which assigns $0 \in U$ to every $v \in V$. Then, for any $v, w \in V$ and any $k \in K$, we have

$$F(v + w) = 0 = 0 + 0 = F(v) + F(w) \quad \text{and} \quad F(kv) = 0 = k0 = kF(v)$$

Thus $F$ is linear. We call $F$ the *zero mapping* and shall usually denote it by $0$.

**Example 6.14:** Consider the identity mapping $I: V \to V$ which maps each $v \in V$ into itself. Then, for any $v, w \in V$ and any $a, b \in K$, we have

$$I(av + bw) = av + bw = aI(v) + bI(w)$$

Thus $I$ is linear.

**Example 6.15:** Let $V$ be the vector space of polynomials in the variable $t$ over the real field $\mathbf{R}$. Then the differential mapping $D: V \to V$ and the integral mapping $\mathcal{J}: V \to \mathbf{R}$ defined in Examples 6.4 and 6.5 are linear. For it is proven in calculus that for any $u, v \in V$ and $k \in \mathbf{R}$,

$$\frac{d(u + v)}{dt} = \frac{du}{dt} + \frac{dv}{dt} \quad \text{and} \quad \frac{d(ku)}{dt} = k\frac{du}{dt}$$

that is, $D(u + v) = D(u) + D(v)$ and $D(ku) = k\,D(u)$; and also,

$$\int_0^1 (u(t) + v(t))\,dt = \int_0^1 u(t)\,dt + \int_0^1 v(t)\,dt$$

and

$$\int_0^1 k\,u(t)\,dt = k\int_0^1 u(t)\,dt$$

that is, $\mathcal{J}(u + v) = \mathcal{J}(u) + \mathcal{J}(v)$ and $\mathcal{J}(ku) = k\,\mathcal{J}(u)$.

**Example 6.16:** Let $F: V \to U$ be a linear mapping which is both one-one and onto. Then an inverse mapping $F^{-1}: U \to V$ exists. We will show (Problem 6.17) that this inverse mapping is also linear.

When we investigated the coordinates of a vector relative to a basis, we also introduced the notion of two spaces being isomorphic. We now give a formal definition.

**Definition:** A linear mapping $F: V \to U$ is called an *isomorphism* if it is one-to-one. The vector spaces $V$, $U$ are said to be *isomorphic* if there is an isomorphism of $V$ onto $U$.

**Example 6.17:** Let $V$ be a vector space over $K$ of dimension $n$ and let $\{e_1, \ldots, e_n\}$ be a basis of $V$. Then as noted previously the mapping $v \mapsto [v]_e$, i.e. which maps each $v \in V$ into its coordinate vector relative to the basis $\{e_i\}$, is an isomorphism of $V$ onto $K^n$.

Our next theorem gives us an abundance of examples of linear mappings; in particular, it tells us that a linear mapping is completely determined by its values on the elements of a basis.

**Theorem 6.2:** Let $V$ and $U$ be vector spaces over a field $K$. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of $V$ and let $u_1, u_2, \ldots, u_n$ be any vectors in $U$. Then there exists a unique linear mapping $F: V \to U$ such that $F(v_1) = u_1$, $F(v_2) = u_2$, $\ldots$, $F(v_n) = u_n$.

v.v.9

We emphasize that the vectors $u_1, \ldots, u_n$ in the preceding theorem are completely arbitrary; they may be linearly dependent or they may even be equal to each other.

## KERNEL AND IMAGE OF A LINEAR MAPPING

We begin by defining two concepts.

**Definition:** Let $F: V \to U$ be a linear mapping. The *image* of $F$, written $\operatorname{Im} F$, is the set of image points in $U$:

$$\operatorname{Im} F = \{u \in U : F(v) = u \text{ for some } v \in V\}$$

The *kernel* of $F$, written $\operatorname{Ker} F$, is the set of elements in $V$ which map into $0 \in U$:

$$\operatorname{Ker} F = \{v \in V : F(v) = 0\}$$

The following theorem is easily proven (Problem 6.22).

**Theorem 6.3:** Let $F: V \to U$ be a linear mapping. Then the image of $F$ is a subspace of $U$ and the kernel of $F$ is a subspace of $V$.

v.v.9

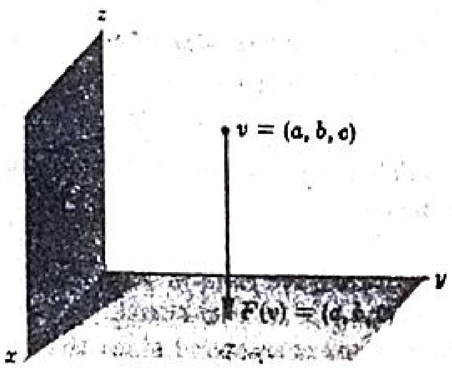
**Example 6.18:** Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the projection mapping into the $xy$ plane: $F(x, y, z) = (x, y, 0)$. Clearly the image of $F$ is the entire $xy$ plane:

$$\operatorname{Im} F = \{(a, b, 0) : a, b \in \mathbf{R}\}$$

Note that the kernel of $F$ is the $z$ axis:

$$\operatorname{Ker} F = \{(0, 0, c) : c \in \mathbf{R}\}$$

since these points and only these points map into the zero vector $0 = (0, 0, 0)$.



Now suppose that the vectors $v_1, \ldots, v_n$ generate $V$ and that $F: V \to U$ is linear. We show that the vectors $F(v_1), \ldots, F(v_n) \in U$ generate $\operatorname{Im} F$. For suppose $u \in \operatorname{Im} F$; then $F(v) = u$ for some vector $v \in V$. Since the $v_i$ generate $V$ and since $v \in V$, there exist scalars $a_1, \ldots, a_n$ for which $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Accordingly,

$$u = F(v) = F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n)$$

and hence the vectors $F(v_1), \ldots, F(v_n)$ generate $\operatorname{Im} F$.

**Example 6.19:** Consider an arbitrary $4 \times 3$ matrix $A$ over a field $K$:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}$$

which we view as a linear mapping $A: K^3 \to K^4$. Now the usual basis $\{e_1, e_2, e_3\}$ of $K^3$ generates $K^3$ and so their values $Ae_1, Ae_2, Ae_3$ under $A$ generate the image of $A$. But the vectors $Ae_1, Ae_2$, and $Ae_3$ are the columns of $A$:

$$Ae_1 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix}, \qquad Ae_2 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{pmatrix}$$

$$Ae_3 = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \\ d_3 \end{pmatrix}$$

thus the image of $A$ is precisely the column space of $A$.

We emphasize that if $A$ is any $m \times n$ matrix over $K$ viewed as a linear mapping $A: K^n \to K^m$, then the image of $A$ is precisely the column space of $A$.

So far we have not related the notion of dimension to that of a linear mapping $F: V \to U$. In the case that $V$ is of finite dimension, we have the following fundamental relationship.

**Theorem 6.4:** Let $V$ be of finite dimension and let $F: V \to U$ be a linear mapping. Then

$$\dim V = \dim(\operatorname{Ker} F) + \dim(\operatorname{Im} F)$$

That is, the sum of the dimensions of the image and kernel of a linear mapping is equal to the dimension of its domain. This formula is easily seen to hold for the projection mapping $F$ in Example 6.18. There the image ($xy$ plane) and the kernel ($z$ axis) of $F$ have dimensions 2 and 1 respectively, whereas the domain $\mathbf{R}^3$ of $F$ has dimension 3.

**Remark:** Let $F: V \to U$ be a linear mapping. Then the *rank* of $F$ is defined to be the dimension of its image, and the *nullity* of $F$ is defined to be the dimension of its kernel:

$$\operatorname{rank}(F) = \dim(\operatorname{Im} F) \quad \text{and} \quad \operatorname{nullity}(F) = \dim(\operatorname{Ker} F)$$

Thus the preceding theorem yields the following formula for $F$ when $V$ has finite dimension:

$$\operatorname{rank}(F) + \operatorname{nullity}(F) = \dim V$$

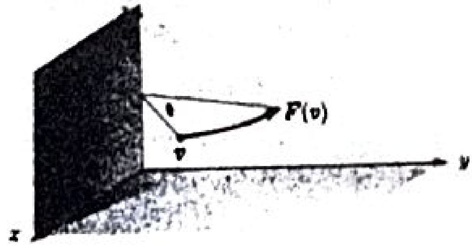
Recall that the rank of a matrix $A$ was originally defined to be the dimension of its column space and of its row space. Observe that if we now view $A$ as a linear mapping, then both definitions correspond since the image of $A$ is precisely its column space.

## SINGULAR AND NONSINGULAR MAPPINGS

A linear mapping $F: V \to U$ is said to be *singular* if the image of some nonzero vector under $F$ is 0, i.e. if there exists $v \in V$ for which $v \neq 0$ but $F(v) = 0$. Thus $F: V \to U$ is *nonsingular* if only $0 \in V$ maps into $0 \in U$ or, equivalently, if its kernel consists only of the zero vector: $\operatorname{Ker} F = \{0\}$.

**Example 6.20:** Let $F: \mathbf{R}^3 \to \mathbf{R}^3$ be the linear mapping which rotates a vector about the $z$ axis through an angle $\theta$:

$$F(x, y, z) = (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta,\ z)$$




Observe that only the zero vector is mapped into the zero vector; hence $F$ is nonsingular.

Now if the linear mapping $F: V \to U$ is one-to-one, then only $0 \in V$ can map into $0 \in U$ and so $F$ is nonsingular. The converse is also true. For suppose $F$ is nonsingular and $F(v) = F(w)$; then $F(v - w) = F(v) - F(w) = 0$ and hence $v - w = 0$ or $v = w$. Thus $F(v) = F(w)$ implies $v = w$, that is, $F$ is one-to-one. By definition (page 125), a one-to-one linear mapping is called an isomorphism. Thus we have proven

**Theorem 6.5:** A linear mapping $F: V \to U$ is an isomorphism if and only if it is nonsingular.

We remark that nonsingular mappings can also be characterized as those mappings which carry independent sets into independent sets (Problem 6.26).

## LINEAR MAPPINGS AND SYSTEMS OF LINEAR EQUATIONS

Consider system of $m$ linear equations in $n$ unknowns over a field $K$:

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

which is equivalent to the matrix equation

$$Ax = b$$

where $A = (a_{ij})$ is the coefficient matrix, and $x = (x_i)$ and $b = (b_i)$ are the column vectors of the unknowns and of the constants, respectively. Now the matrix $A$ may also be viewed as the linear mapping

$$A: K^n \to K^m$$

Thus the solution of the equation $Ax = b$ may be viewed as the preimage of $b \in K^m$ under the linear mapping $A: K^n \to K^m$. Furthermore, the solution of the associated homogeneous equation $Ax = 0$ may be viewed as the kernel of the linear mapping $A: K^n \to K^m$.

By Theorem 6.4,

$$\dim(\operatorname{Ker} A) = \dim K^n - \dim(\operatorname{Im} A) = n - \operatorname{rank} A$$

But $n$ is exactly the number of unknowns in the homogeneous system $Ax = 0$. Thus we have the following theorem on linear equations appearing in Chapter 5.

a

**Theorem 5.11:** The dimension of the solution space $W$ of the homogeneous system of linear equations $AX = 0$ is $n - r$ where $n$ is the number of unknowns and $r$ is the rank of the coefficient matrix $A$.

## OPERATIONS WITH LINEAR MAPPINGS

We are able to combine linear mappings in various ways to obtain new linear mappings. These operations are very important and shall be used throughout the text.

Suppose $F: V \to U$ and $G: V \to U$ are linear mappings of vector spaces over a field $K$. We define the sum $F + G$ to be the mapping from $V$ into $U$ which assigns $F(v) + G(v)$ to $v \in V$:

$$(F + G)(v) = F(v) + G(v)$$

Furthermore, for any scalar $k \in K$, we define the product $kF$ to be the mapping from $V$ into $U$ which assigns $kF(v)$ to $v \in V$:

$$(kF)(v) = kF(v)$$

We show that if $F$ and $G$ are linear, then $F + G$ and $kF$ are also linear. We have, for any vectors $v, w \in V$ and any scalars $a, b \in K$,

$$\begin{aligned}(F + G)(av + bw) &= F(av + bw) + G(av + bw) \\ &= aF(v) + bF(w) + aG(v) + bG(w) \\ &= a(F(v) + G(v)) + b(F(w) + G(w)) \\ &= a(F + G)(v) + b(F + G)(w)\end{aligned}$$

and
$$\begin{aligned}(kF)(av + bw) &= kF(av + bw) = k(aF(v) + bF(w)) \\ &= akF(v) + bkF(w) = a(kF)(v) + b(kF)(w)\end{aligned}$$

Thus $F + G$ and $kF$ are linear.

The following theorem applies.

**Theorem 6.6:** Let $V$ and $U$ be vector spaces over a field $K$. Then the collection of all linear mappings from $V$ into $U$ with the above operations of addition and scalar multiplication form a vector space over $K$.

The space in the above theorem is usually denoted by

$$\mathrm{Hom}(V, U)$$

Here Hom comes from the word homomorphism. In the case that $V$ and $U$ are of finite dimension, we have the following theorem.

**Theorem 6.7:** Suppose $\dim V = m$ and $\dim U = n$. Then $\dim \mathrm{Hom}(V, U) = mn$.

Now suppose that $V$, $U$ and $W$ are vector spaces over the same field $K$, and that $F: V \to U$ and $G: U \to W$ are linear mappings:

$$V \xrightarrow{F} U \xrightarrow{G} W$$

Recall that the composition function $G \circ F$ is the mapping from $V$ into $W$ defined by $(G \circ F)(v) = G(F(v))$. We show that $G \circ F$ is linear whenever $F$ and $G$ are linear. We have, for any vectors $v, w \in V$ and any scalars $a, b \in K$,

$$\begin{aligned}(G \circ F)(av + bw) &= G(F(av + bw)) = G(aF(v) + bF(w)) \\ &= aG(F(v)) + bG(F(w)) = a(G \circ F)(v) + b(G \circ F)(w)\end{aligned}$$

That is, $G \circ F$ is linear.

The composition of linear mappings and that of addition and scalar multiplication are related as follows:

**Theorem 6.8:** Let $V$, $U$ and $W$ be vector spaces over $K$. Let $F, F'$ be linear mappings from $V$ into $U$ and $G, G'$ linear mappings from $U$ into $W$, and let $k \in K$. Then:

(i) $G \circ (F + F') = G \circ F + G \circ F'$

(ii) $(G + G') \circ F = G \circ F + G' \circ F$

(iii) $k(G \circ F) = (kG) \circ F = G \circ (kF)$.

## ALGEBRA OF LINEAR OPERATORS

Let $V$ be a vector space over a field $K$. We now consider the special case of linear mappings $T : V \to V$, i.e. from $V$ into itself. They are also called *linear operators* or *linear transformations* on $V$. We will write $A(V)$, instead of $\text{Hom}(V, V)$, for the space of all such mappings.

By Theorem 6.6, $A(V)$ is a vector space over $K$; it is of dimension $n^2$ if $V$ is of dimension $n$. Now if $T, S \in A(V)$, then the composition $S \circ T$ exists and is also a linear mapping from $V$ into itself, i.e. $S \circ T \in A(V)$. Thus we have a "multiplication" defined in $A(V)$. (We shall write $ST$ for $S \circ T$ in the space $A(V)$.)

We remark that an *algebra* $A$ over a field $K$ is a vector space over $K$ in which an operation of multiplication is defined satisfying, for every $F, G, H \in A$ and every $k \in K$,

(i) $F(G + H) = FG + FH$

(ii) $(G + H)F = GF + HF$

(iii) $k(GF) = (kG)F = G(kF)$.

If the associative law also holds for the multiplication, i.e. if for every $F, G, H \in A$,

(iv) $(FG)H = F(GH)$

then the algebra $A$ is said to be *associative*. Thus by Theorems 6.8 and 6.1, $A(V)$ is an associative algebra over $K$ with respect to composition of mappings; hence it is frequently called the *algebra of linear operators* on $V$.

Observe that the identity mapping $I : V \to V$ belongs to $A(V)$. Also, for any $T \in A(V)$, we have $TI = IT = T$. We note that we can also form "powers" of $T$; we use the notation $T^2 = T \circ T$, $T^3 = T \circ T \circ T$, .... Furthermore, for any polynomial

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \qquad a_i \in K$$

we can form the operator $p(T)$ defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n$$

(For a scalar $k \in K$, the operator $kI$ is frequently denoted by simply $k$.) In particular, if $p(T) = 0$, the zero mapping, then $T$ is said to be a *zero* of the polynomial $p(x)$.

**Example 6.21:** Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be defined by $T(x, y, z) = (0, x, y)$. Now if $(a, b, c)$ is any element of $\mathbf{R}^3$, then:

$$(T + I)(a, b, c) = (0, a, b) + (a, b, c) = (a, a + b, b + c)$$

and

$$T^3(a, b, c) = T^2(0, a, b) = T(0, 0, a) = (0, 0, 0)$$

Thus we see that $T^3 = 0$, the zero mapping from $V$ into itself. In other words, $T$ is a zero of the polynomial $p(x) = x^3$.

## INVERTIBLE OPERATORS

A linear operator $T: V \to V$ is said to be *invertible* if it has an inverse, i.e. if there exists $T^{-1} \in A(V)$ such that $TT^{-1} = T^{-1}T = I$.

Now $T$ is invertible if and only if it is one-one and onto. Thus in particular, if $T$ is invertible then only $0 \in V$ can map into itself, i.e. $T$ is nonsingular. On the other hand, suppose $T$ is nonsingular, i.e. $\operatorname{Ker} T = \{0\}$. Recall (page 127) that $T$ is also one-one. Moreover, assuming $V$ has finite dimension, we have, by Theorem 6.4,

$$\begin{aligned}\dim V &= \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T) = \dim(\operatorname{Im} T) + \dim(\{0\}) \\ &= \dim(\operatorname{Im} T) + 0 = \dim(\operatorname{Im} T)\end{aligned}$$

Then $\operatorname{Im} T = V$, i.e. the image of $T$ is $V$; thus $T$ is onto. Hence $T$ is both one-one and onto and so is invertible. We have just proven

**Theorem 6.9:** A linear operator $T: V \to V$ on a vector space of finite dimension is invertible if and only if it is nonsingular.

**Example 6.22:** Let $T$ be the operator on $\mathbf{R}^2$ defined by $T(x, y) = (y, 2x - y)$. The kernel of $T$ is $\{(0, 0)\}$, hence $T$ is nonsingular and, by the preceding theorem, invertible. We now find a formula for $T^{-1}$. Suppose $(s, t)$ is the image of $(x, y)$ under $T$, hence $(x, y)$ is the image of $(s, t)$ under $T^{-1}$: $T(x, y) = (s, t)$ and $T^{-1}(s, t) = (x, y)$. We have

$$T(x, y) = (y, 2x - y) = (s, t) \quad \text{and so} \quad y = s,\ 2x - y = t$$

Solving for $x$ and $y$ in terms of $s$ and $t$, we obtain $x = \frac{1}{2}s + \frac{1}{2}t$, $y = s$. Thus $T^{-1}$ is given by the formula $T^{-1}(s, t) = (\frac{1}{2}s + \frac{1}{2}t, s)$.

The finiteness of the dimensionality of $V$ in the preceding theorem is necessary as seen in the next example.

**Example 6.23:** Let $V$ be the vector space of polynomials over $K$, and let $T$ be the operator on $V$ defined by

$$T(a_0 + a_1 t + \cdots + a_n t^n) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$$

i.e. $T$ increases the exponent of $t$ in each term by 1. Now $T$ is a linear mapping and is nonsingular. However, $T$ is not onto and so is not invertible.

We now give an important application of the above theorem to systems of linear equations over $K$. Consider a system with the same number of equations as unknowns, say $n$. We can represent this system by the matrix equation

$$Ax = b \tag{$*$}$$

where $A$ is an $n$-square matrix over $K$ which we view as a linear operator on $K^n$. Suppose the matrix $A$ is *nonsingular*, i.e. the matrix equation $Ax = 0$ has only the zero solution. Then, by Theorem 6.9, the linear mapping $A$ is one-to-one and onto. This means that the system $(*)$ has a unique solution for any $b \in K^n$. On the other hand, suppose the matrix $A$ is *singular*, i.e. the matrix equation $Ax = 0$ has a nonzero solution. Then the linear mapping $A$ is not onto. This means that there exist $b \in K^n$ for which $(*)$ does not have a solution. Furthermore, if a solution exists it is not unique. Thus we have proven the following fundamental result:

**Theorem 6.10:** Consider the following system of linear equations:

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

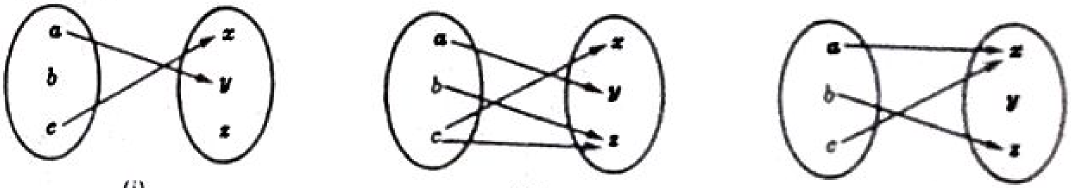
(i) If the corresponding homogeneous system has only the zero solution, then the above system has a unique solution for any values of the $b_i$.

(ii) If the corresponding homogeneous system has a nonzero solution, then: (i) there are values for the $b_i$ for which the above system does not have a solution; (ii) whenever a solution of the above system exists, it is not unique.

## Solved Problems

### MAPPINGS

**6.1.** State whether or not each diagram defines a mapping from $A = \{a, b, c\}$ into $B = \{x, y, z\}$.




(i) (ii) (iii)

(i) No. There is nothing assigned to the element $b \in A$.

(ii) No. Two elements, $x$ and $z$, are assigned to $c \in A$.

(iii) Yes.

**6.2.** Use a formula to define each of the following functions from $\mathbf{R}$ into $\mathbf{R}$.

(i) To each number let $f$ assign its cube.

(ii) To each number let $g$ assign the number 5.

(iii) To each positive number let $h$ assign its square, and to each nonpositive number let $h$ assign the number 6.

Also, find the value of each function at 4, $-2$ and 0.

(i) Since $f$ assigns to any number $x$ its cube $x^3$, we can define $f$ by $f(x) = x^3$. Also:

$$f(4) = 4^3 = 64, \quad f(-2) = (-2)^3 = -8, \quad f(0) = 0^3 = 0$$

(ii) Since $g$ assigns 5 to any number $x$, we can define $g$ by $g(x) = 5$. Thus the value of $g$ at each number 4, $-2$ and 0 is 5:

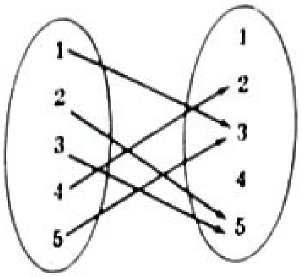
$$g(4) = 5, \quad g(-2) = 5, \quad g(0) = 5$$

(iii) Two different rules are used to define $h$ as follows:

$$h(x) = \begin{cases} x^2 & \text{if } x > 0 \\ 6 & \text{if } x \leq 0 \end{cases}$$

Since $4 > 0$, $h(4) = 4^2 = 16$. On the other hand, $-2, 0 \leq 0$ and so $h(-2) = 6$, $h(0) = 6$.

**6.3.** Let $A = \{1,2,3,4,5\}$ and let $f: A \to A$ be the mapping defined by the diagram on the right. (i) Find the image of $f$. (ii) Find the graph of $f$.




(i) The image $f(A)$ of the mapping $f$ consists of all the points assigned to elements of $A$. Now only 2, 3 and 5 appear as the image of any elements of $A$, hence $f(A) = \{2,3,5\}$.

(ii) The graph of $f$ consists of the ordered pairs $(a, f(a))$, where $a \in A$. Now $f(1) = 3$, $f(2) = 5$, $f(3) = 5$, $f(4) = 2$, $f(5) = 3$; hence the graph of

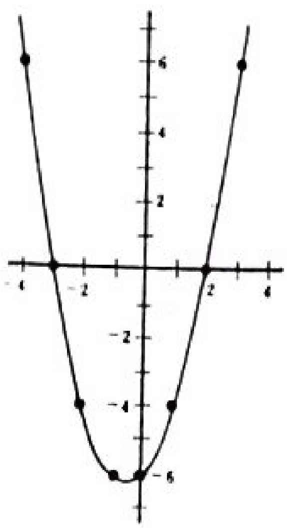
$$f = \{(1,3), (2,5), (3,5), (4,2), (5,3)\}$$

**6.4.** Sketch the graph of (i) $f(x) = x^2 + x - 6$, (ii) $g(x) = x^3 - 3x^2 - x + 3$.

Note that these are "polynomial functions". In each case set up a table of values for $x$ and then find the corresponding values of $f(x)$. Plot the points in a coordinate diagram and then draw a smooth continuous curve through the points.

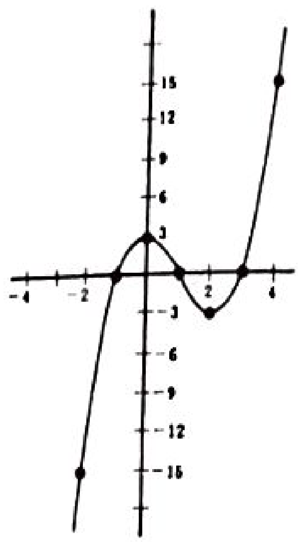
(i)

| $x$ | $f(x)$ |
|---|---|
| $-4$ | 6 |
| $-3$ | 0 |
| $-2$ | $-4$ |
| $-1$ | $-6$ |
| 0 | $-6$ |
| 1 | $-4$ |
| 2 | 0 |
| 3 | 6 |

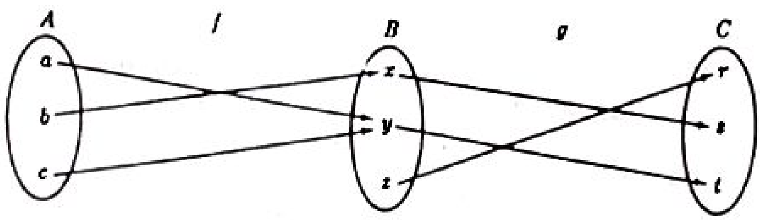



(ii)

| $x$ | $g(x)$ |
|---|---|
| $-2$ | $-15$ |
| $-1$ | 0 |
| 0 | 3 |
| 1 | 0 |
| 2 | $-3$ |
| 3 | 0 |
| 4 | 15 |




**6.5.** Let the mappings $f: A \to B$ and $g: B \to C$ be defined by the diagram




(i) Find the composition mapping $(g \circ f): A \to C$. (ii) Find the image of each mapping: $f$, $g$ and $g \circ f$.

(i) We use the definition of the composition mapping to compute:

$$(g \circ f)(a) = g(f(a)) = g(y) = t$$
$$(g \circ f)(b) = g(f(b)) = g(x) = s$$
$$(g \circ f)(c) = g(f(c)) = g(y) = t$$

Observe that we arrive at the same answer if we "follow the arrows" in the diagram:

$$a \to y \to t, \quad b \to x \to s, \quad c \to y \to t$$

(ii) By the diagram, the image values under the mapping $f$ are $x$ and $y$, and the image values under $g$ are $r$, $s$ and $t$; hence

$$\text{image of } f = \{x, y\} \quad \text{and} \quad \text{image of } g = \{r, s, t\}$$

By (i), the image values under the composition mapping $g \circ f$ are $t$ and $s$; hence image of $g \circ f = \{s, t\}$. Note that the images of $g$ and $g \circ f$ are different.

**6.6.** Let the mappings $f$ and $g$ be defined by $f(x) = 2x + 1$ and $g(x) = x^2 - 2$. (i) Find $(g \circ f)(4)$ and $(f \circ g)(4)$. (ii) Find formulae defining the composition mappings $g \circ f$ and $f \circ g$.

(i) $f(4) = 2 \cdot 4 + 1 = 9$. Hence $(g \circ f)(4) = g(f(4)) = g(9) = 9^2 - 2 = 79$.
$g(4) = 4^2 - 2 = 14$. Hence $(f \circ g)(4) = f(g(4)) = f(14) = 2 \cdot 14 + 1 = 29$.

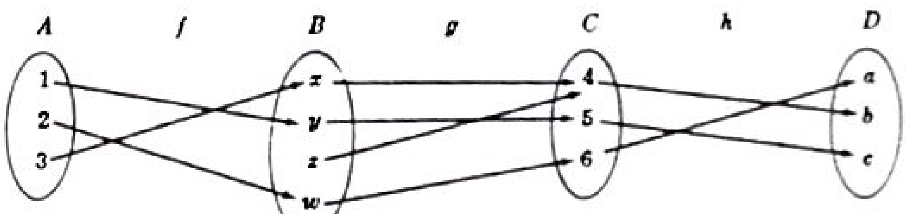
(ii) Compute the formula for $g \circ f$ as follows:

$$(g \circ f)(x) = g(f(x)) = g(2x + 1) = (2x + 1)^2 - 2 = 4x^2 + 4x - 1$$

Observe that the same answer can be found by writing $y = f(x) = 2x + 1$ and $z = g(y) = y^2 - 2$, and then eliminating $y$: $z = y^2 - 2 = (2x + 1)^2 - 2 = 4x^2 + 4x - 1$.

$(f \circ g)(x) = f(g(x)) = f(x^2 - 2) = 2(x^2 - 2) + 1 = 2x^2 - 3$. Observe that $f \circ g \neq g \circ f$.

**6.7.** Let the mappings $f : A \to B$, $g : B \to C$ and $h : C \to D$ be defined by the diagram




Determine if each mapping (i) is one-one, (ii) is onto, (iii) has an inverse.

(i) The mapping $f : A \to B$ is one-one since each element of $A$ has a different image. The mapping $g : B \to C$ is not one-one since $x$ and $z$ both map into the same element 4. The mapping $h : C \to D$ is one-one.

(ii) The mapping $f : A \to B$ is not onto since $z \in B$ is not the image of any element of $A$. The mapping $g : B \to C$ is onto since each element of $C$ is the image of some element of $B$. The mapping $h : C \to D$ is also onto.

(iii) A mapping has an inverse if and only if it is both one-one and onto. Hence only $h$ has an inverse.

**6.8.** Suppose $f : A \to B$ and $g : B \to C$; hence the composition mapping $(g \circ f) : A \to C$ exists. Prove the following. (i) If $f$ and $g$ are one-one, then $g \circ f$ is one-one. (ii) If $f$ and $g$ are onto, then $g \circ f$ is onto. (iii) If $g \circ f$ is one-one, then $f$ is one-one. (iv) If $g \circ f$ is onto, then $g$ is onto.

(i) Suppose $(g \circ f)(x) = (g \circ f)(y)$. Then $g(f(x)) = g(f(y))$. Since $g$ is one-one, $f(x) = f(y)$. Since $f$ is one-one, $x = y$. We have proven that $(g \circ f)(x) = (g \circ f)(y)$ implies $x = y$; hence $g \circ f$ is one-one.

(ii) Suppose $c \in C$. Since $g$ is onto, there exists $b \in B$ for which $g(b) = c$. Since $f$ is onto, there exists $a \in A$ for which $f(a) = b$. Thus $(g \circ f)(a) = g(f(a)) = g(b) = c$; hence $g \circ f$ is onto.

(iii) Suppose $f$ is not one-one. Then there exists distinct elements $x, y \in A$ for which $f(x) = f(y)$. Thus $(g \circ f)(x) = g(f(x)) = g(f(y)) = (g \circ f)(y)$; hence $g \circ f$ is not one-one. Accordingly if $g \circ f$ is one-one, then $f$ must be one-one.

(iv) If $a \in A$, then $(g \circ f)(a) = g(f(a)) \in g(B)$; hence $(g \circ f)(A) \subset g(B)$. Suppose $g$ is not onto. Then $g(B)$ is properly contained in $C$ and so $(g \circ f)(A)$ is properly contained in $C$; thus $g \circ f$ is not onto. Accordingly if $g \circ f$ is onto, then $g$ must be onto.

**6.9.** Prove that a mapping $f: A \to B$ has an inverse if and only if it is one-to-one and onto.

Suppose $f$ has an inverse, i.e. there exists a function $f^{-1}: B \to A$ for which $f^{-1} \circ f = 1_A$ and $f \circ f^{-1} = 1_B$. Since $1_A$ is one-to-one, $f$ is one-to-one by Problem 6.8(iii); and since $1_B$ is onto, $f$ is onto by Problem 6.8(iv). That is, $f$ is both one-to-one and onto.

Now suppose $f$ is both one-to-one and onto. Then each $b \in B$ is the image of a unique element in $A$, say $\hat{b}$. Thus if $f(a) = b$, then $a = \hat{b}$; hence $f(\hat{b}) = b$. Now let $g$ denote the mapping from $B$ to $A$ defined by $b \mapsto \hat{b}$. We have:

(i) $(g \circ f)(a) = g(f(a)) = g(b) = \hat{b} = a$, for every $a \in A$; hence $g \circ f = 1_A$.

(ii) $(f \circ g)(b) = f(g(b)) = f(\hat{b}) = b$, for every $b \in B$; hence $f \circ g = 1_B$.

Accordingly, $f$ has an inverse. Its inverse is the mapping $g$.

**6.10.** Let $f: \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = 2x - 3$. Now $f$ is one-to-one and onto; hence has an inverse mapping $f^{-1}$. Find a formula for $f^{-1}$.

Let $y$ be the image of $x$ under the mapping $f$: $y = f(x) = 2x - 3$. Consequently $x$ will be the image of $y$ under the inverse mapping $f^{-1}$. Thus solve for $x$ in terms of $y$ in the above equation: $x = (y+3)/2$. Then the formula defining the inverse function is $f^{-1}(y) = (y+3)/2$.

## LINEAR MAPPINGS

**6.11.** Show that the following mappings $F$ are linear:

(i) $F: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (x + y, x)$.

(ii) $F: \mathbf{R}^3 \to \mathbf{R}$ defined by $F(x, y, z) = 2x - 3y + 4z$.

(i) Let $v = (a, b)$ and $w = (a', b')$; hence

$$v + w = (a + a', b + b') \quad \text{and} \quad kv = (ka, kb), \quad k \in \mathbf{R}$$

We have $F(v) = (a + b, a)$ and $F(w) = (a' + b', a')$. Thus

$$\begin{aligned} F(v + w) &= F(a + a', b + b') = (a + a' + b + b', a + a') \\ &= (a + b, a) + (a' + b', a') = F(v) + F(w) \end{aligned}$$

and
$$F(kv) = F(ka, kb) = (ka + kb, ka) = k(a + b, a) = kF(v)$$

Since $v$, $w$ and $k$ were arbitrary, $F$ is linear.

(ii) Let $v = (a, b, c)$ and $w = (a', b', c')$; hence

$$v + w = (a + a', b + b', c + c') \quad \text{and} \quad kv = (ka, kb, kc), \quad k \in \mathbf{R}$$

We have $F(v) = 2a - 3b + 4c$ and $F(w) = 2a' - 3b' + 4c'$. Thus

$$\begin{aligned} F(v + w) &= F(a + a', b + b', c + c') = 2(a + a') - 3(b + b') + 4(c + c') \\ &= (2a - 3b + 4c) + (2a' - 3b' + 4c') = F(v) + F(w) \end{aligned}$$

and
$$F(kv) = F(ka, kb, kc) = 2ka - 3kb + 4kc = k(2a - 3b + 4c) = kF(v)$$

Accordingly, $F$ is linear.

**6.12.** Show that the following mappings $F$ are not linear:

(i) $F: \mathbf{R}^2 \to \mathbf{R}$ defined by $F(x, y) = xy$.

(ii) $F: \mathbf{R}^2 \to \mathbf{R}^3$ defined by $F(x, y) = (x + 1, 2y, x + y)$.

(iii) $F: \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (|x|, 0)$.

(i) Let $v = (1, 2)$ and $w = (3, 4)$; then $v + w = (4, 6)$.

We have $F(v) = 1 \cdot 2 = 2$ and $F(w) = 3 \cdot 4 = 12$. Hence

$$F(v+w) = F(4,6) = 4 \cdot 6 = 24 \neq F(v) + F(w)$$

Accordingly, $F$ is not linear.

(ii) Since $F(0,0) = (1,0,0) \neq (0,0,0)$, $F$ cannot be linear.

(iii) Let $v = (1,2,3)$ and $k = -3$; hence $kv = (-3,-6,-9)$.

We have $F(v) = (1,0)$ and so $kF(v) = -3(1,0) = (-3,0)$. Then

$$F(kv) = F(-3,-6,-9) = (3,0) \neq kF(v)$$

and hence $F$ is not linear.

v.v9 **6.13.** Let $V$ be the vector space of $n$-square matrices over $K$. Let $M$ be an arbitrary matrix in $V$. Let $T: V \to V$ be defined by $T(A) = AM + MA$, where $A \in V$. Show that $T$ is linear.

For any $A, B \in V$ and any $k \in K$, we have

$$T(A+B) = (A+B)M + M(A+B) = AM + BM + MA + MB$$
$$= (AM + MA) + (BM + MB) = T(A) + T(B)$$

and

$$T(kA) = (kA)M + M(kA) = k(AM) + k(MA) = k(AM + MA) = kT(A)$$

Accordingly, $T$ is linear.

A M v.v.9 **6.14.** Prove Theorem 6.2: Let $V$ and $U$ be vector spaces over a field $K$. Let $\{v_1, \ldots, v_n\}$ be a basis of $V$ and let $u_1, \ldots, u_n$ be any arbitrary vectors in $U$. Then there exists a unique linear mapping $F: V \to U$ such that $F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

There are three steps to the proof of the theorem: (1) Define a mapping $F: V \to U$ such that $F(v_i) = u_i$, $i = 1, \ldots, n$. (2) Show that $F$ is linear. (3) Show that $F$ is unique.

*Step (1).* Let $v \in V$. Since $\{v_1, \ldots, v_n\}$ is a basis of $V$, there exist unique scalars $a_1, \ldots, a_n \in K$ for which $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. We define $F: V \to U$ by $F(v) = a_1u_1 + a_2u_2 + \cdots + a_nu_n$. (Since the $a_i$ are unique, the mapping $F$ is well-defined.) Now, for $i = 1, \ldots, n$,

$$v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_n$$

Hence

$$F(v_i) = 0u_1 + \cdots + 1u_i + \cdots + 0u_n = u_i$$

Thus the first step of the proof is complete.

*Step (2).* Suppose $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ and $w = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Then

$$v + w = (a_1+b_1)v_1 + (a_2+b_2)v_2 + \cdots + (a_n+b_n)v_n$$

and, for any $k \in K$, $kv = ka_1v_1 + ka_2v_2 + \cdots + ka_nv_n$. By definition of the mapping $F$,

$$F(v) = a_1u_1 + a_2u_2 + \cdots + a_nu_n \quad \text{and} \quad F(w) = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

Hence

$$F(v+w) = (a_1+b_1)u_1 + (a_2+b_2)u_2 + \cdots + (a_n+b_n)u_n$$
$$= (a_1u_1 + a_2u_2 + \cdots + a_nu_n) + (b_1u_1 + b_2u_2 + \cdots + b_nu_n)$$
$$= F(v) + F(w)$$

and

$$F(kv) = k(a_1u_1 + a_2u_2 + \cdots + a_nu_n) = kF(v)$$

Thus $F$ is linear.

*Step (3).* Now suppose $G: V \to U$ is linear and $G(v_i) = u_i$, $i = 1, \ldots, n$. If $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, then

$$G(v) = G(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1G(v_1) + a_2G(v_2) + \cdots + a_nG(v_n)$$
$$= a_1u_1 + a_2u_2 + \cdots + a_nu_n = F(v)$$

Since $G(v) = F(v)$ for every $v \in V$, $G = F$. Thus $F$ is unique and the theorem is proved.

**6.15.** Let $T: \mathbf{R}^2 \to \mathbf{R}$ be the linear mapping for which

$$T(1, 1) = 3 \quad \text{and} \quad T(0, 1) = -2 \tag{1}$$

(Since $\{(1,1), (0,1)\}$ is a basis of $\mathbf{R}^2$, such a linear mapping exists and is unique by Theorem 6.2.) Find $T(a, b)$.

First we write $(a, b)$ as a linear combination of $(1, 1)$ and $(0, 1)$ using unknown scalars $x$ and $y$:

$$(a, b) = x(1, 1) + y(0, 1) \tag{2}$$

Then $$(a, b) = (x, x) + (0, y) = (x, x+y) \quad \text{and so} \quad x = a,\ x + y = b$$

Solving for $x$ and $y$ in terms of $a$ and $b$, we obtain

$$x = a \quad \text{and} \quad y = b - a \tag{3}$$

Now using (1) and (2) we have

$$T(a, b) = T(x(1, 1) + y(0, 1)) = xT(1, 1) + yT(0, 1) = 3x - 2y$$

Finally, using (3) we have $T(a, b) = 3x - 2y = 3(a) - 2(b - a) = 5a - 2b$.

**6.16.** Let $T: V \to U$ be linear, and suppose $v_1, \ldots, v_n \in V$ have the property that their images $T(v_1), \ldots, T(v_n)$ are linearly independent. Show that the vectors $v_1, \ldots, v_n$ are also linearly independent.

Suppose that, for scalars $a_1, \ldots, a_n$, $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Then

$$0 = T(0) = T(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n)$$

Since the $T(v_i)$ are linearly independent, all the $a_i = 0$. Thus the vectors $v_1, \ldots, v_n$ are linearly independent.

**6.17.** Suppose the linear mapping $F: V \to U$ is one-to-one and onto. Show that the inverse mapping $F^{-1}: U \to V$ is also linear.

Suppose $u, u' \in U$. Since $F$ is one-to-one and onto, there exist unique vectors $v, v' \in V$ for which $F(v) = u$ and $F(v') = u'$. Since $F$ is linear, we also have

$$F(v + v') = F(v) + F(v') = u + u' \quad \text{and} \quad F(kv) = kF(v) = ku$$

By definition of the inverse mapping, $F^{-1}(u) = v$, $F^{-1}(u') = v'$, $F^{-1}(u + u') = v + v'$ and $F^{-1}(ku) = kv$. Then

$$F^{-1}(u + u') = v + v' = F^{-1}(u) + F^{-1}(u') \quad \text{and} \quad F^{-1}(ku) = kv = kF^{-1}(u)$$

and thus $F^{-1}$ is linear.

## IMAGE AND KERNEL OF LINEAR MAPPINGS

A **6.18.** Let $F: \mathbf{R}^4 \to \mathbf{R}^3$ be the linear mapping defined by

$$F(x, y, s, t) = (x - y + s + t,\ x + 2s - t,\ x + y + 3s - 3t)$$

Find a basis and the dimension of the (i) image $U$ of $F$, (ii) kernel $W$ of $F$.

(i) The images of the following generators of $\mathbf{R}^4$ generate the image $U$ of $F$:

$$F(1, 0, 0, 0) = (1, 1, 1) \qquad F(0, 0, 1, 0) = (1, 2, 3)$$
$$F(0, 1, 0, 0) = (-1, 0, 1) \qquad F(0, 0, 0, 1) = (1, -1, -3)$$

Form the matrix whose rows are the generators of $U$ and row reduce to echelon form:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \\ 1 & -1 & -3 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{pmatrix} \text{ to } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus $\{(1, 1, 1), (0, 1, 2)\}$ is a basis of $U$; hence $\dim U = 2$.

(ii) We seek the set of $(x, y, s, t)$ such that $F(x, y, s, t) = (0, 0, 0)$, i.e.,

$$F(x, y, s, t) = (x - y + s + t,\ x + 2s - t,\ x + y + 3s - 3t) = (0, 0, 0)$$

Set corresponding components equal to each other to form the following homogeneous system whose solution space is the kernel $W$ of $F$:

$$\begin{aligned} x - y + s + t &= 0 \\ x \quad + 2s - t &= 0 \\ x + y + 3s - 3t &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \\ 2y + 2s - 4t &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x - y + s + t &= 0 \\ y + s - 2t &= 0 \end{aligned}$$

The free variables are $s$ and $t$; hence $\dim W = 2$. Set

(a) $s = -1,\ t = 0$ to obtain the solution $(2, 1, -1, 0)$.

(b) $s = 0,\ t = 1$ to obtain the solution $(1, 2, 0, 1)$.

Thus $\{(2, 1, -1, 0), (1, 2, 0, 1)\}$ is a basis of $W$. (Observe that $\dim U + \dim W = 2 + 2 = 4$, which is the dimension of the domain $\mathbf{R}^4$ of $F$.)

A 89 86 M

**6.19.** Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be the linear mapping defined by

$$T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z)$$

Find a basis and the dimension of the (i) image $U$ of $T$, (ii) kernel $W$ of $T$.

(i) The images of generators of $\mathbf{R}^3$ generate the image $U$ of $T$:

$$T(1, 0, 0) = (1, 0, 1), \quad T(0, 1, 0) = (2, 1, 1), \quad T(0, 0, 1) = (-1, 1, -2)$$

Form the matrix whose rows are the generators of $U$ and row reduce to echelon form:

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & -2 \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus $\{(1, 0, 1), (0, 1, -1)\}$ is a basis of $U$, and so $\dim U = 2$.

(ii) We seek the set of $(x, y, z)$ such that $T(x, y, z) = (0, 0, 0)$, i.e.,

$$T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z) = (0, 0, 0)$$

Set corresponding components equal to each other to form the homogeneous system whose solution space is the kernel $W$ of $T$:

$$\begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ x + y - 2z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \\ -y - z &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y - z &= 0 \\ y + z &= 0 \end{aligned}$$

The only free variable is $z$; hence $\dim W = 1$. Let $z = 1$; then $y = -1$ and $x = 3$. Thus $\{(3, -1, 1)\}$ is a basis of $W$. (Observe that $\dim U + \dim W = 2 + 1 = 3$, which is the dimension of the domain $\mathbf{R}^3$ of $T$.)

**6.20.** Find a linear map $F : \mathbf{R}^3 \to \mathbf{R}^4$ whose image is generated by $(1, 2, 0, -4)$ and $(2, 0, -1, -3)$.

Method 1.

Consider the usual basis of $\mathbf{R}^3$: $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$. Set $F(e_1) = (1, 2, 0, -4)$, $F(e_2) = (2, 0, -1, -3)$ and $F(e_3) = (0, 0, 0, 0)$. By Theorem 6.2, such a linear map $F$ exists and is unique. Furthermore, the image of $F$ is generated by the $F(e_i)$; hence $F$ has the required property. We find a general formula for $F(x, y, z)$:

$$\begin{aligned} F(x, y, z) &= F(xe_1 + ye_2 + ze_3) = xF(e_1) + yF(e_2) + zF(e_3) \\ &= x(1, 2, 0, -4) + y(2, 0, -1, -3) + z(0, 0, 0, 0) \\ &= (x + 2y,\ 2x,\ -y,\ -4x - 3y) \end{aligned}$$

**Method 2.**

Form a $4 \times 3$ matrix $A$ whose columns consist only of the given vectors, say,

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 0 & 0 \\ 0 & -1 & -1 \\ -4 & -3 & -3 \end{pmatrix}$$

Recall that $A$ determines a linear map $A : \mathbf{R}^3 \to \mathbf{R}^4$ whose image is generated by the columns of $A$. Thus $A$ satisfies the required condition.

ME

**6.21.** Let $V$ be the vector space of 2 by 2 matrices over $\mathbf{R}$ and let $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Let $F : V \to V$ be the linear map defined by $F(A) = AM - MA$. Find a basis and the dimension of the kernel $W$ of $F$.

We seek the set of $\begin{pmatrix} x & y \\ s & t \end{pmatrix}$ such that $F\begin{pmatrix} x & y \\ s & t \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

$$\begin{aligned} F\begin{pmatrix} x & y \\ s & t \end{pmatrix} &= \begin{pmatrix} x & y \\ s & t \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x & y \\ s & t \end{pmatrix} \\ &= \begin{pmatrix} x & 2x + 3y \\ s & 2s + 3t \end{pmatrix} - \begin{pmatrix} x + 2s & y + 2t \\ 3s & 3t \end{pmatrix} \\ &= \begin{pmatrix} -2s & 2x + 2y - 2t \\ -2s & 2s \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Thus
$$\begin{aligned} 2x + 2y - 2t &= 0 \\ 2s &= 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + y - t &= 0 \\ s &= 0 \end{aligned}$$

The free variables are $y$ and $t$; hence $\dim W = 2$. To obtain a basis of $W$ set

(a) $y = -1$, $t = 0$ to obtain the solution $x = 1$, $y = -1$, $s = 0$, $t = 0$;

(b) $y = 0$, $t = 1$ to obtain the solution $x = 1$, $y = 0$, $s = 0$, $t = 1$.

Thus $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ is a basis of $W$.

**6.22.** v.v. Prove Theorem 6.3: Let $F : V \to U$ be a linear mapping. Then (i) the image of $F$ is a subspace of $U$ and (ii) the kernel of $F$ is a subspace of $V$.

(i) Since $F(0) = 0$, $0 \in \operatorname{Im} F$. Now suppose $u, u' \in \operatorname{Im} F$ and $a, b \in K$. Since $u$ and $u'$ belong to the image of $F$, there exist vectors $v, v' \in V$ such that $F(v) = u$ and $F(v') = u'$. Then

$$F(av + bv') = aF(v) + bF(v') = au + bu' \in \operatorname{Im} F$$

Thus the image of $F$ is a subspace of $U$.

(ii) Since $F(0) = 0$, $0 \in \operatorname{Ker} F$. Now suppose $v, w \in \operatorname{Ker} F$ and $a, b \in K$. Since $v$ and $w$ belong to the kernel of $F$, $F(v) = 0$ and $F(w) = 0$. Thus

$$F(av + bw) = aF(v) + bF(w) = a0 + b0 = 0 \quad \text{and so} \quad av + bw \in \operatorname{Ker} F$$

Thus the kernel of $F$ is a subspace of $V$.

93 All

**6.23.** Prove Theorem 6.4: Let $V$ be of finite dimension, and let $F : V \to U$ be a linear mapping with image $U'$ and kernel $W$. Then $\dim U' + \dim W = \dim V$.

Suppose $\dim V = n$. Since $W$ is a subspace of $V$, its dimension is finite; say, $\dim W = r \leq n$. Thus we need prove that $\dim U' = n - r$.

Let $\{w_1, \ldots, w_r\}$ be a basis of $W$. We extend $\{w_i\}$ to a basis of $V$:

$$\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}$$

Let

$$B = \{F(v_1), F(v_2), \ldots, F(v_{n-r})\}$$

The theorem is proved if we show that $B$ is a basis of the image $U'$ of $F$.

*Proof that $B$ generates $U'$.* Let $u \in U'$. Then there exists $v \in V$ such that $F(v) = u$. Since $\{w_i, v_i\}$ generates $V$ and since $v \in V$,

$$v = a_1w_1 + \cdots + a_rw_r + b_1v_1 + \cdots + b_{n-r}v_{n-r}$$

where the $a_i, b_i$ are scalars. Note that $F(w_i) = 0$ since the $w_i$ belong to the kernel of $F$. Thus

$$\begin{aligned} u = F(v) &= F(a_1w_1 + \cdots + a_rw_r + b_1v_1 + \cdots + b_{n-r}v_{n-r}) \\ &= a_1F(w_1) + \cdots + a_rF(w_r) + b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \\ &= a_10 + \cdots + a_r0 + b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \\ &= b_1F(v_1) + \cdots + b_{n-r}F(v_{n-r}) \end{aligned}$$

Accordingly, the $F(v_i)$ generate the image of $F$.

*Proof that $B$ is linearly independent.* Suppose

$$a_1F(v_1) + a_2F(v_2) + \cdots + a_{n-r}F(v_{n-r}) = 0$$

Then $F(a_1v_1 + a_2v_2 + \cdots + a_{n-r}v_{n-r}) = 0$ and so $a_1v_1 + \cdots + a_{n-r}v_{n-r}$ belongs to the kernel $W$ of $F$. Since $\{w_i\}$ generates $W$, there exist scalars $b_1, \ldots, b_r$ such that

$$a_1v_1 + a_2v_2 + \cdots + a_{n-r}v_{n-r} = b_1w_1 + b_2w_2 + \cdots + b_rw_r$$

or

$$a_1v_1 + \cdots + a_{n-r}v_{n-r} - b_1w_1 - \cdots - b_rw_r = 0 \tag{*}$$

Since $\{w_i, v_i\}$ is a basis of $V$, it is linearly independent; hence the coefficients of the $w_i$ and $v_i$ in (*) are all 0. In particular, $a_1 = 0, \ldots, a_{n-r} = 0$. Accordingly, the $F(v_i)$ are linearly independent.

Thus $B$ is a basis of $U'$, and so $\dim U' = n - r$ and the theorem is proved.

**6.24.** Suppose $f: V \to U$ is linear with kernel $W$, and that $f(v) = u$. Show that the "coset" $v + W = \{v + w : w \in W\}$ is the preimage of $u$, that is, $f^{-1}(u) = v + W$.

We must prove that (i) $f^{-1}(u) \subset v + W$ and (ii) $v + W \subset f^{-1}(u)$. We first prove (i). Suppose $v' \in f^{-1}(u)$. Then $f(v') = u$ and so $f(v' - v) = f(v') - f(v) = u - u = 0$, that is, $v' - v \in W$. Thus $v' = v + (v' - v) \in v + W$ and hence $f^{-1}(u) \subset v + W$.

Now we prove (ii). Suppose $v' \in v + W$. Then $v' = v + w$ where $w \in W$. Since $W$ is the kernel of $f$, $f(w) = 0$. Accordingly, $f(v') = f(v + w) = f(v) + f(w) = f(v) + 0 = f(v) = u$. Thus $v' \in f^{-1}(u)$ and so $v + W \subset f^{-1}(u)$.

## SINGULAR AND NONSINGULAR MAPPINGS

**6.25.** Suppose $F: V \to U$ is linear and that $V$ is of finite dimension. Show that $V$ and the image of $F$ have the same dimension if and only if $F$ is nonsingular. Determine all nonsingular mappings $T: \mathbf{R}^4 \to \mathbf{R}^3$.

By Theorem 6.4, $\dim V = \dim(\operatorname{Im} F) + \dim(\operatorname{Ker} F)$. Hence $V$ and $\operatorname{Im} F$ have the same dimension if and only if $\dim(\operatorname{Ker} F) = 0$ or $\operatorname{Ker} F = \{0\}$, i.e. if and only if $F$ is nonsingular.

Since the dimension of $\mathbf{R}^3$ is less than the dimension of $\mathbf{R}^4$, so is the dimension of the image of $T$. Accordingly, no linear mapping $T: \mathbf{R}^4 \to \mathbf{R}^3$ can be nonsingular.

**6.26.** Prove that a linear mapping $F: V \to U$ is nonsingular if and only if the image of any independent set is independent.

Suppose $F$ is nonsingular and suppose $\{v_1, \ldots, v_n\}$ is an independent subset of $V$. We claim that the vectors $F(v_1), \ldots, F(v_n)$ are independent. Suppose $a_1F(v_1) + a_2F(v_2) + \cdots + a_nF(v_n) = 0$, where $a_i \in K$. Since $F$ is linear, $F(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = 0$; hence

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n \in \operatorname{Ker} F$$

But $F$ is nonsingular, i.e. $\operatorname{Ker} F = \{0\}$; hence $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Since the $v_i$ are linearly independent, all the $a_i$ are 0. Accordingly, the $F(v_i)$ are linearly independent. In other words, the image of the independent set $\{v_1, \ldots, v_n\}$ is independent.

On the other hand, suppose the image of any independent set is independent. If $v \in V$ is nonzero, then $\{v\}$ is independent. Then $\{F(v)\}$ is independent and so $F(v) \neq 0$. Accordingly, $F$ is nonsingular.

## OPERATIONS WITH LINEAR MAPPINGS

**6.27.** Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ and $G: \mathbf{R}^3 \to \mathbf{R}^2$ be defined by $F(x, y, z) = (2x, y + z)$ and $G(x, y, z) = (x - z, y)$. Find formulae defining the mappings $F + G$, $3F$ and $2F - 5G$.

$$(F + G)(x, y, z) = F(x, y, z) + G(x, y, z)$$
$$= (2x, y + z) + (x - z, y) = (3x - z, 2y + z)$$
$$(3F)(x, y, z) = 3F(x, y, z) = 3(2x, y + z) = (6x, 3y + 3z)$$
$$(2F - 5G)(x, y, z) = 2F(x, y, z) - 5G(x, y, z) = 2(2x, y + z) - 5(x - z, y)$$
$$= (4x, 2y + 2z) + (-5x + 5z, -5y) = (-x + 5z, -3y + 2z)$$

**6.28.** Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ and $G: \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $F(x, y, z) = (2x, y + z)$ and $G(x, y) = (y, x)$. Derive formulae defining the mappings $G \circ F$ and $F \circ G$.

$$(G \circ F)(x, y, z) = G(F(x, y, z)) = G(2x, y + z) = (y + z, 2x)$$

The mapping $F \circ G$ is not defined since the image of $G$ is not contained in the domain of $F$.

**6.29.** Show: (i) the zero mapping 0, defined by $0(v) = 0$ for every $v \in V$, is the zero element of $\operatorname{Hom}(V, U)$; (ii) the negative of $F \in \operatorname{Hom}(V, U)$ is the mapping $(-1)F$, i.e. $-F = (-1)F$.

(i) Let $F \in \operatorname{Hom}(V, U)$. Then, for every $v \in V$,

$$(F + 0)(v) = F(v) + 0(v) = F(v) + 0 = F(v)$$

Since $(F + 0)(v) = F(v)$ for every $v \in V$, $F + 0 = F$.

(ii) For every $v \in V$,

$$(F + (-1)F)(v) = F(v) + (-1)F(v) = F(v) - F(v) = 0 = 0(v)$$

Since $(F + (-1)F)(v) = 0(v)$ for every $v \in V$, $F + (-1)F = 0$. Thus $(-1)F$ is the negative of $F$.

**6.30.** Show that for $F_1, \ldots, F_n \in \operatorname{Hom}(V, U)$ and $a_1, \ldots, a_n \in K$, and for any $v \in V$,

$$(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) = a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)$$

By definition of the mapping $a_1F_1$, $(a_1F_1)(v) = a_1F_1(v)$; hence the theorem holds for $n = 1$. Thus by induction,

$$(a_1F_1 + a_2F_2 + \cdots + a_nF_n)(v) = (a_1F_1)(v) + (a_2F_2 + \cdots + a_nF_n)(v)$$
$$= a_1F_1(v) + a_2F_2(v) + \cdots + a_nF_n(v)$$

**6.31.** Let $F: \mathbf{R}^3 \to \mathbf{R}^2$, $G: \mathbf{R}^3 \to \mathbf{R}^2$ and $H: \mathbf{R}^3 \to \mathbf{R}^2$ be defined by $F(x, y, z) = (x + y + z, x + y)$, $G(x, y, z) = (2x + z, x + y)$ and $H(x, y, z) = (2y, x)$. Show that $F, G, H \in \operatorname{Hom}(\mathbf{R}^3, \mathbf{R}^2)$ are linearly independent.

Suppose, for scalars $a, b, c \in K$,

$$aF + bG + cH = 0 \qquad (1)$$

(Here 0 is the zero mapping.) For $e_1 = (1, 0, 0) \in \mathbf{R}^3$, we have

$$(aF + bG + cH)(e_1) = aF(1, 0, 0) + bG(1, 0, 0) + cH(1, 0, 0)$$
$$= a(1, 1) + b(2, 1) + c(0, 1) = (a + 2b, a + b + c)$$

and $0(e_1) = (0,0)$. Thus by (1), $(a+2b, a+b+c) = (0,0)$ and so

$$a + 2b = 0 \quad \text{and} \quad a + b + c = 0 \tag{2}$$

Similarly for $e_2 = (0,1,0) \in \mathbf{R}^3$, we have

$$\begin{aligned}(aF + bG + cH)(e_2) &= aF(0,1,0) + bG(0,1,0) + cH(0,1,0)\\ &= a(1,1) + b(0,1) + c(2,0) = (a+2c, a+b) = 0(e_2) = (0,0)\end{aligned}$$

Thus $$a + 2c = 0 \quad \text{and} \quad a + b = 0 \tag{3}$$

Using (2) and (3) we obtain $$a = 0, \; b = 0, \; c = 0 \tag{4}$$

Since (1) implies (4), the mappings $F$, $G$ and $H$ are linearly independent.

M Sir

**6.32.** Prove Theorem 6.7: Suppose $\dim V = m$ and $\dim U = n$. Then $\dim \operatorname{Hom}(V, U) = mn$.

Suppose $\{v_1, \ldots, v_m\}$ is a basis of $V$ and $\{u_1, \ldots, u_n\}$ is a basis of $U$. By Theorem 6.2, a linear mapping in $\operatorname{Hom}(V, U)$ is uniquely determined by arbitrarily assigning elements of $U$ to the basis elements $v_i$ of $V$. We define

$$F_{ij} \in \operatorname{Hom}(V, U), \quad i = 1, \ldots, m, \; j = 1, \ldots, n$$

to be the linear mapping for which $F_{ij}(v_i) = u_j$, and $F_{ij}(v_k) = 0$ for $k \neq i$. That is, $F_{ij}$ maps $v_i$ into $u_j$ and the other $v$'s into 0. Observe that $\{F_{ij}\}$ contains exactly $mn$ elements; hence the theorem is proved if we show that it is a basis of $\operatorname{Hom}(V, U)$.

*Proof that $\{F_{ij}\}$ generates* $\operatorname{Hom}(V, U)$. Let $F \in \operatorname{Hom}(V, U)$. Suppose $F(v_1) = w_1, F(v_2) = w_2, \ldots, F(v_m) = w_m$. Since $w_k \in U$, it is a linear combination of the $u$'s; say,

$$w_k = a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n, \quad k = 1, \ldots, m, \quad a_{ij} \in K \tag{1}$$

Consider the linear mapping $G = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}$. Since $G$ is a linear combination of the $F_{ij}$, the proof that $\{F_{ij}\}$ generates $\operatorname{Hom}(V, U)$ is complete if we show that $F = G$.

We now compute $G(v_k)$, $k = 1, \ldots, m$. Since $F_{ij}(v_k) = 0$ for $k \neq i$ and $F_{ki}(v_k) = u_i$,

$$\begin{aligned}G(v_k) &= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j\\ &= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n\end{aligned}$$

Thus by (1), $G(v_k) = w_k$ for each $k$. But $F(v_k) = w_k$ for each $k$. Accordingly, by Theorem 6.2, $F = G$; hence $\{F_{ij}\}$ generates $\operatorname{Hom}(V, U)$.

*Proof that $\{F_{ij}\}$ is linearly independent.* Suppose, for scalars $a_{ij} \in K$,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij} = 0$$

For $v_k$, $k = 1, \ldots, m$,

$$\begin{aligned}0 = 0(v_k) &= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}F_{ij}(v_k) = \sum_{j=1}^{n} a_{kj}F_{kj}(v_k) = \sum_{j=1}^{n} a_{kj}u_j\\ &= a_{k1}u_1 + a_{k2}u_2 + \cdots + a_{kn}u_n\end{aligned}$$

But the $u_i$ are linearly independent; hence for $k = 1, \ldots, m$, we have $a_{k1} = 0, a_{k2} = 0, \ldots, a_{kn} = 0$. In other words, all the $a_{ij} = 0$ and so $\{F_{ij}\}$ is linearly independent.

Thus $\{F_{ij}\}$ is a basis of $\operatorname{Hom}(V, U)$; hence $\dim \operatorname{Hom}(V, U) = mn$.

**6.33.** Prove Theorem 6.8: Let $V$, $U$ and $W$ be vector spaces over $K$. Let $F, F'$ be linear mappings from $V$ into $U$ and let $G, G'$ be linear mappings from $U$ into $W$; and let $k \in K$. Then: (i) $G \circ (F + F') = G \circ F + G \circ F'$; (ii) $(G + G') \circ F = G \circ F + G' \circ F$; (iii) $k(G \circ F) = (kG) \circ F = G \circ (kF)$.

(i) For every $v \in V$,

$$(G \circ (F + F'))(v) = G((F + F')(v)) = G(F(v) + F'(v))$$
$$= G(F(v)) + G(F'(v)) = (G \circ F)(v) + (G \circ F')(v) = (G \circ F + G \circ F')(v)$$

Since $(G \circ (F + F'))(v) = (G \circ F + G \circ F')(v)$, for every $v \in V$, $G \circ (F + F') = G \circ F + G \circ F'$.

(ii) For every $v \in V$,

$$((G + G') \circ F)(v) = (G + G')(F(v)) = G(F(v)) + G'(F(v))$$
$$= (G \circ F)(v) + (G' \circ F)(v) = (G \circ F + G' \circ F)(v)$$

Since $((G + G') \circ F)(v) = (G \circ F + G \circ F')(v)$ for every $v \in V$, $(G + G') \circ F = G \circ F + G' \circ F$.

(iii) For every $v \in V$,

$$(k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = (kG)(F(v)) = (kG \circ F)(v)$$

and $\quad (k(G \circ F))(v) = k(G \circ F)(v) = k(G(F(v))) = G(kF(v)) = G((kF)(v)) = (G \circ kF)(v)$

Accordingly, $k(G \circ F) = (kG) \circ F = G \circ (kF)$. (We emphasize that two mappings are shown to be equal by showing that they assign the same image to each point in the domain.)

**6.34.** Let $F: V \to U$ and $G: U \to W$ be linear. Hence $(G \circ F): V \to W$ is linear. Show that (i) rank $(G \circ F) \leq$ rank $G$, (ii) rank $(G \circ F) \leq$ rank $F$.

(i) Since $F(V) \subset U$, we also have $G(F(V)) \subset G(U)$ and so $\dim G(F(V)) \leq \dim G(U)$. Then

$$\text{rank } (G \circ F) = \dim ((G \circ F)(V)) = \dim (G(F(V))) \leq \dim G(U) = \text{rank } G$$

(ii) By Theorem 6.4, $\dim (G(F(V))) \leq \dim F(V)$. Hence

$$\text{rank } (G \circ F) = \dim ((G \circ F)(V)) = \dim (G(F(V))) \leq \dim F(V) = \text{rank } F$$

## ALGEBRA OF LINEAR OPERATORS

**6.35.** Let $S$ and $T$ be the linear operators on $\mathbf{R}^2$ defined by $S(x, y) = (y, x)$ and $T(x, y) = (0, x)$. Find formulae defining the operators $S + T$, $2S - 3T$, $ST$, $TS$, $S^2$ and $T^2$.

$(S + T)(x, y) = S(x, y) + T(x, y) = (y, x) + (0, x) = (y, 2x)$.

$(2S - 3T)(x, y) = 2S(x, y) - 3T(x, y) = 2(y, x) - 3(0, x) = (2y, -x)$.

$(ST)(x, y) = S(T(x, y)) = S(0, x) = (x, 0)$.

$(TS)(x, y) = T(S(x, y)) = T(y, x) = (0, y)$.

$S^2(x, y) = S(S(x, y)) = S(y, x) = (x, y)$. Note $S^2 = I$, the identity mapping.

$T^2(x, y) = T(T(x, y)) = T(0, x) = (0, 0)$. Note $T^2 = 0$, the zero mapping.

**6.36.** Let $T$ be the linear operator on $\mathbf{R}^2$ defined by

$$T(3, 1) = (2, -4) \quad \text{and} \quad T(1, 1) = (0, 2) \tag{1}$$

(By Theorem 6.2, such a linear operator exists and is unique.) Find $T(a, b)$. In particular, find $T(7, 4)$.

First write $(a, b)$ as a linear combination of $(3, 1)$ and $(1, 1)$ using unknown scalars $x$ and $y$:

$$(a, b) = x(3, 1) + y(1, 1) \tag{2}$$

Hence $\quad (a, b) = (3x, x) + (y, y) = (3x + y, x + y)$ and so $\begin{cases} 3x + y = a \\ x + y = b \end{cases}$

Solving for $x$ and $y$ in terms of $a$ and $b$,

$$x = \tfrac{1}{2}a - \tfrac{1}{2}b \quad \text{and} \quad y = -\tfrac{1}{2}a + \tfrac{3}{2}b \tag{3}$$

Now using (2), (1) and (3),

$$T(a, b) = xT(3, 1) + yT(1, 1) = x(2, -4) + y(0, 2)$$
$$= (2x, -4x) + (0, 2y) = (2x, -4x + 2y) = (a - b, 5b - 3a)$$

Thus $T(7, 4) = (7 - 4, 20 - 21) = (3, -1)$.

**6.37.** Let $T$ be the operator on $\mathbf{R}^3$ defined by $T(x, y, z) = (2x, 4x - y, 2x + 3y - z)$. (i) Show that $T$ is invertible. (ii) Find a formula for $T^{-1}$.

(i) The kernel $W$ of $T$ is the set of all $(x, y, z)$ such that $T(x, y, z) = (0, 0, 0)$, i.e.,

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (0, 0, 0)$$

Thus $W$ is the solution space of the homogeneous system

$$2x = 0, \quad 4x - y = 0, \quad 2x + 3y - z = 0$$

which has only the trivial solution $(0, 0, 0)$. Thus $W = \{0\}$; hence $T$ is nonsingular and so by Theorem 6.9 is invertible.

(ii) Let $(r, s, t)$ be the image of $(x, y, z)$ under $T$; then $(x, y, z)$ is the image of $(r, s, t)$ under $T^{-1}$: $T(x, y, z) = (r, s, t)$ and $T^{-1}(r, s, t) = (x, y, z)$. We will find the values of $x$, $y$ and $z$ in terms of $r$, $s$ and $t$, and then substitute in the above formula for $T^{-1}$. From

$$T(x, y, z) = (2x, 4x - y, 2x + 3y - z) = (r, s, t)$$

we find $x = \frac{1}{2}r$, $y = 2r - s$, $z = 7r - 3s - t$. Thus $T^{-1}$ is given by

$$T^{-1}(r, s, t) = (\tfrac{1}{2}r, 2r - s, 7r - 3s - t)$$

**6.38.** Let $V$ be of finite dimension and let $T$ be a linear operator on $V$. Recall that $T$ is invertible if and only if $T$ is nonsingular or one-to-one. Show that $T$ is invertible if and only if $T$ is onto.

By Theorem 6.4, $\dim V = \dim(\operatorname{Im} T) + \dim(\operatorname{Ker} T)$. Hence the following statements are equivalent: (i) $T$ is onto, (ii) $\operatorname{Im} T = V$, (iii) $\dim(\operatorname{Im} T) = \dim V$, (iv) $\dim(\operatorname{Ker} T) = 0$, (v) $\operatorname{Ker} T = \{0\}$, (vi) $T$ is nonsingular, (vii) $T$ is invertible.

**6.39.** Let $V$ be of finite dimension and let $T$ be a linear operator on $V$ for which $TS = I$, for some operator $S$ on $V$. (We call $S$ a right inverse of $T$.) (i) Show that $T$ is invertible. (ii) Show that $S = T^{-1}$. (iii) Give an example showing that the above need not hold if $V$ is of infinite dimension.

(i) Let $\dim V = n$. By the preceding problem, $T$ is invertible if and only if $T$ is onto; hence $T$ is invertible if and only if $\operatorname{rank} T = n$. We have $n = \operatorname{rank} I = \operatorname{rank} TS \leq \operatorname{rank} T \leq n$. Hence $\operatorname{rank} T = n$ and $T$ is invertible.

(ii) $TT^{-1} = T^{-1}T = I$. Then $S = IS = (T^{-1}T)S = T^{-1}(TS) = T^{-1}I = T^{-1}$.

(iii) Let $V$ be the space of polynomials in $t$ over $K$; say, $p(t) = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n$. Let $T$ and $S$ be the operators on $V$ defined by

$$T(p(t)) = 0 + a_1 + a_2t + \cdots + a_nt^{n-1} \quad \text{and} \quad S(p(t)) = a_0t + a_1t^2 + \cdots + a_nt^{n+1}$$

We have

$$(TS)(p(t)) = T(S(p(t))) = T(a_0t + a_1t^2 + \cdots + a_nt^{n+1})$$
$$= a_0 + a_1t + \cdots + a_nt^n = p(t)$$

and so $TS = I$, the identity mapping. On the other hand, if $k \in K$ and $k \neq 0$, then $(ST)(k) = S(T(k)) = S(0) = 0 \neq k$. Accordingly, $ST \neq I$.

**6.40.** Let $S$ and $T$ be the linear operators on $\mathbf{R}^2$ defined by $S(x, y) = (0, x)$ and $T(x, y) = (x, 0)$. Show that $TS = 0$ but $ST \neq 0$. Also show that $T^2 = T$.

$(TS)(x, y) = T(S(x, y)) = T(0, x) = (0, 0)$. Since $TS$ assigns $0 = (0, 0)$ to every $(x, y) \in \mathbf{R}^2$, it is the zero mapping: $TS = 0$.

$(ST)(x, y) = S(T(x, y)) = S(x, 0) = (0, x)$. For example, $(ST)(4, 2) = (0, 4)$. Thus $ST \neq 0$, since it does not assign $0 = (0, 0)$ to every element of $\mathbf{R}^2$.

For any $(x, y) \in \mathbf{R}^2$, $T^2(x, y) = T(T(x, y)) = T(x, 0) = (x, 0) = T(x, y)$. Hence $T^2 = T$.

## MISCELLANEOUS PROBLEMS

**6.41.** Let $\{e_1, e_2, e_3\}$ be a basis of $V$ and $\{f_1, f_2\}$ a basis of $U$. Let $T: V \to U$ be linear. Furthermore, suppose

$$\begin{aligned} T(e_1) &= a_1f_1 + a_2f_2 \\ T(e_2) &= b_1f_1 + b_2f_2 \\ T(e_3) &= c_1f_1 + c_2f_2 \end{aligned} \qquad \text{and} \qquad A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

Show that, for any $v \in V$, $A[v]_e = [T(v)]_f$ where the vectors in $K^2$ and $K^3$ are written as column vectors.

Suppose $v = k_1e_1 + k_2e_2 + k_3e_3$; then $[v]_e = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}$. Also,

$$\begin{aligned} T(v) &= k_1T(e_1) + k_2T(e_2) + k_3T(e_3) \\ &= k_1(a_1f_1 + a_2f_2) + k_2(b_1f_1 + b_2f_2) + k_3(c_1f_1 + c_2f_2) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)f_1 + (a_2k_1 + b_2k_2 + c_2k_3)f_2 \end{aligned}$$

Accordingly,
$$[T(v)]_f = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \end{pmatrix}$$

Computing, we obtain

$$A[v]_e = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \end{pmatrix} = [T(v)]_f$$

**6.42.** Let $k$ be a nonzero scalar. Show that a linear map $T$ is singular if and only if $kT$ is singular. Hence $T$ is singular if and only if $-T$ is singular.

Suppose $T$ is singular. Then $T(v) = 0$ for some vector $v \neq 0$. Hence $(kT)(v) = kT(v) = k0 = 0$ and so $kT$ is singular.

Now suppose $kT$ is singular. Then $(kT)(w) = 0$ for some vector $w \neq 0$; hence $T(kw) = kT(w) = (kT)(w) = 0$. But $k \neq 0$ and $w \neq 0$ implies $kw \neq 0$; thus $T$ is also singular.

**6.43.** Let $E$ be a linear operator on $V$ for which $E^2 = E$. (Such an operator is termed a projection.) Let $U$ be the image of $E$ and $W$ the kernel. Show that: (i) if $u \in U$, then $E(u) = u$, i.e. $E$ is the identity map on $U$; (ii) if $E \neq I$, then $E$ is singular, i.e. $E(v) = 0$ for some $v \neq 0$; (iii) $V = U \oplus W$.

(i) If $u \in U$, the image of $E$, then $E(v) = u$ for some $v \in V$. Hence using $E^2 = E$, we have
$$u = E(v) = E^2(v) = E(E(v)) = E(u)$$

(ii) If $E \neq I$ then, for some $v \in V$, $E(v) = u$ where $v \neq u$. By (i), $E(u) = u$. Thus
$$E(v - u) = E(v) - E(u) = u - u = 0 \qquad \text{where} \quad v - u \neq 0$$

(iii) We first show that $V = U + W$. Let $v \in V$. Set $u = E(v)$ and $w = v - E(v)$. Then
$$v = E(v) + v - E(v) = u + w$$

By definition, $u = E(v) \in U$, the image of $E$. We now show that $w \in W$, the kernel of $E$:
$$E(w) = E(v - E(v)) = E(v) - E^2(v) = E(v) - E(v) = 0$$
and thus $w \in W$. Hence $V = U + W$.

We next show that $U \cap W = \{0\}$. Let $v \in U \cap W$. Since $v \in U$, $E(v) = v$ by (i). Since $v \in W$, $E(v) = 0$. Thus $v = E(v) = 0$ and so $U \cap W = \{0\}$.

The above two properties imply that $V = U \oplus W$.

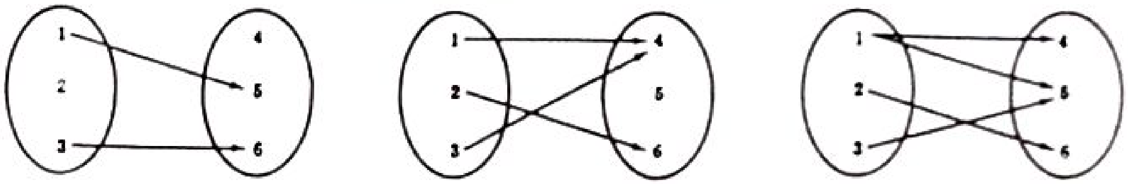
**6.44.** Show that a square matrix $A$ is invertible if and only if it is nonsingular. (Compare with Theorem 6.9, page 130.)

Recall that $A$ is invertible if and only if $A$ is row equivalent to the identity matrix $I$. Thus the following statements are equivalent. (i) $A$ is invertible. (ii) $A$ and $I$ are row equivalent. (iii) The equations $AX = 0$ and $IX = 0$ have the same solution space. (iv) $AX = 0$ has only the zero solution. (v) $A$ is nonsingular.

## Supplementary Problems

### MAPPINGS

**6.45.** State whether each diagram defines a mapping from $\{1, 2, 3\}$ into $\{4, 5, 6\}$.




**6.46.** Define each of the following mappings $f: \mathbf{R} \to \mathbf{R}$ by a formula:

(i) To each number let $f$ assign its square plus 3.

(ii) To each number let $f$ assign its cube plus twice the number.

(iii) To each number $\geq 3$ let $f$ assign the number squared, and to each number $< 3$ let $f$ assign the number $-2$.

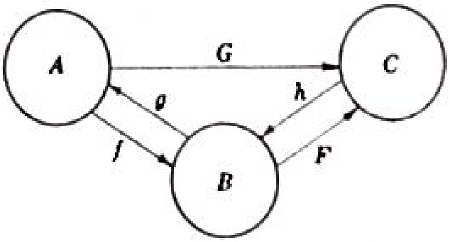
**6.47.** Let $f: \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = x^2 - 4x + 3$. Find (i) $f(4)$, (ii) $f(-3)$, (iii) $f(y - 2x)$, (iv) $f(x - 2)$.

**6.48.** Determine the number of different mappings from $\{a, b\}$ into $\{1, 2, 3\}$.

**6.49.** Let the mapping $g$ assign to each name in the set {Betty, Martin, David, Alan, Rebecca} the number of different letters needed to spell the name. Find (i) the graph of $g$, (ii) the image of $g$.

**6.50.** Sketch the graph of each mapping: (i) $f(x) = \frac{1}{2}x - 1$, (ii) $g(x) = 2x^2 - 4x - 3$.

**6.51.** The mappings $f: A \to B$, $g: B \to A$, $h: C \to B$, $F: B \to C$ and $G: A \to C$ are illustrated in the diagram below.




Determine whether each of the following defines a composition mapping and, if it does, find its domain and co-domain: (i) $g \circ f$, (ii) $h \circ f$, (iii) $F \circ f$, (iv) $G \circ f$, (v) $g \circ h$, (vi) $h \circ G \circ g$.

**6.52.** Let $f: \mathbf{R} \to \mathbf{R}$ and $g: \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = x^2 + 3x + 1$ and $g(x) = 2x - 3$. Find formulae defining the composition mappings (i) $f \circ g$, (ii) $g \circ f$, (iii) $g \circ g$, (iv) $f \circ f$.

**6.53.** For any mapping $f: A \to B$, show that $1_B \circ f = f = f \circ 1_A$.

**6.54.** For each of the following mappings $f : \mathbf{R} \to \mathbf{R}$ find a formula for the inverse mapping: (i) $f(x) = 3x - 7$, (ii) $f(x) = x^3 + 2$.

## LINEAR MAPPINGS

**6.55.** Show that the following mappings $F$ are linear:

(i) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (2x - y, x)$.

(ii) $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (z, x + y)$.

(iii) $F : \mathbf{R} \to \mathbf{R}^2$ defined by $F(x) = (2x, 3x)$.

(iv) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (ax + by, cx + dy)$ where $a, b, c, d \in \mathbf{R}$.

**6.56.** Show that the following mappings $F$ are not linear:

(i) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (x^2, y^2)$.

(ii) $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (x + 1, y + z)$.

(iii) $F : \mathbf{R} \to \mathbf{R}^2$ defined by $F(x) = (x, 1)$.

(iv) $F : \mathbf{R}^2 \to \mathbf{R}$ defined by $F(x, y) = |x - y|$.

**6.57.** Let $V$ be the vector space of polynomials in $t$ over $K$. Show that the mappings $T : V \to V$ and $S : V \to V$ defined below are linear:

$$T(a_0 + a_1 t + \cdots + a_n t^n) = a_0 t + a_1 t^2 + \cdots + a_n t^{n+1}$$
$$S(a_0 + a_1 t + \cdots + a_n t^n) = 0 + a_1 + a_2 t + \cdots + a_n t^{n-1}$$

**6.58.** Let $V$ be the vector space of $n \times n$ matrices over $K$, and let $M$ be an arbitrary matrix in $V$. Show that the first two mappings $T : V \to V$ are linear, but the third is not linear (unless $M = 0$): (i) $T(A) = MA$, (ii) $T(A) = MA - AM$, (iii) $T(A) = M + A$.

**6.59.** Find $T(a, b)$ where $T : \mathbf{R}^2 \to \mathbf{R}^3$ is defined by $T(1, 2) = (3, -1, 5)$ and $T(0, 1) = (2, 1, -1)$.

**6.60.** Find $T(a, b, c)$ where $T : \mathbf{R}^3 \to \mathbf{R}$ is defined by

$$T(1, 1, 1) = 3, \quad T(0, 1, -2) = 1 \quad \text{and} \quad T(0, 0, 1) = -2$$

**6.61.** Suppose $F : V \to U$ is linear. Show that, for any $v \in V$, $F(-v) = -F(v)$.

**6.62.** Let $W$ be a subspace of $V$. Show that the *inclusion map of $W$ into $V$*, denoted by $i : W \subset V$ and defined by $i(w) = w$, is linear.

## KERNEL AND IMAGE OF LINEAR MAPPINGS

**6.63.** For each of the following linear mappings $F$, find a basis and the dimension of (*a*) its image $U$ and (*b*) its kernel $W$:

(i) $F : \mathbf{R}^3 \to \mathbf{R}^3$ defined by $F(x, y, z) = (x + 2y, y - z, x + 2z)$.

(ii) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(x, y) = (x + y, x + y)$.

(iii) $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (x + y, y + z)$.

**6.64.** Let $V$ be the vector space of $2 \times 2$ matrices over $\mathbf{R}$ and let $M = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}$. Let $F : V \to V$ be the linear map defined by $F(A) = MA$. Find a basis and the dimension of (i) the kernel $W$ of $F$ and (ii) the image $U$ of $F$.

93, 90

**6.65.** Find a linear mapping $F : \mathbf{R}^3 \to \mathbf{R}^3$ whose image is generated by $(1, 2, 3)$ and $(4, 5, 6)$.

**6.66.** Find a linear mapping $F : \mathbf{R}^4 \to \mathbf{R}^3$ whose kernel is generated by $(1, 2, 3, 4)$ and $(0, 1, 1, 1)$.

**6.67.** Let $V$ be the vector space of polynomials in $t$ over $\mathbf{R}$. Let $D : V \to V$ be the differential operator: $D(f) = df/dt$. Find the kernel and image of $D$.

**6.68.** Let $F : V \to U$ be linear. Show that (i) the image of any subspace of $V$ is a subspace of $U$ and (ii) the preimage of any subspace of $U$ is a subspace of $V$.

**6.69.** Each of the following matrices determines a linear map from $\mathbf{R}^4$ into $\mathbf{R}^3$:

$$\text{(i)} \quad A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & -1 & 2 & -1 \\ 1 & -3 & 2 & -2 \end{pmatrix} \qquad \text{(ii)} \quad B = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{pmatrix}$$

Find a basis and the dimension of the image $U$ and the kernel $W$ of each map.

**6.70.** Let $T: \mathbf{C} \to \mathbf{C}$ be the conjugate mapping on the complex field $\mathbf{C}$. That is, $T(z) = \bar{z}$ where $z \in C$, or $T(a+bi) = a - bi$ where $a, b \in \mathbf{R}$. (i) Show that $T$ is not linear if $\mathbf{C}$ is viewed as a vector space over itself. (ii) Show that $T$ is linear if $\mathbf{C}$ is viewed as a vector space over the real field $\mathbf{R}$.

## OPERATIONS WITH LINEAR MAPPINGS

**6.71.** Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ and $G: \mathbf{R}^3 \to \mathbf{R}^2$ be defined by $F(x, y, z) = (y, x+z)$ and $G(x, y, z) = (2z, x - y)$. Find formulae defining the mappings $F + G$ and $3F - 2G$.

**6.72.** Let $H: \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $H(x, y) = (y, 2x)$. Using the mappings $F$ and $G$ in the preceding problem, find formulae defining the mappings: (i) $H \circ F$ and $H \circ G$, (ii) $F \circ H$ and $G \circ H$, (iii) $H \circ (F + G)$ and $H \circ F + H \circ G$.

**6.73.** Show that the following mappings $F$, $G$ and $H$ are linearly independent:

(i) $F, G, H \in \operatorname{Hom}(\mathbf{R}^2, \mathbf{R}^2)$ defined by

$$F(x, y) = (x, 2y), \quad G(x, y) = (y, x + y), \quad H(x, y) = (0, x).$$

(ii) $F, G, H \in \operatorname{Hom}(\mathbf{R}^3, \mathbf{R})$ defined by

$$F(x, y, z) = x + y + z, \quad G(x, y, z) = y + z, \quad H(x, y, z) = x - z.$$

**6.74.** For $F, G \in \operatorname{Hom}(V, U)$, show that $\operatorname{rank}(F + G) \leq \operatorname{rank} F + \operatorname{rank} G$. (Here $V$ has finite dimension.)

**6.75.** Let $F: V \to U$ and $G: U \to V$ be linear. Show that if $F$ and $G$ are nonsingular then $G \circ F$ is nonsingular. Give an example where $G \circ F$ is nonsingular but $G$ is not.

**6.76.** Prove that $\operatorname{Hom}(V, U)$ does satisfy all the required axioms of a vector space. That is, prove Theorem 6.6, page 128.

## ALGEBRA OF LINEAR OPERATORS

**6.77.** Let $S$ and $T$ be the linear operators on $\mathbf{R}^2$ defined by $S(x, y) = (x + y, 0)$ and $T(x, y) = (-y, x)$. Find formulae defining the operators $S + T$, $5S - 3T$, $ST$, $TS$, $S^2$ and $T^2$.

**6.78.** Let $T$ be the linear operator on $\mathbf{R}^2$ defined by $T(x, y) = (x + 2y, 3x + 4y)$. Find $p(T)$ where $p(t) = t^2 - 5t - 2$.

**6.79.** Show that each of the following operators $T$ on $\mathbf{R}^3$ is invertible, and find a formula for $T^{-1}$: (i) $T(x, y, z) = (x - 3y - 2z, y - 4z, z)$; (ii) $T(x, y, z) = (x + z, x - z, y)$.

**6.80.** Suppose $S$ and $T$ are linear operators on $V$ and that $S$ is nonsingular. Assume $V$ has finite dimension. Show that $\operatorname{rank}(ST) = \operatorname{rank}(TS) = \operatorname{rank} T$.

**6.81.** Suppose $V = U \oplus W$. Let $E_1$ and $E_2$ be the linear operators on $V$ defined by $E_1(v) = u$, $E_2(v) = w$, where $v = u + w$, $u \in U$, $w \in W$. Show that (i) $E_1^2 = E_1$ and $E_2^2 = E_2$, i.e. that $E_1$ and $E_2$ are "projections"; (ii) $E_1 + E_2 = I$, the identity mapping; (iii) $E_1E_2 = 0$ and $E_2E_1 = 0$.

**6.82.** Let $E_1$ and $E_2$ be linear operators on $V$ satisfying (i), (ii) and (iii) of Problem 6.81. Show that $V$ is the direct sum of the image of $E_1$ and the image of $E_2$: $V = \operatorname{Im} E_1 \oplus \operatorname{Im} E_2$.

**6.83.** Show that if the linear operators $S$ and $T$ are invertible, then $ST$ is invertible and $(ST)^{-1} = T^{-1}S^{-1}$.

6.84. Let $V$ have finite dimension, and let $T$ be a linear operator on $V$ such that $\text{rank}(T^2) = \text{rank}\, T$. Show that $\text{Ker}\, T \cap \text{Im}\, T = \{0\}$.

## MISCELLANEOUS PROBLEMS

6.85. Suppose $T: K^n \to K^m$ is a linear mapping. Let $\{e_1, \ldots, e_n\}$ be the usual basis of $K^n$ and let $A$ be the $m \times n$ matrix whose columns are the vectors $T(e_1), \ldots, T(e_n)$ respectively. Show that, for every vector $v \in K^n$, $T(v) = Av$, where $v$ is written as a column vector.

6.86. Suppose $F: V \to U$ is linear and $k$ is a nonzero scalar. Show that the maps $F$ and $kF$ have the same kernel and the same image.

6.87. Show that if $F: V \to U$ is onto, then $\dim U \le \dim V$. Determine all linear maps $T: \mathbf{R}^3 \to \mathbf{R}^4$ which are onto.

6.88. Find those theorems of Chapter 3 which prove that the space of $n$-square matrices over $K$ is an associative algebra over $K$.

6.89. Let $T: V \to U$ be linear and let $W$ be a subspace of $V$. The *restriction of $T$ to $W$* is the map $T_W: W \to U$ defined by $T_W(w) = T(w)$, for every $w \in W$. Prove the following. (i) $T_W$ is linear. (ii) $\text{Ker}\, T_W = \text{Ker}\, T \cap W$. (iii) $\text{Im}\, T_W = T(W)$.

6.90. Two operators $S, T \in A(V)$ are said to be *similar* if there exists an invertible operator $P \in A(V)$ for which $S = P^{-1}TP$. Prove the following. (i) Similarity of operators is an equivalence relation. (ii) Similar operators have the same rank (when $V$ has finite dimension).

# Answers to Supplementary Problems

6.45. (i) No. (ii) Yes. (iii) No.

6.46. (i) $f(x) = x^2 + 3$, (ii) $f(x) = x^3 + 2x$, (iii) $f(x) = \begin{cases} x^2 & \text{if } x \ge 3 \\ -2 & \text{if } x < 3 \end{cases}$

6.47 (i) 3, (ii) 24, (iii) $y^2 - 4xy + 4x^2 - 4y + 8x + 3$, (iv) $x^2 - 8x + 15$.

6.48. Nine.

6.49. (i) {(Betty, 4), (Martin, 6), (David, 4), (Alan, 3), (Rebecca, 5)}.
(ii) Image of $g = \{3, 4, 5, 6\}$.

6.51. (i) $(g \circ f): A \to A$, (ii) No, (iii) $(F \circ f): A \to C$, (iv) No, (v) $(g \circ h): C \to A$, (vi) $(h \circ G \circ g): B \to B$.

6.52. (i) $(f \circ g)(x) = 4x^2 - 6x + 1$ (iii) $(g \circ g)(x) = 4x - 9$
(ii) $(g \circ f)(x) = 2x^2 + 6x - 1$ (iv) $(f \circ f)(x) = x^4 + 6x^3 + 14x^2 + 15x + 5$

6.54. (i) $f^{-1}(x) = (x + 7)/3$, (ii) $f^{-1}(x) = \sqrt[3]{x - 2}$.

6.59. $T(a, b) = (-a + 2b, -3a + b, 7a - b)$.

6.60. $T(a, b, c) = 8a - 3b - 2c$.

6.61. $F(v) + F(-v) = F(v + (-v)) = F(0) = 0$; hence $F(-v) = -F(v)$.

6.63. (i) (a) $\{(1, 0, 1), (0, 1, -2)\}$, $\dim U = 2$; (b) $\{(2, -1, -1)\}$, $\dim W = 1$.
(ii) (a) $\{(1, 1)\}$, $\dim U = 1$; (b) $\{(1, -1)\}$, $\dim W = 1$.
(iii) (a) $\{(1, 0), (0, 1)\}$, $\dim U = 2$; (b) $\{(1, -1, 1)\}$, $\dim W = 1$.

6.64. (i) $\left\{\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\right\}$ basis of $\operatorname{Ker} F$; $\dim(\operatorname{Ker} F) = 2$.

(ii) $\left\{\begin{pmatrix} 1 & 0 \\ -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -2 \end{pmatrix}\right\}$ basis of $\operatorname{Im} F$; $\dim(\operatorname{Im} F) = 2$.

6.65. $F(x, y, z) = (x + 4y,\ 2x + 5y,\ 3x + 6y)$.

6.66. $F(x, y, z, w) = (x + y - z,\ 2x + y - w,\ 0)$.

6.67. The kernel of $D$ is the set of constant polynomials. The image of $D$ is the entire space $V$.

6.69. (i) (a) $\{(1, 2, 1), (0, 1, 1)\}$ basis of $\operatorname{Im} A$; $\dim(\operatorname{Im} A) = 2$.
(b) $\{(4, -2, -5, 0), (1, -3, 0, 5)\}$ basis of $\operatorname{Ker} A$; $\dim(\operatorname{Ker} A) = 2$.
(ii) (a) $\operatorname{Im} B = \mathbf{R}^3$. (b) $\{(-1, 2/3, 1, 1)\}$ basis of $\operatorname{Ker} B$; $\dim(\operatorname{Ker} B) = 1$.

6.71. $(F + G)(x, y, z) = (y + 2z,\ 2x - y + z)$, $(3F - 2G)(x, y, z) = (3y - 4z,\ x + 2y + 3z)$.

6.72. (i) $(H \circ F)(x, y, z) = (x + z,\ 2y)$, $(H \circ G)(x, y, z) = (x - y,\ 4z)$. (ii) Not defined.
(iii) $(H \circ (F + G))(x, y, z) = (H \circ F + H \circ G)(x, y, z) = (2x - y + z,\ 2y + 4z)$.

6.77. $(S + T)(x, y) = (x, x)$ $\qquad$ $(ST)(x, y) = (x - y, 0)$
$(5S - 3T)(x, y) = (5x + 8y,\ -3x)$ $\qquad$ $(TS)(x, y) = (0,\ x + y)$

$S^2(x, y) = (x + y, 0)$; note that $S^2 = S$.
$T^2(x, y) = (-x, -y)$; note that $T^2 + I = 0$, hence $T$ is a zero of $x^2 + 1$.

6.78. $p(T) = 0$

6.79. (i) $T^{-1}(r, s, t) = (14t + 3s + r,\ 4t + s,\ t)$, (ii) $T^{-1}(r, s, t) = (\frac{1}{2}r + \frac{1}{2}s,\ t,\ \frac{1}{2}r - \frac{1}{2}s)$.

6.87. There are no linear maps from $\mathbf{R}^3$ into $\mathbf{R}^4$ which are onto.

# Chapter 7

## Matrices and Linear Operators

### INTRODUCTION

Suppose $\{e_1, \ldots, e_n\}$ is a basis of a vector space $V$ over a field $K$ and, for $v \in V$, suppose $v = a_1e_1 + a_2e_2 + \cdots + a_ne_n$. Then the coordinate vector of $v$ relative to $\{e_i\}$, which we write as a column vector unless otherwise specified or implied, is

$$[v]_e = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$$

Recall that the mapping $v \mapsto [v]_e$, determined by the basis $\{e_i\}$, is an isomorphism from $V$ onto the space $K^n$.

In this chapter we show that there is also an isomorphism, determined by the basis $\{e_i\}$, from the algebra $A(V)$ of linear operators on $V$ onto the algebra $\mathcal{A}$ of $n$-square matrices over $K$.

A similar result also holds for linear mappings $F: V \to U$, from one space into another.

### MATRIX REPRESENTATION OF A LINEAR OPERATOR

Let $T$ be a linear operator on a vector space $V$ over a field $K$ and suppose $\{e_1, \ldots, e_n\}$ is a basis of $V$. Now $T(e_1), \ldots, T(e_n)$ are vectors in $V$ and so each is a linear combination of the elements of the basis $\{e_i\}$:

$$\begin{aligned} T(e_1) &= a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n \\ T(e_2) &= a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(e_n) &= a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{nn}e_n \end{aligned}$$

The following definition applies.

**Definition:** The transpose of the above matrix of coefficients, denoted by $[T]_e$ or $[T]$, is called the *matrix representation of $T$ relative to the basis $\{e_i\}$* or simply the *matrix of $T$ in the basis $\{e_i\}$*:

$$[T]_e = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}$$

**Example 7.1:** Let $V$ be the vector space of polynomials in $t$ over $\mathbf{R}$ of degree $\le 3$, and let $D: V \to V$ be the differential operator defined by $D(p(t)) = d(p(t))/dt$. We compute the matrix of $D$ in the basis $\{1, t, t^2, t^3\}$. We have:

$$\begin{aligned} D(1) &= 0 = 0 + 0t + 0t^2 + 0t^3 \\ D(t) &= 1 = 1 + 0t + 0t^2 + 0t^3 \\ D(t^2) &= 2t = 0 + 2t + 0t^2 + 0t^3 \\ D(t^3) &= 3t^2 = 0 + 0t + 3t^2 + 0t^3 \end{aligned}$$

Accordingly,

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

**Example 7.2:** Let $T$ be the linear operator on $\mathbf{R}^2$ defined by $T(x, y) = (4x - 2y, 2x + y)$. We compute the matrix of $T$ in the basis $\{f_1 = (1, 1), f_2 = (-1, 0)\}$. We have

$$T(f_1) = T(1, 1) = (2, 3) = 3(1, 1) + (-1, 0) = 3f_1 + f_2$$

$$T(f_2) = T(-1, 0) = (-4, -2) = -2(1, 1) + 2(-1, 0) = -2f_1 + 2f_2$$

Accordingly, $[T]_f = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$.

**Remark:** Recall that any $n$-square matrix $A$ over $K$ defines a linear operator on $K^n$ by the map $v \mapsto Av$ (where $v$ is written as a column vector). We show (Problem 7.7) that the matrix representation of this operator is precisely the matrix $A$ if we use the usual basis of $K^n$.

Our first theorem tells us that the "action" of an operator $T$ on a vector $v$ is preserved by its matrix representation:

**Theorem 7.1:** Let $\{e_1, \ldots, e_n\}$ be a basis of $V$ and let $T$ be any operator on $V$. Then, for any vector $v \in V$, $[T]_e [v]_e = [T(v)]_e$.

Sit

That is, if we multiply the coordinate vector of $v$ by the matrix representation of $T$, then we obtain the coordinate vector of $T(v)$.

**Example 7.3:** Consider the differential operator $D: V \to V$ in Example 7.1. Let

$$p(t) = a + bt + ct^2 + dt^3 \quad \text{and so} \quad D(p(t)) = b + 2ct + 3dt^2$$

Hence, relative to the basis $\{1, t, t^2, t^3\}$,

$$[p(t)] = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \text{and} \quad [D(p(t))] = \begin{pmatrix} b \\ 2c \\ 3d \\ 0 \end{pmatrix}$$

We show that Theorem 7.1 does hold here:

$$[D][p(t)] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} b \\ 2c \\ 3d \\ 0 \end{pmatrix} = [D(p(t))]$$

**Example 7.4:** Consider the linear operator $T: \mathbf{R}^2 \to \mathbf{R}^2$ in Example 7.2: $T(x, y) = (4x - 2y, 2x + y)$. Let $v = (5, 7)$. Then

$$v = (5, 7) = 7(1, 1) + 2(-1, 0) = 7f_1 + 2f_2$$

$$T(v) = (6, 17) = 17(1, 1) + 11(-1, 0) = 17f_1 + 11f_2$$

where $f_1 = (1, 1)$ and $f_2 = (-1, 0)$. Hence, relative to the basis $\{f_1, f_2\}$,

$$[v]_f = \begin{pmatrix} 7 \\ 2 \end{pmatrix} \quad \text{and} \quad [T(v)]_f = \begin{pmatrix} 17 \\ 11 \end{pmatrix}$$

Using the matrix $[T]_f$ in Example 7.2, we verify that Theorem 7.1 holds here:

$$[T]_f [v]_f = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 2 \end{pmatrix} = \begin{pmatrix} 17 \\ 11 \end{pmatrix} = [T(v)]_f$$

matrix

Now we have associated a matrix $[T]_e$ to each $T$ in $A(V)$, the algebra of linear operators on $V$. By our first theorem the action of an individual operator $T$ is preserved by this representation. The next two theorems tell us that the three basic operations with these operators

(i) addition, (ii) scalar multiplication, (iii) composition

are also preserved.

**Theorem 7.2:** Let $\{e_1, \ldots, e_n\}$ be a basis of $V$ over $K$, and let $\mathcal{A}$ be the algebra of $n$-square matrices over $K$. Then the mapping $T \mapsto [T]_e$ is a vector space isomorphism from $A(V)$ onto $\mathcal{A}$. That is, the mapping is one-one and onto and, for any $S, T \in A(V)$ and any $k \in K$,

$$[T+S]_e = [T]_e + [S]_e \quad \text{and} \quad [kT]_e = k[T]_e$$

3/ Si

**Theorem 7.3:** For any operators $S, T \in A(V)$, $[ST]_e = [S]_e[T]_e$.

We illustrate the above theorems in the case $\dim V = 2$. Suppose $\{e_1, e_2\}$ is a basis of $V$, and $T$ and $S$ are operators on $V$ for which

$$T(e_1) = a_1e_1 + a_2e_2 \qquad S(e_1) = c_1e_1 + c_2e_2$$
$$T(e_2) = b_1e_1 + b_2e_2 \qquad S(e_2) = d_1e_1 + d_2e_2$$

Then $$[T]_e = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad \text{and} \quad [S]_e = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}$$

Now we have $$(T+S)(e_1) = T(e_1) + S(e_1) = a_1e_1 + a_2e_2 + c_1e_1 + c_2e_2 = (a_1+c_1)e_1 + (a_2+c_2)e_2$$

$$(T+S)(e_2) = T(e_2) + S(e_2) = b_1e_1 + b_2e_2 + d_1e_1 + d_2e_2 = (b_1+d_1)e_1 + (b_2+d_2)e_2$$

Thus
$$[T+S]_e = \begin{pmatrix} a_1+c_1 & b_1+d_1 \\ a_2+c_2 & b_2+d_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} + \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = [T]_e + [S]_e$$

Also, for $k \in K$, we have

$$(kT)(e_1) = kT(e_1) = k(a_1e_1 + a_2e_2) = ka_1e_1 + ka_2e_2$$
$$(kT)(e_2) = kT(e_2) = k(b_1e_1 + b_2e_2) = kb_1e_1 + kb_2e_2$$

Hence $$[kT]_e = \begin{pmatrix} ka_1 & kb_1 \\ ka_2 & kb_2 \end{pmatrix} = k\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = k[T]_e$$

Finally, we have

$$(ST)(e_1) = S(T(e_1)) = S(a_1e_1 + a_2e_2) = a_1S(e_1) + a_2S(e_2)$$
$$= a_1(c_1e_1 + c_2e_2) + a_2(d_1e_1 + d_2e_2)$$
$$= (a_1c_1 + a_2d_1)e_1 + (a_1c_2 + a_2d_2)e_2$$

$$(ST)(e_2) = S(T(e_2)) = S(b_1e_1 + b_2e_2) = b_1S(e_1) + b_2S(e_2)$$
$$= b_1(c_1e_1 + c_2e_2) + b_2(d_1e_1 + d_2e_2)$$
$$= (b_1c_1 + b_2d_1)e_1 + (b_1c_2 + b_2d_2)e_2$$

Accordingly,

$$[ST]_e = \begin{pmatrix} a_1c_1 + a_2d_1 & b_1c_1 + b_2d_1 \\ a_1c_2 + a_2d_2 & b_1c_2 + b_2d_2 \end{pmatrix} = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = [S]_e[T]_e$$

## CHANGE OF BASIS

We have shown that we can represent vectors by $n$-tuples (column vectors) and linear operators by matrices once we have selected a basis. We ask the following natural question: How does our representation change if we select another basis? In order to answer this question, we first need a definition.

**Definition:** Let $\{e_1, \ldots, e_n\}$ be a basis of $V$ and let $\{f_1, \ldots, f_n\}$ be another basis. Suppose

$$\begin{aligned} f_1 &= a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n \\ f_2 &= a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n &= a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{nn}e_n \end{aligned}$$

Then the transpose $P$ of the above matrix of coefficients is termed the *transition matrix* from the "old" basis $\{e_i\}$ to the "new" basis $\{f_i\}$:

$$P = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

We comment that since the vectors $f_1, \ldots, f_n$ are linearly independent, the matrix $P$ is invertible (Problem 5.47). In fact, its inverse $P^{-1}$ is the transition matrix from the basis $\{f_i\}$ back to the basis $\{e_i\}$.

**Example 7.5:** Consider the following two bases of $\mathbf{R}^2$:

$$\{e_1 = (1,0),\ e_2 = (0,1)\} \quad \text{and} \quad \{f_1 = (1,1),\ f_2 = (-1,0)\}$$

Then
$$\begin{aligned} f_1 &= (1,1) = (1,0) + (0,1) = e_1 + e_2 \\ f_2 &= (-1,0) = -(1,0) + 0(0,1) = -e_1 + 0e_2 \end{aligned}$$

Hence the transition matrix $P$ from the basis $\{e_i\}$ to the basis $\{f_i\}$ is

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

We also have
$$\begin{aligned} e_1 &= (1,0) = 0(1,1) - (-1,0) = 0f_1 - f_2 \\ e_2 &= (0,1) = (1,1) + (-1,0) = f_1 + f_2 \end{aligned}$$

Hence the transition matrix $Q$ from the basis $\{f_i\}$ back to the basis $\{e_i\}$ is

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

Observe that $P$ and $Q$ are inverses:

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

We now show how coordinate vectors are affected by a change of basis.

**Theorem 7.4:** Let $P$ be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space $V$. Then, for any vector $v \in V$, $P[v]_f = [v]_e$. Hence $[v]_f = P^{-1}[v]_e$.

We emphasize that even though $P$ is called the transition matrix from the old basis $\{e_i\}$ to the new basis $\{f_i\}$, its effect is to transform the coordinates of a vector in the new basis $\{f_i\}$ back to the coordinates in the old basis $\{e_i\}$.

We illustrate the above theorem in the case $\dim V = 3$. Suppose $P$ is the transition matrix from a basis $\{e_1, e_2, e_3\}$ of $V$ to a basis $\{f_1, f_2, f_3\}$ of $V$; say,

$$\begin{aligned} f_1 &= a_1e_1 + a_2e_2 + a_3e_3 \\ f_2 &= b_1e_1 + b_2e_2 + b_3e_3 \\ f_3 &= c_1e_1 + c_2e_2 + c_3e_3 \end{aligned} \qquad \text{Hence} \quad P = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Now suppose $v \in V$ and, say, $v = k_1f_1 + k_2f_2 + k_3f_3$. Then, substituting for the $f_i$ from above, we obtain

$$\begin{aligned} v &= k_1(a_1e_1 + a_2e_2 + a_3e_3) + k_2(b_1e_1 + b_2e_2 + b_3e_3) + k_3(c_1e_1 + c_2e_2 + c_3e_3) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)e_1 + (a_2k_1 + b_2k_2 + c_2k_3)e_2 + (a_3k_1 + b_3k_2 + c_3k_3)e_3 \end{aligned}$$

Thus

$$[v]_f = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} \qquad \text{and} \qquad [v]_e = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix}$$

Accordingly,

$$P[v]_f = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{pmatrix} = [v]_e$$

Also, multiplying the above equation by $P^{-1}$, we have

$$P^{-1}[v]_e = P^{-1}P[v]_f = I[v]_f = [v]_f$$

**Example 7.6:** Let $v = (a, b) \in \mathbf{R}^2$. Then, for the bases of $\mathbf{R}^2$ in the preceding example,

$$\begin{aligned} v &= (a, b) = a(1, 0) + b(0, 1) = ae_1 + be_2 \\ v &= (a, b) = b(1, 1) + (b - a)(-1, 0) = bf_1 + (b - a)f_2 \end{aligned}$$

Hence

$$[v]_e = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{and} \qquad [v]_f = \begin{pmatrix} b \\ b - a \end{pmatrix}$$

By the preceding example, the transition matrix $P$ from $\{e_i\}$ to $\{f_i\}$ and its inverse $P^{-1}$ are given by

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad P^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

We verify the result of Theorem 7.4:

$$P[v]_f = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} b \\ b - a \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} = [v]_e$$

$$P^{-1}[v]_e = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ b - a \end{pmatrix} = [v]_f$$

The next theorem shows how matrix representations of linear operators are affected by a change of basis.

**Theorem 7.5:** Let $P$ be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space $V$. Then for any linear operator $T$ on $V$, $[T]_f = P^{-1}[T]_eP$.

**Example 7.7:** Let $T$ be the linear operator on $\mathbf{R}^2$ defined by $T(x, y) = (4x - 2y, 2x + y)$. Then for the bases of $\mathbf{R}^2$ in Example 7.5, we have

$$\begin{aligned} T(e_1) &= T(1, 0) = (4, 2) = 4(1, 0) + 2(0, 1) = 4e_1 + 2e_2 \\ T(e_2) &= T(0, 1) = (-2, 1) = -2(1, 0) + (0, 1) = -2e_1 + e_2 \end{aligned}$$

Accordingly,

$$[T]_e = \begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}$$

We compute $[T]_f$ using Theorem 7.5:

$$[T]_f = P^{-1}[T]_e P = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}$$

Note that this agrees with the derivation of $[T]_f$ in Example 7.2.

**Remark:** Suppose $P = (a_{ij})$ is any $n$-square invertible matrix over a field $K$. Now if $\{e_1, \ldots, e_n\}$ is a basis of a vector space $V$ over $K$, then the $n$ vectors

$$f_i = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, \ldots, n$$

are linearly independent (Problem 5.47) and so form another basis of $V$. Furthermore, $P$ is the transition matrix from the basis $\{e_i\}$ to the basis $\{f_i\}$. Accordingly, if $A$ is any matrix representation of a linear operator $T$ on $V$, then the matrix $B = P^{-1}AP$ is also a matrix representation of $T$.

## SIMILARITY

Suppose $A$ and $B$ are square matrices for which there exists an invertible matrix $P$ such that $B = P^{-1}AP$. Then $B$ is said to be *similar* to $A$ or is said to be obtained from $A$ by a *similarity transformation*. We show (Problem 7.16) that similarity of matrices is an equivalence relation. Thus by Theorem 7.5 and the above remark, we have the following basic result.

**Theorem 7.6:** Two matrices $A$ and $B$ represent the same linear operator $T$ if and only if they are similar to each other.

That is, all the matrix representations of the linear operator $T$ form an equivalence class of similar matrices.

A linear operator $T$ is said to be *diagonalizable* if for some basis $\{e_i\}$ it is represented by a diagonal matrix; the basis $\{e_i\}$ is then said to *diagonalize* $T$. The preceding theorem gives us the following result.

**Theorem 7.7:** Let $A$ be a matrix representation of a linear operator $T$. Then $T$ is diagonalizable if and only if there exists an invertible matrix $P$ such that $P^{-1}AP$ is a diagonal matrix.

That is, $T$ is diagonalizable if and only if its matrix representation can be diagonalized by a similarity transformation.

We emphasize that not every operator is diagonalizable. However, we will show (Chapter 10) that every operator $T$ can be represented by certain "standard" matrices called its *normal* or *canonical* forms. We comment now that that discussion will require some theory of fields, polynomials and determinants.

Now suppose $f$ is a function on square matrices which assigns the same value to similar matrices; that is, $f(A) = f(B)$ whenever $A$ is similar to $B$. Then $f$ induces a function, also denoted by $f$, on linear operators $T$ in the following natural way: $f(T) = f([T]_e)$, where $\{e_i\}$ is any basis. The function is well-defined by the preceding theorem.

The *determinant* is perhaps the most important example of the above type of functions. Another important example follows.

**Example 7.8:** The *trace* of a square matrix $A = (a_{ij})$, written $\operatorname{tr}(A)$, is defined to be the sum of its diagonal elements:

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

We show (Problem 7.17) that similar matrices have the same trace. Thus we can speak of the trace of a linear operator $T$; it is the trace of any one of its matrix representations: $\operatorname{tr}(T) = \operatorname{tr}([T]_e)$.

## MATRICES AND LINEAR MAPPINGS

We now consider the general case of linear mappings from one space into another. Let $V$ and $U$ be vector spaces over the same field $K$ and, say, $\dim V = m$ and $\dim U = n$. Furthermore, let $\{e_1, \ldots, e_m\}$ and $\{f_1, \ldots, f_n\}$ be arbitrary but fixed bases of $V$ and $U$ respectively.

Suppose $F: V \to U$ is a linear mapping. Then the vectors $F(e_1), \ldots, F(e_m)$ belong to $U$ and so each is a linear combination of the $f_i$:

$$\begin{aligned} F(e_1) &= a_{11}f_1 + a_{12}f_2 + \cdots + a_{1n}f_n \\ F(e_2) &= a_{21}f_1 + a_{22}f_2 + \cdots + a_{2n}f_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F(e_m) &= a_{m1}f_1 + a_{m2}f_2 + \cdots + a_{mn}f_n \end{aligned}$$

The transpose of the above matrix of coefficients, denoted by $[F]_e^f$ is called the *matrix representation* of $F$ relative to the bases $\{e_i\}$ and $\{f_i\}$, or the matrix of $F$ in the bases $\{e_i\}$ and $\{f_i\}$:

$$[F]_e^f = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}$$

The following theorems apply.

**Theorem 7.8:** For any vector $v \in V$, $[F]_e^f [v]_e = [F(v)]_f$.

That is, multiplying the coordinate vector of $v$ in the basis $\{e_i\}$ by the matrix $[F]_e^f$, we obtain the coordinate vector of $F(v)$ in the basis $\{f_i\}$.

**Theorem 7.9:** The mapping $F \mapsto [F]_e^f$ is an isomorphism from $\operatorname{Hom}(V, U)$ onto the vector space of $n \times m$ matrices over $K$. That is, the mapping is one-one and onto and, for any $F, G \in \operatorname{Hom}(V, U)$ and any $k \in K$,

$$[F+G]_e^f = [F]_e^f + [G]_e^f \quad \text{and} \quad [kF]_e^f = k[F]_e^f$$

**Remark:** Recall that any $n \times m$ matrix $A$ over $K$ has been identified with the linear mapping from $K^m$ into $K^n$ given by $v \mapsto Av$. Now suppose $V$ and $U$ are vector spaces over $K$ of dimensions $m$ and $n$ respectively, and suppose $\{e_i\}$ is a basis of $V$ and $\{f_i\}$ is a basis of $U$. Then in view of the preceding theorem, we shall also identify $A$ with the linear mapping $F: V \to U$ given by $[F(v)]_f = A[v]_e$. We comment that if other bases of $V$ and $U$ are given, then $A$ is identified with another linear mapping from $V$ into $U$.

**Theorem 7.10:** Let $\{e_i\}$, $\{f_i\}$ and $\{g_i\}$ be bases of $V$, $U$ and $W$ respectively. Let $F: V \to U$ and $G: U \to W$ be linear mappings. Then

$$[G \circ F]_e^g = [G]_f^g [F]_e^f$$

That is, relative to the appropriate bases, the matrix representation of the composition of two linear mappings is equal to the product of the matrix representations of the individual mappings.

We lastly show how the matrix representation of a linear mapping $F: V \to U$ is affected when new bases are selected.

**Theorem 7.11:** Let $P$ be the transition matrix from a basis $\{e_i\}$ to a basis $\{e_i'\}$ in $V$, and let $Q$ be the transition matrix from a basis $\{f_i\}$ to a basis $\{f_i'\}$ in $U$. Then for any linear mapping $F: V \to U$,

$$[F]_{e'}^{f'} = Q^{-1} [F]_e^f P$$

Thus in particular,

$$[F]_{e}^{f'} = Q^{-1}[F]_{e}^{f}$$

i.e. when the change of basis only takes place in $U$; and

$$[F]_{e'}^{f} = [F]_{e}^{f}P$$

i.e. when the change of basis only takes place in $V$.

Note that Theorems 7.1, 7.2, 7.3 and 7.5 are special cases of Theorems 7.8, 7.9, 7.10 and 7.11 respectively.

The next theorem shows that every linear mapping from one space into another can be represented by a very simple matrix.

**Theorem 7.12:** Let $F: V \to U$ be linear and, say, $\operatorname{rank} F = r$. Then there exist bases of $V$ and of $U$ such that the matrix representation of $F$ has the form

$$A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

where $I$ is the $r$-square identity matrix. We call $A$ the *normal* or *canonical* form of $F$.

## WARNING

As noted previously, some texts write the operator symbol $T$ to the right of the vector $v$ on which it acts, that is,

$$vT \quad \text{instead of} \quad T(v)$$

In such texts, vectors and operators are represented by $n$-tuples and matrices which are the transposes of those appearing here. That is, if

$$v = k_1e_1 + k_2e_2 + \cdots + k_ne_n$$

then they write

$$[v]_e = (k_1, k_2, \ldots, k_n) \quad \text{instead of} \quad [v]_e = \begin{pmatrix} k_1 \\ k_2 \\ \cdot\cdot \\ k_n \end{pmatrix}$$

And if

$$\begin{aligned} T(e_1) &= a_1e_1 + a_2e_2 + \cdots + a_ne_n \\ T(e_2) &= b_1e_1 + b_2e_2 + \cdots + b_ne_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(e_n) &= c_1e_1 + c_2e_2 + \cdots + c_ne_n \end{aligned}$$

then they write

$$[T]_e \quad \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ b_1 & b_2 & \ldots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ c_1 & c_2 & \ldots & c_n \end{pmatrix} \quad \text{instead of} \quad [T]_e = \begin{pmatrix} a_1 & b_1 & \ldots & c_1 \\ a_2 & b_2 & \ldots & c_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_n & b_n & \ldots & c_n \end{pmatrix}$$

This is also true for the transition matrix from one basis to another and for matrix representations of linear mappings $F: V \to U$. We comment that such texts have theorems which are analogous to the ones appearing here.

# Solved Problems

## MATRIX REPRESENTATIONS OF LINEAR OPERATORS

**7.1.** Find the matrix representation of each of the following operators $T$ on $\mathbf{R}^2$ relative to the usual basis $\{e_1 = (1,0),\ e_2 = (0,1)\}$:

(i) $T(x, y) = (2y, 3x - y)$, (ii) $T(x, y) = (3x - 4y, x + 5y)$.

Note first that if $(a, b) \in \mathbf{R}^2$, then $(a, b) = ae_1 + be_2$.

(i) $$\begin{aligned} T(e_1) &= T(1,0) = (0,3) = 0e_1 + 3e_2 \\ T(e_2) &= T(0,1) = (2,-1) = 2e_1 - e_2 \end{aligned} \quad \text{and} \quad [T]_e = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$$

(ii) $$\begin{aligned} T(e_1) &= T(1,0) = (3,1) = 3e_1 + e_2 \\ T(e_2) &= T(0,1) = (-4,5) = -4e_1 + 5e_2 \end{aligned} \quad \text{and} \quad [T]_e = \begin{pmatrix} 3 & -4 \\ 1 & 5 \end{pmatrix}$$

**7.2.** Find the matrix representation of each operator $T$ in the preceding problem relative to the basis $\{f_1 = (1,3),\ f_2 = (2,5)\}$.

We must first find the coordinates of an arbitrary vector $(a, b) \in \mathbf{R}^2$ with respect to the basis $\{f_i\}$. We have

$$(a, b) = x(1,3) + y(2,5) = (x + 2y, 3x + 5y)$$

or $$x + 2y = a \quad \text{and} \quad 3x + 5y = b$$

or $$x = 2b - 5a \quad \text{and} \quad y = 3a - b$$

Thus $$(a, b) = (2b - 5a)f_1 + (3a - b)f_2$$

(i) We have $T(x, y) = (2y, 3x - y)$. Hence

$$\begin{aligned} T(f_1) &= T(1,3) = (6,0) = -30f_1 + 18f_2 \\ T(f_2) &= T(2,5) = (10,1) = -48f_1 + 29f_2 \end{aligned} \quad \text{and} \quad [T]_f = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$$

(ii) We have $T(x, y) = (3x - 4y, x + 5y)$. Hence

$$\begin{aligned} T(f_1) &= T(1,3) = (-9,16) = 77f_1 - 43f_2 \\ T(f_2) &= T(2,5) = (-14,27) = 124f_1 - 69f_2 \end{aligned} \quad \text{and} \quad [T]_f = \begin{pmatrix} 77 & 124 \\ -43 & -69 \end{pmatrix}$$

**7.3.** Suppose that $T$ is the linear operator on $\mathbf{R}^3$ defined by

$$T(x, y, z) = (a_1x + a_2y + a_3z,\ b_1x + b_2y + b_3z,\ c_1x + c_2y + c_3z)$$

Show that the matrix of $T$ in the usual basis $\{e_i\}$ is given by

$$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

That is, the rows of $[T]_e$ are obtained from the coefficients of $x$, $y$ and $z$ in the components of $T(x, y, z)$.

$$\begin{aligned} T(e_1) &= T(1,0,0) = (a_1, b_1, c_1) = a_1e_1 + b_1e_2 + c_1e_3 \\ T(e_2) &= T(0,1,0) = (a_2, b_2, c_2) = a_2e_1 + b_2e_2 + c_2e_3 \\ T(e_3) &= T(0,0,1) = (a_3, b_3, c_3) = a_3e_1 + b_3e_2 + c_3e_3 \end{aligned}$$

Accordingly,

$$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

**Remark:** This property holds for any space $K^n$ but only relative to the usual basis

$$\{e_1 = (1,0,\ldots,0),\ e_2 = (0,1,0,\ldots,0),\ \ldots,\ e_n = (0,\ldots,0,1)\}$$

**7.4.** Find the matrix representation of each of the following linear operators $T$ on $\mathbf{R}^3$ relative to the usual basis $\{e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)\}$:

(i) $T(x, y, z) = (2x - 3y + 4z,\ 5x - y - 2z,\ 4x + 7y)$.

(ii) $T(x, y, z) = (2y + z,\ x - 4y,\ 3x)$.

By Problem 7.3. (i) $[T]_e = \begin{pmatrix} 2 & -3 & 4 \\ 5 & -1 & 2 \\ 4 & 7 & 0 \end{pmatrix}$, (ii) $[T]_e = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$

91 M 7

**7.5.** Let $T$ be the linear operator on $\mathbf{R}^3$ defined by $T(x, y, z) = (2y + z,\ x - 4y,\ 3x)$.

(i) Find the matrix of $T$ in the basis $\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\}$

(ii) Verify that $[T]_f[v]_f = [T(v)]_f$ for any vector $v \in \mathbf{R}^3$.

We must first find the coordinates of an arbitrary vector $(a, b, c) \in \mathbf{R}^3$ with respect to the basis $\{f_1, f_2, f_3\}$. Write $(a, b, c)$ as a linear combination of the $f_i$ using unknown scalars $x$, $y$ and $z$:

$$\begin{aligned}(a, b, c) &= x(1,1,1) + y(1,1,0) + z(1,0,0) \\ &= (x + y + z,\ x + y,\ x)\end{aligned}$$

Set corresponding components equal to each other to obtain the system of equations

$$x + y + z = a, \qquad x + y = b, \qquad x = c$$

Solve the system for $x$, $y$ and $z$ in terms of $a$, $b$ and $c$ to find $x = c$, $y = b - c$, $z = a - b$. Thus

$$(a, b, c) = cf_1 + (b - c)f_2 + (a - b)f_3$$

(i) Since $T(x, y, z) = (2y + z,\ x - 4y,\ 3x)$

$$\begin{aligned} T(f_1) &= T(1,1,1) = (3,-3,3) = 3f_1 - 6f_2 + 6f_3 \\ T(f_2) &= T(1,1,0) = (2,-3,3) = 3f_1 - 6f_2 + 5f_3 \\ T(f_3) &= T(1,0,0) = (0,1,3) = 3f_1 - 2f_2 - f_3 \end{aligned} \quad \text{and} \quad [T]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$$

(ii) Suppose $v = (a, b, c)$; then

$$v = (a, b, c) = cf_1 + (b - c)f_2 + (a - b)f_3 \quad \text{and so} \quad [v]_f = \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix}$$

Also,

$$\begin{aligned} T(v) &= T(a, b, c) = (2b + c,\ a - 4b,\ 3a) \\ &= 3af_1 + (-2a - 4b)f_2 + (-a + 6b + c)f_3 \end{aligned} \quad \text{and so} \quad [T(v)]_f = \begin{pmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{pmatrix}$$

Thus

$$[T]_f[v]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}\begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix} = \begin{pmatrix} 3a \\ -2a - 4b \\ -a + 6b + c \end{pmatrix} = [T(v)]_f$$

**7.6.** Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and let $T$ be the linear operator on $\mathbf{R}^2$ defined by $T(v) = Av$ (where $v$ is written as a column vector). Find the matrix of $T$ in each of the following bases:

(i) $\{e_1 = (1, 0),\ e_2 = (0, 1)\}$, i.e. the usual basis;

(ii) $\{f_1 = (1, 3),\ f_2 = (2, 5)\}$.

(i)
$$\begin{aligned} T(e_1) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 1e_1 + 3e_2 \\ T(e_2) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2e_1 + 4e_2 \end{aligned} \quad \text{and thus} \quad [T]_e = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Observe that the matrix of $T$ in the usual basis is precisely the original matrix $A$ which defined $T$. This is not unusual. In fact, we show in the next problem that this is true for any matrix $A$ when using the usual basis.

(ii) By Problem 7.2, $(a, b) = (2b - 5a)f_1 + (3a - b)f_2$. Hence

$$T(f_1) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 15 \end{pmatrix} = -5f_1 + 6f_2 \qquad \text{and thus} \qquad [T]_f = \begin{pmatrix} -5 & -8 \\ 6 & 10 \end{pmatrix}$$

$$T(f_2) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 12 \\ 26 \end{pmatrix} = -8f_1 + 10f_2$$

90 **7.7.** Recall that any $n$-square matrix $A = (a_{ij})$ may be viewed as the linear operator $T$ on $K^n$ defined by $T(v) = Av$, where $v$ is written as a column vector. Show that the matrix representation of $T$ relative to the usual basis $\{e_i\}$ of $K^n$ is the matrix $A$, that is, $[T]_e = A$.

$$T(e_1) = Ae_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{pmatrix} = a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n$$

$$T(e_2) = Ae_2 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{n2} \end{pmatrix} = a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n$$

..................................................................................

$$T(e_n) = Ae_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{nn} \end{pmatrix} = a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n$$

(That is, $T(e_i) = Ae_i$ is the $i$th column of $A$.) Accordingly,

$$[T]_e = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = A$$

**7.8.** Each of the sets (i) $\{1, t, e^t, te^t\}$ and (ii) $\{e^{3t}, te^{3t}, t^2e^{3t}\}$ is a basis of a vector space $V$ of functions $f: \mathbf{R} \to \mathbf{R}$. Let $D$ be the differential operator on $V$, that is, $D(f) = df/dt$. Find the matrix of $D$ in the given basis.

(i)
$$\begin{aligned} D(1) &= 0 &&= 0(1) + 0(t) + 0(e^t) + 0(te^t) \\ D(t) &= 1 &&= 1(1) + 0(t) + 0(e^t) + 0(te^t) \\ D(e^t) &= e^t &&= 0(1) + 0(t) + 1(e^t) + 0(te^t) \\ D(te^t) &= e^t + te^t &&= 0(1) + 0(t) + 1(e^t) + 1(te^t) \end{aligned} \qquad \text{and} \qquad [D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(ii)
$$\begin{aligned} D(e^{3t}) &= 3e^{3t} &&= 3(e^{3t}) + 0(te^{3t}) + 0(t^2e^{3t}) \\ D(te^{3t}) &= e^{3t} + 3te^{3t} &&= 1(e^{3t}) + 3(te^{3t}) + 0(t^2e^{3t}) \\ D(t^2e^{3t}) &= 2te^{3t} + 3t^2e^{3t} &&= 0(e^{3t}) + 2(te^{3t}) + 3(t^2e^{3t}) \end{aligned} \qquad \text{and} \qquad [D] = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

**7.9.** Prove Theorem 7.1: Suppose $\{e_1, \ldots, e_n\}$ is a basis of $V$ and $T$ is a linear operator on $V$. Then for any $v \in V$, $[T]_e[v]_e = [T(v)]_e$. 3

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

$$T(e_i) = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n = \sum_{j=1}^{n} a_{ij}e_j$$

Then $[T]_e$ is the $n$-square matrix whose $j$th row is

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Now suppose
$$v = k_1e_1 + k_2e_2 + \cdots + k_ne_n = \sum_{i=1}^{n} k_ie_i$$

Writing a column vector as the transpose of a row vector,

$$[v]_e = (k_1, k_2, \ldots, k_n)^t \tag{2}$$

Furthermore, using the linearity of $T$,

$$T(v) = T\left(\sum_{i=1}^{n} k_ie_i\right) = \sum_{i=1}^{n} k_iT(e_i) = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}e_j\right)$$
$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)e_j = \sum_{j=1}^{n}(a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)e_j$$

Thus $[T(v)]_e$ is the column vector whose $j$th entry is

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \tag{3}$$

On the other hand, the $j$th entry of $[T]_e[v]_e$ is obtained by multiplying the $j$th row of $[T]_e$ by $[v]_e$, i.e. (1) by (2). But the product of (1) and (2) is (3); hence $[T]_e[v]_e$ and $[T(v)]_e$ have the same entries. Thus $[T]_e[v]_e = [T(v)]_e$.

**7.10.** Prove Theorem 7.2: Let $\{e_1, \ldots, e_n\}$ be a basis of $V$ over $K$, and let $\mathcal{A}$ be the algebra of $n$-square matrices over $K$. Then the mapping $T \mapsto [T]_e$ is a vector space isomorphism from $A(V)$ onto $\mathcal{A}$. That is, the mapping is one-one and onto and, for any $S, T \in A(V)$ and any $k \in K$, $[T+S]_e = [T]_e + [S]_e$ and $[kT]_e = k[T]_e$.

The mapping is one-one since, by Theorem 8.1, a linear mapping is completely determined by its values on a basis. The mapping is onto since each matrix $M \in \mathcal{A}$ is the image of the linear operator

$$F(e_i) = \sum_{j=1}^{n} m_{ij}e_j \qquad i = 1, \ldots, n$$

where $(m_{ij})$ is the transpose of the matrix $M$.

Now suppose, for $i = 1, \ldots, n$,

$$T(e_i) = \sum_{j=1}^{n} a_{ij}e_j \quad \text{and} \quad S(e_i) = \sum_{j=1}^{n} b_{ij}e_j$$

Let $A$ and $B$ be the matrices $A = (a_{ij})$ and $B = (b_{ij})$. Then $[T]_e = A^t$ and $[S]_e = B^t$. We have, for $i = 1, \ldots, n$,

$$(T+S)(e_i) = T(e_i) + S(e_i) = \sum_{j=1}^{n}(a_{ij} + b_{ij})e_j$$

Observe that $A + B$ is the matrix $(a_{ij} + b_{ij})$. Accordingly,

$$[T+S]_e = (A+B)^t = A^t + B^t = [T]_e + [S]_e$$

We also have, for $i = 1, \ldots, n$,

$$(kT)(e_i) = k\,T(e_i) = k\sum_{j=1}^{n} a_{ij}e_j = \sum_{j=1}^{n}(ka_{ij})e_j$$

Observe that $kA$ is the matrix $(ka_{ij})$. Accordingly,

$$[kT]_e = (kA)^t = kA^t = k[T]_e$$

Thus the theorem is proved.

93

**7.11.** Prove Theorem 7.3: Let $\{e_1, \ldots, e_n\}$ be a basis of $V$. Then for any linear operators $S, T \in A(V)$, $[ST]_e = [S]_e [T]_e$.

Suppose $T(e_i) = \sum_{j=1}^{n} a_{ij} e_j$ and $S(e_j) = \sum_{k=1}^{n} b_{jk} e_k$. Let $A$ and $B$ be the matrices $A = (a_{ij})$ and $B = (b_{jk})$. Then $[T]_e = A^t$ and $[S]_e = B^t$. We have

$$(ST)(e_i) = S(T(e_i)) = S\left(\sum_{j=1}^{n} a_{ij} e_j\right) = \sum_{j=1}^{n} a_{ij} S(e_j)$$

$$= \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk} e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij} b_{jk}\right) e_k$$

Recall that $AB$ is the matrix $AB = (c_{ik})$ where $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$. Accordingly,

$$[ST]_e = (AB)^t = B^t A^t = [S]_e [T]_e$$

## CHANGE OF BASIS, SIMILAR MATRICES

M 90 (2)

**7.12.** Consider these bases of $\mathbb{R}^2$: $\{e_1 = (1,0),\ e_2 = (0,1)\}$ and $\{f_1 = (1,3),\ f_2 = (2,5)\}$. (i) Find the transition matrix $P$ from $\{e_i\}$ to $\{f_i\}$. (ii) Find the transition matrix $Q$ from $\{f_i\}$ to $\{e_i\}$. (iii) Verify that $Q = P^{-1}$. (iv) Show that $[v]_f = P^{-1}[v]_e$ for any vector $v \in \mathbb{R}^2$. (v) Show that $[T]_f = P^{-1}[T]_e P$ for the operator $T$ on $\mathbb{R}^2$ defined by $T(x, y) = (2y, 3x - y)$. (See Problems 7.1 and 7.2.)

(i)
$$f_1 = (1,3) = 1e_1 + 3e_2 \qquad\qquad f_2 = (2,5) = 2e_1 + 5e_2 \qquad \text{and} \qquad P = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

(ii) By Problem 7.2, $(a, b) = (2b - 5a)f_1 + (3a - b)f_2$. Thus

$$e_1 = (1,0) = -5f_1 + 3f_2 \qquad\qquad e_2 = (0,1) = 2f_1 - f_2 \qquad \text{and} \qquad Q = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}$$

(iii)
$$PQ = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

(iv) If $v = (a, b)$, then $[v]_e = \begin{pmatrix} a \\ b \end{pmatrix}$ and $[v]_f = \begin{pmatrix} 2b - 5a \\ 3a - b \end{pmatrix}$. Hence

$$P^{-1}[v]_e = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -5a + 2b \\ 3a - b \end{pmatrix} = [v]_f$$

(v) By Problems 7.1 and 7.2; $[T]_e = \begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}$ and $[T]_f = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix}$. Hence

$$P^{-1}[T]_e P = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} -30 & -48 \\ 18 & 29 \end{pmatrix} = [T]_f$$

M 93 (7)

**7.13.** Consider the following bases of $\mathbb{R}^3$: $\{e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)\}$ and $\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\}$. (i) Find the transition matrix $P$ from $\{e_i\}$ to $\{f_i\}$. (ii) Find the transition matrix $Q$ from $\{f_i\}$ to $\{e_i\}$. (iii) Verify that $Q = P^{-1}$. (iv) Show that $[v]_f = P^{-1}[v]_e$ for any vector $v \in \mathbb{R}^3$. (v) Show that $[T]_f = P^{-1}[T]_e P$ for the $T$ defined by $T(x, y, z) = (2y + z,\ x - 4y,\ 3x)$. (See Problems 7.4 and 7.5.)

(i)
$$f_1 = (1,1,1) = 1e_1 + 1e_2 + 1e_3 \qquad f_2 = (1,1,0) = 1e_1 + 1e_2 + 0e_3 \qquad f_3 = (1,0,0) = 1e_1 + 0e_2 + 0e_3 \qquad \text{and} \qquad P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(ii) By Problem 7.5, $(a,b,c) = cf_1 + (b-c)f_2 + (a-b)f_3$. Thus

$$\begin{aligned} e_1 &= (1,0,0) = 0f_1 + 0f_2 + 1f_3 \\ e_2 &= (0,1,0) = 0f_1 + 1f_2 - 1f_3 \\ e_3 &= (0,0,1) = 1f_1 - 1f_2 + 0f_3 \end{aligned} \qquad \text{and} \qquad Q = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

(iii)
$$PQ = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$

(iv) If $v = (a,b,c)$, then $[v]_e = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and $[v]_f = \begin{pmatrix} c \\ b-c \\ a-b \end{pmatrix}$. Thus

$$P^{-1}[v]_e = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} c \\ b-c \\ a-b \end{pmatrix} = [v]_f$$

(v) By Problems 7.4(ii) and 7.5, $[T]_e = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$ and $[T]_f = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix}$ Thus

$$P^{-1}[T]_eP = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 \\ 1 & -4 & 0 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{pmatrix} = [T]_f$$

**7.14. Prove Theorem 7.4: Let $P$ be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space $V$. Then for any $v \in V$, $P[v]_f = [v]_e$. Also, $[v]_f = P^{-1}[v]_e$.**

Suppose, for $i = 1, \ldots, n$, $f_i = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n = \sum_{j=1}^{n} a_{ij}e_j$. Then $P$ is the $n$-square matrix whose $j$th row is

$$(a_{1j}, a_{2j}, \ldots, a_{nj}) \tag{1}$$

Also suppose $v = k_1f_1 + k_2f_2 + \cdots + k_nf_n = \sum_{i=1}^{n} k_if_i$. Then writing a column vector as the transpose of a row vector,

$$[v]_f = (k_1, k_2, \ldots, k_n)^t \tag{2}$$

Substituting for $f_i$ in the equation for $v$,

$$\begin{aligned} v &= \sum_{i=1}^{n} k_if_i = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}e_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}k_i\right)e_j \\ &= \sum_{j=1}^{n} (a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)e_j \end{aligned}$$

Accordingly, $[v]_e$ is the column vector whose $j$th entry is

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \tag{3}$$

On the other hand, the $j$th entry of $P[v]_f$ is obtained by multiplying the $j$th row of $P$ by $[v]_f$, i.e. (1) by (2). But the product of (1) and (2) is (3); hence $P[v]_f$ and $[v]_e$ have the same entries and thus $P[v]_f = [v]_e$.

Furthermore, multiplying the above by $P^{-1}$ gives $P^{-1}[v]_e = P^{-1}P[v]_f = [v]_f$.

**7.15. Prove Theorem 7.5: Let $P$ be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space $V$. Then, for any linear operator $T$ on $V$, $[T]_f = P^{-1}[T]_eP$.**

For any vector $v \in V$, $P^{-1}[T]_eP[v]_f = P^{-1}[T]_e[v]_e = P^{-1}[T(v)]_e = [T(v)]_f$.

But $[T]_f[v]_f = [T(v)]_f$; hence $P^{-1}[T]_e P[v]_f = [T]_f[v]_f$.

Since the mapping $v \mapsto [v]_f$ is onto $K^n$, $P^{-1}[T]_e PX = [T]_f X$ for every $X \in K^n$.

Accordingly, $P^{-1}[T]_e P = [T]_f$.

**7.16.** Show that similarity of matrices is an equivalence relation, that is: (i) $A$ is similar to $A$; (ii) if $A$ is similar to $B$, then $B$ is similar to $A$; (iii) if $A$ is similar to $B$ and $B$ is similar to $C$ then $A$ is similar to $C$.

(i) The identity matrix $I$ is invertible and $I = I^{-1}$. Since $A = I^{-1}AI$, $A$ is similar to $A$.

(ii) Since $A$ is similar to $B$ there exists an invertible matrix $P$ such that $A = P^{-1}BP$. Hence $B = PAP^{-1} = (P^{-1})^{-1}AP^{-1}$ and $P^{-1}$ is invertible. Thus $B$ is similar to $A$.

(iii) Since $A$ is similar to $B$ there exists an invertible matrix $P$ such that $A = P^{-1}BP$, and since $B$ is similar to $C$ there exists an invertible matrix $Q$ such that $B = Q^{-1}CQ$. Hence $A = P^{-1}BP = P^{-1}(Q^{-1}CQ)P = (QP)^{-1}C(QP)$ and $QP$ is invertible. Thus $A$ is similar to $C$.

## TRACE

**7.17.** The *trace* of a square matrix $A = (a_{ij})$, written $\operatorname{tr}(A)$, is the sum of its diagonal elements: $\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$. Show that (i) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, (ii) if $A$ is similar to $B$ then $\operatorname{tr}(A) = \operatorname{tr}(B)$.

(i) Suppose $A = (a_{ij})$ and $B = (b_{ij})$. Then $AB = (c_{ik})$ where $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$. Thus

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$$

On the other hand, $BA = (d_{jk})$ where $d_{jk} = \sum_{i=1}^{n} b_{ji}a_{ik}$. Thus

$$\operatorname{tr}(BA) = \sum_{j=1}^{n} d_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \operatorname{tr}(AB)$$

(ii) If $A$ is similar to $B$, there exists an invertible matrix $P$ such that $A = P^{-1}BP$. Using (i),

$$\operatorname{tr}(A) = \operatorname{tr}(P^{-1}BP) = \operatorname{tr}(BPP^{-1}) = \operatorname{tr}(B)$$

**7.18.** Find the trace of the following operator on $\mathbf{R}^3$:

$$T(x, y, z) = (a_1x + a_2y + a_3z,\ b_1x + b_2y + b_3z,\ c_1x + c_2y + c_3z)$$

We first must find a matrix representation of $T$. Choosing the usual basis $\{e_i\}$,

$$[T]_e = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

and $\operatorname{tr}(T) = \operatorname{tr}([T]_e) = a_1 + b_2 + c_3$.

**7.19.** Let $V$ be the space of $2 \times 2$ matrices over $\mathbf{R}$, and let $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Let $T$ be the linear operator on $V$ defined by $T(A) = MA$. Find the trace of $T$.

We must first find a matrix representation of $T$. Choose the usual basis of $V$:

$$\left\{E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

Then

$$T(E_1) = ME_1 = \begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix} = \begin{pmatrix}1&0\\3&0\end{pmatrix} = 1E_1 + 0E_2 + 3E_3 + 0E_4$$

$$T(E_2) = ME_2 = \begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}0&1\\0&0\end{pmatrix} = \begin{pmatrix}0&1\\0&3\end{pmatrix} = 0E_1 + 1E_2 + 0E_3 + 3E_4$$

$$T(E_3) = ME_3 = \begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}0&0\\1&0\end{pmatrix} = \begin{pmatrix}2&0\\4&0\end{pmatrix} = 2E_1 + 0E_2 + 4E_3 + 0E_4$$

$$T(E_4) = ME_4 = \begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}0&2\\0&4\end{pmatrix} = 0E_1 + 2E_2 + 0E_3 + 4E_4$$

Hence

$$[T]_E = \begin{pmatrix}1&0&2&0\\0&1&0&2\\3&0&4&0\\0&3&0&4\end{pmatrix}$$

and $\operatorname{tr}(T) = 1 + 1 + 4 + 4 = 10$.

## MATRIX REPRESENTATIONS OF LINEAR MAPPINGS

**7.20.** Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ be the linear mapping defined by $F(x, y, z) = (3x + 2y - 4z,\ x - 5y + 3z)$.

(i) Find the matrix of $F$ in the following bases of $\mathbf{R}^3$ and $\mathbf{R}^2$:

$$\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\}, \quad \{g_1 = (1,3),\ g_2 = (2,5)\}$$

(ii) Verify that the action of $F$ is preserved by its matrix representation; that is, for any $v \in \mathbf{R}^3$, $[F]_f^g [v]_f = [F(v)]_g$.

(i) By Problem 7.2, $(a, b) = (2b - 5a)g_1 + (3a - b)g_2$. Hence

$$\begin{aligned} F(f_1) &= F(1,1,1) = (1,-1) = -7g_1 + 4g_2 \\ F(f_2) &= F(1,1,0) = (5,-4) = -33g_1 + 19g_2 \\ F(f_3) &= F(1,0,0) = (3,1) = -13g_1 + 8g_2 \end{aligned} \quad \text{and} \quad [F]_f^g = \begin{pmatrix}-7&-33&-13\\4&19&8\end{pmatrix}$$

(ii) If $v = (x, y, z)$ then, by Problem 7.5, $v = zf_1 + (y - z)f_2 + (x - y)f_3$. Also,

$$F(v) = (3x + 2y - 4z,\ x - 5y + 3z) = (-13x - 20y + 26z)g_1 + (8x + 11y - 15z)g_2$$

Hence $[v]_f = \begin{pmatrix}z\\y-z\\x-y\end{pmatrix}$ and $[F(v)]_g = \begin{pmatrix}-13x - 20y + 26z\\8x + 11y - 15z\end{pmatrix}$. Thus

$$[F]_f^g [v]_f = \begin{pmatrix}-7&-33&-13\\4&19&8\end{pmatrix}\begin{pmatrix}z\\y-z\\x-y\end{pmatrix} = \begin{pmatrix}-13x - 20y + 26z\\8x + 11y - 15z\end{pmatrix} = [F(v)]_g$$

**7.21.** Let $F: K^n \to K^m$ be the linear mapping defined by

$$F(x_1, x_2, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n,\ a_{21}x_1 + \cdots + a_{2n}x_n,\ \ldots,\ a_{m1}x_1 + \cdots + a_{mn}x_n)$$

Show that the matrix representation of $F$ relative to the usual bases of $K^n$ and of $K^m$ is given by

$$[F] = \begin{pmatrix}a_{11}&a_{12}&\ldots&a_{1n}\\a_{21}&a_{22}&\ldots&a_{2n}\\ \ldots&\ldots&\ldots&\ldots\\a_{m1}&a_{m2}&\ldots&a_{mn}\end{pmatrix}$$

That is, the rows of $[F]$ are obtained from the coefficients of the $x_i$ in the components of $F(x_1, \ldots, x_n)$, respectively.

$$\begin{aligned} F(1,0,\ldots,0) &= (a_{11}, a_{21}, \ldots, a_{m1}) \\ F(0,1,\ldots,0) &= (a_{12}, a_{22}, \ldots, a_{m2}) \\ &\cdots\cdots\cdots \\ F(0,0,\ldots,1) &= (a_{1n}, a_{2n}, \ldots, a_{mn}) \end{aligned} \quad \text{and} \quad [F] = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

**7.22.** Find the matrix representation of each of the following linear mappings relative to the usual bases of $\mathbf{R}^n$:

(i) $F: \mathbf{R}^2 \to \mathbf{R}^3$ defined by $F(x, y) = (3x - y, 2x + 4y, 5x - 6y)$

(ii) $F: \mathbf{R}^4 \to \mathbf{R}^2$ defined by $F(x, y, s, t) = (3x - 4y + 2s - 5t, 5x + 7y - s - 2t)$

(iii) $F: \mathbf{R}^3 \to \mathbf{R}^4$ defined by $F(x, y, z) = (2x + 3y - 8z, x + y + z, 4x - 5z, 6y)$

By Problem 7.21, we need only look at the coefficients of the unknowns in $F(x, y, \ldots)$. Thus

$$\text{(i)}\ [F] = \begin{pmatrix} 3 & -1 \\ 2 & 4 \\ 5 & -6 \end{pmatrix} \qquad \text{(ii)}\ [F] = \begin{pmatrix} 3 & -4 & 2 & -5 \\ 5 & 7 & -1 & -2 \end{pmatrix} \qquad \text{(iii)}\ [F] = \begin{pmatrix} 2 & 3 & -8 \\ 1 & 1 & 1 \\ 4 & 0 & -5 \\ 0 & 6 & 0 \end{pmatrix}$$

**7.23.** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $T(x, y) = (2x - 3y, x + 4y)$. Find the matrix of $T$ in the bases $\{e_1 = (1, 0), e_2 = (0, 1)\}$ and $\{f_1 = (1, 3), f_2 = (2, 5)\}$ of $\mathbf{R}^2$ respectively. (We can view $T$ as a linear mapping from one space into another, each having its own basis.)

By Problem 7.2, $(a, b) = (2b - 5a)f_1 + (3a - b)f_2$. Then

$$\begin{aligned} T(e_1) &= T(1,0) = (2,1) &&= -8f_1 + 5f_2 \\ T(e_2) &= T(0,1) = (-3,4) &&= 23f_1 - 13f_2 \end{aligned} \quad \text{and} \quad [T]_e^f = \begin{pmatrix} -8 & 23 \\ 5 & -13 \end{pmatrix}$$

**7.24.** Let $A = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}$. Recall that $A$ determines a linear mapping $F: \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(v) = Av$ where $v$ is written as a column vector.

(i) Show that the matrix representation of $F$ relative to the usual basis of $\mathbf{R}^3$ and of $\mathbf{R}^2$ is the matrix $A$ itself: $[F] = A$.

(ii) Find the matrix representation of $F$ relative to the following bases of $\mathbf{R}^3$ and $\mathbf{R}^2$.

$$\{f_1 = (1, 1, 1),\ f_2 = (1, 1, 0),\ f_3 = (1, 0, 0)\}, \quad \{g_1 = (1, 3),\ g_2 = (2, 5)\}$$

(i)

$$F(1,0,0) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2e_1 + 1e_2$$

$$F(0,1,0) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix} = 5e_1 - 4e_2$$

$$F(0,0,1) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 7 \end{pmatrix} = -3e_1 + 7e_2$$

from which $[F] = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix} = A$. (Compare with Problem 7.7.)

(ii) By Problem 7.2, $(a, b) = (2b - 5a)g_1 + (3a - b)g_2$. Then

$$F(f_1) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = -12g_1 + 8g_2$$

$$F(f_2) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \end{pmatrix} = -41g_1 + 24g_2$$

$$F(f_3) = \begin{pmatrix} 2 & 5 & -3 \\ 1 & -4 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -8g_1 + 5g_2$$

and $[F]^g_f = \begin{pmatrix} -12 & -41 & -8 \\ 8 & 24 & 5 \end{pmatrix}$.

**7.25.** Prove Theorem 7.12: Let $F: V \to U$ be linear. Then there exists a basis of $V$ and a basis of $U$ such that the matrix representation $A$ of $F$ has the form $A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ where $I$ is the $r$-square identity matrix and $r$ is the rank of $F$.

Suppose $\dim V = m$ and $\dim U = n$. Let $W$ be the kernel of $F$ and $U'$ the image of $F$. We are given that $\operatorname{rank} F = r$; hence the dimension of the kernel of $F$ is $m - r$. Let $\{w_1, \ldots, w_{m-r}\}$ be a basis of the kernel of $F$ and extend this to a basis of $V$:

$$\{v_1, \ldots, v_r, w_1, \ldots, w_{m-r}\}$$

Set
$$u_1 = F(v_1),\ u_2 = F(v_2),\ \ldots,\ u_r = F(v_r)$$

We note that $\{u_1, \ldots, u_r\}$ is a basis of $U'$, the image of $F$. Extend this to a basis

$$\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$$

of $U$. Observe that

$$\begin{aligned}
F(v_1) &= u_1 = 1u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
F(v_2) &= u_2 = 0u_1 + 1u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
F(v_r) &= u_r = 0u_1 + 0u_2 + \cdots + 1u_r + 0u_{r+1} + \cdots + 0u_n \\
F(w_1) &= 0 = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
F(w_{m-r}) &= 0 = 0u_1 + 0u_2 + \cdots + 0u_r + 0u_{r+1} + \cdots + 0u_n
\end{aligned}$$

Thus the matrix of $F$ in the above bases has the required form.

## Supplementary Problems

### MATRIX REPRESENTATIONS OF LINEAR OPERATORS

**7.26.** Find the matrix of each of the following linear operators $T$ on $\mathbf{R}^2$ with respect to the usual basis $\{e_1 = (1,0),\ e_2 = (0,1)\}$: (i) $T(x,y) = (2x - 3y,\ x + y)$, (ii) $T(x,y) = (5x + y,\ 3x - 2y)$.

**7.27.** Find the matrix of each operator $T$ in the preceding problem with respect to the basis $\{f_1 =$ $f_2 = (2,3)\}$. In each case, verify that $[T]_f[v]_f = [T(v)]_f$ for any $v \in \mathbf{R}^2$.

**7.28.** Find the matrix of each operator $T$ in Problem 7.26 in the basis $\{g_1 = (1,3),\ g_2 = (1,4)\}$.

. Find the matrix representation of each of the following linear operators $T$ on $\mathbf{R}^3$ relative to the usual basis:

(i) $T(x, y, z) = (x, y, 0)$

(ii) $T(x, y, z) = (2x - 7y - 4z, 3x + y + 4z, 6x - 8y + z)$

(iii) $T(x, y, z) = (z, y + z, x + y + z)$

7.30. Let $D$ be the differential operator, i.e. $D(f) = df/dt$. Each of the following sets is a basis of a vector space $V$ of functions $f: \mathbf{R} \to \mathbf{R}$. Find the matrix of $D$ in each basis: (i) $\{e^t, e^{2t}, te^{2t}\}$, (ii) $\{\sin t, \cos t\}$, (iii) $\{e^{3t}, te^{3t}, t^2e^{3t}\}$, (iv) $\{1, t, \sin 3t, \cos 3t\}$.

7.31. Consider the complex field $\mathbf{C}$ as a vector space over the real field $\mathbf{R}$. Let $T$ be the conjugation operator on $\mathbf{C}$, i.e. $T(z) = \bar{z}$. Find the matrix of $T$ in each basis: (i) $\{1, i\}$, (ii) $\{1 + i, 1 + 2i\}$.

7.32. Let $V$ be the vector space of $2 \times 2$ matrices over $\mathbf{R}$ and let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Find the matrix of each of the following linear operators $T$ on $V$ in the usual basis (see Problem 7.19) of $V$: (i) $T(A) = MA$, (ii) $T(A) = AM$, (iii) $T(A) = MA - AM$.

M

7.33. Let $1_V$ and $0_V$ denote the identity and zero operators, respectively, on a vector space $V$. Show that, for any basis $\{e_i\}$ of $V$, (i) $[1_V]_e = I$, the identity matrix, (ii) $[0_V]_e = 0$, the zero matrix.

## :GE OF BA IS, SIMILAR MATRICES

Consider the following bases of $\mathbf{R}^2$: $\{e_1 = (1, 0), e_2 = (0, 1)\}$ and $\{f_1 = (1, 2), f_2 = (2, 3)\}$.

(i) Find the transition matrices $P$ and $Q$ from $\{e_i\}$ to $\{f_i\}$ and from $\{f_i\}$ to $\{e_i\}$, respectively. Verify $Q = P^{-1}$.

(ii) Show that $[v]_e = P[v]_f$ for any vector $v \in \mathbf{R}^2$.

(iii) Show that $[T]_f = P^{-1}[T]_e P$ for each operator $T$ in Problem 7.26.

7.35. Repeat Problem 7.34 for the bases $\{f_1 = (1, 2), f_2 = (2, 3)\}$ and $\{g_1 = (1, 3), g_2 = (1, 4)\}$.

36. Suppose $\{e_1, e_2\}$ is a basis of $V$ and $T: V \to V$ is the linear operator for which $T(e_1) = 3e_1 - 2e_2$ and $T(e_2) = e_1 + 4e_2$. Suppose $\{f_1, f_2\}$ is the basis of $V$ for which $f_1 = e_1 + e_2$ and $f_2 = 2e_1 + 3e_2$. Find the matrix of $T$ in the basis $\{f_1, f_2\}$.

Consider the bases $B = \{1, i\}$ and $B' = \{1 + i, 1 + 2i\}$ of the complex field $\mathbf{C}$ over the real field $\mathbf{R}$. (i) Find the transition matrices $P$ and $Q$ from $B$ to $B'$ and from $B'$ to $B$, respectively. Verify that $Q = P^{-1}$. (ii) Show that $[T]_{B'} = P^{-1}[T]_B P$ for the conjugation operator $T$ in Problem 7.31.

Suppose $\{e_i\}$, $\{f_i\}$ and $\{g_i\}$ are bases of $V$, and that $P$ and $Q$ are the transition matrices from $\{e_i\}$ to $\{f_i\}$ and from $\{f_i\}$ to $\{g_i\}$, respectively. Show that $PQ$ is the transition matrix from $\{e_i\}$ to $\{g_i\}$.

Let $A$ be a 2 by 2 matrix such that only $A$ is similar to itself. Show that $A$ has the form

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

Generalize to $n \times n$ matrices.

M 9?

Show tha' ill the matrices similar to an invertible matrix are invertible. More generally, show that lar matrices have the same rank.

## basEPRESENTATIONS OF LINEAR MAPPINGS

the matrix representation of the linear mappings relative to the usual bases for $\mathbf{R}^n$:

(1, 2 $F: \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (2x - 4y + 9z, 5x + 3y - 2z)$

$F: \mathbf{R}^2 \to \mathbf{R}^4$ defined by $F(x, y) = (3x + 4y, 5x - 2y, x + 7y, 4x)$

$F: \mathbf{R}^4 \to \mathbf{R}$ defined by $F(x, y, s, t) = 2x + 3y - 7s - t$

$F: \mathbf{R} \to \mathbf{R}^2$ defined by $F(x) = (3x, 5x)$

7.42. Let $F: \mathbf{R}^3 \to \mathbf{R}^2$ be the linear mapping defined by $F(x, y, z) = (2x + y - z, 3x - 2y + 4z)$.

(i) Find the matrix of $F$ in the following bases of $\mathbf{R}^3$ and $\mathbf{R}^2$.

$$\{f_1 = (1,1,1),\ f_2 = (1,1,0),\ f_3 = (1,0,0)\} \quad \text{and} \quad \{g_1 = (1,3),\ g_2 = (1,4)\}$$

(ii) Verify that, for any vector $v \in \mathbf{R}^3$, $[F]^g_f [v]_f = [F(v)]_g$.

7.43. Let $\{e_i\}$ and $\{f_i\}$ be bases of $V$, and let $1_V$ be the identity mapping on $V$. Show that the matrix of $1_V$ in the bases $\{e_i\}$ and $\{f_i\}$ is the inverse of the transition matrix $P$ from $\{e_i\}$ to $\{f_i\}$; that is, $[1_V]^f_e = P^{-1}$.

7.44. Prove Theorem 7.7, page 155. (*Hint.* See Problem 7.9, page 161.)

7.45. Prove Theorem 7.8. (*Hint.* See Problem 7.10.)

7.46. Prove Theorem 7.9. (*Hint.* See Problem 7.11.)

7.47. Prove Theorem 7.10. (*Hint.* See Problem 7.15.)

## MISCELLANEOUS PROBLEMS

7.48. Let $T$ be a linear operator on $V$ and let $W$ be a subspace of $V$ *invariant* under $T$, that is, $T(W) \subset W$. Suppose $\dim W = m$. Show that $T$ has a matrix representation of the form $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ where $A$ is an $m \times m$ submatrix.

7.49. Let $V = U \oplus W$, and let $U$ and $W$ each be invariant under a linear operator $T: V \to V$. Suppose $\dim U = m$ and $\dim V = n$. Show that $T$ has a matrix representation of the form $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ where $A$ and $B$ are $m \times m$ and $n \times n$ submatrices, respectively.

7.50. Recall that two linear operators $F$ and $G$ on $V$ are said to be *similar* if there exists an invertible operator $T$ on $V$ such that $G = T^{-1}FT$.

(i) Show that linear operators $F$ and $G$ are similar if and only if, for any basis $\{e_i\}$ of $V$, the matrix representations $[F]_e$ and $[G]_e$ are similar matrices.

(ii) Show that if an operator $F$ is diagonalizable, then any similar operator $G$ is also diagonalizable.

7.51. Two $m \times n$ matrices $A$ and $B$ over $K$ are said to be *equivalent* if there exists an $m$-square invertible matrix $Q$ and an $n$-square invertible matrix $P$ such that $B = QAP$.

(i) Show that equivalence of matrices is an equivalence relation.

(ii) Show that $A$ and $B$ can be matrix representations of the same linear operator $F: V \to U$ if and only if $A$ and $B$ are equivalent.

(iii) Show that every matrix $A$ is equivalent to a matrix of the form $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ where $I$ is the $r$-square identity matrix and $r = \operatorname{rank} A$.

2. Two algebras $\mathbf{A}$ and $\mathbf{B}$ over a field $K$ are said to be *isomorphic* (as algebras) if there exists a bijective mapping $f: \mathbf{A} \to \mathbf{B}$ such that for $u, v \in A$ and $k \in K$, (i) $f(u+v) = f(u) + f(v)$, (ii) $f(ku) = kf(u)$, (iii) $f(uv) = f(u)f(v)$. (That is, $f$ preserves the three operations of an algebra: vector addition, scalar multiplication, and vector multiplication.) The mapping $f$ is then called an *isomorphism* of $\mathbf{A}$ onto $\mathbf{B}$. Show that the relation of algebra isomorphism is an equivalence relation.

Let $\mathcal{A}$ be the algebra of $n$-square matrices over $K$, and let $P$ be an invertible matrix in $\mathcal{A}$. Show that the map $A \mapsto P^{-1}AP$, where $A \in \mathcal{A}$, is an algebra isomorphism of $\mathcal{A}$ onto itself.

8.45. (i) $(t+2)(t-3)(t-4)$, (ii) $(t+2)^2(t-4)$, (iii) $(t+2)^2(t-4)$.

8.46. (i) 3, 4, −2, (ii) 4, −2, (iii) 4, −2.

8.47. (i) 131, (ii) −55.

8.48. (i) 135, (ii) −103, (iii) −31.

8.49. $\operatorname{adj} A = \begin{pmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 2 & -2 & 0 \end{pmatrix}$, $A^{-1} = (\operatorname{adj} A)/|A| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -1 & 1 & 0 \end{pmatrix}$.

8.50. $\operatorname{adj} A = \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix}$.

8.51. (i) $\begin{pmatrix} 16 & 29 & 26 & -2 \\ -30 & -38 & -16 & 29 \\ 8 & 51 & 13 & 1 \\ 13 & 1 & 28 & 18 \end{pmatrix}$ (ii) $\begin{pmatrix} 21 & 14 & 17 & 19 \\ -44 & 11 & 33 & 11 \\ 29 & 1 & 13 & 21 \\ 17 & 7 & 19 & -18 \end{pmatrix}$

8.52. $A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$

8.54. $\det(T) = x$

8.55. (i) 0, (ii) 6, (iii) 1.

8.58. (i) $x = 21/26$, $y = 29/26$, (ii) $x = -5/13$, $y = 1/13$.

8.59. (i) $x = 5$, $y = 1$, $z = 1$. (ii) Since $\Delta = 0$, the system cannot be solved by determinants.

8.61. $\operatorname{sgn} \sigma = 1$, $\operatorname{sgn} \tau = -1$, $\operatorname{sgn} \pi = -1$.

8.62. (i) $\tau \circ \sigma = 53142$, (ii) $\pi \circ \sigma = 52413$, (iii) $\sigma^{-1} = 32154$, (iv) $\tau^{-1} = 14253$.

8.66. (i) Yes, (ii) No, (iii) Yes, (iv) No.

# Chapter 9

## Eigenvalues and Eigenvectors

### INTRODUCTION

In this chapter we investigate the theory of a single linear operator $T$ on a vector space $V$ of finite dimension. In particular, we find conditions under which $T$ is diagonalizable. As was seen in Chapter 7, this question is closely related to the theory of similarity transformations for matrices.

We shall also associate certain polynomials with an operator $T$: its characteristic polynomial and its minimum polynomial. These polynomials and their roots play a major role in the investigation of $T$. We comment that the particular field $K$ also plays an important part in the theory since the existence of roots of a polynomial depends on $K$.

### POLYNOMIALS OF MATRICES AND LINEAR OPERATORS

Consider a polynomial $f(t)$ over a field $K$: $f(t) = a_nt^n + \cdots + a_1t + a_0$. If $A$ is a square matrix over $K$, then we define

$$f(A) = a_nA^n + \cdots + a_1A + a_0I$$

where $I$ is the identity matrix. In particular, we say that $A$ is a *root* or *zero* of the polynomial $f(t)$ if $f(A) = 0$.

**Example 9.1:** Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and let $f(t) = 2t^2 - 3t + 7$, $g(t) = t^2 - 5t - 2$. Then

$$f(A) = 2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 14 \\ 21 & 39 \end{pmatrix}$$

and

$$g(A) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus $A$ is a zero of $g(t)$.

The following theorem applies.

**Theorem 9.1:** Let $f$ and $g$ be polynomials over $K$, and let $A$ be an $n$-square matrix over $K$. Then

(i) $(f + g)(A) = f(A) + g(A)$

(ii) $(fg)(A) = f(A)\,g(A)$

and, for any scalar $k \in K$,

(iii) $(kf)(A) = k\,f(A)$

Furthermore, since $f(t)\,g(t) = g(t)\,f(t)$ for any polynomials $f(t)$ and $g(t)$,

$$f(A)\,g(A) = g(A)\,f(A)$$

That is, any two polynomials in the matrix $A$ commute.

Now suppose $T: V \to V$ is a linear operator on a vector space $V$ over $K$. If $f(t) = a_n t^n + \cdots + a_1 t + a_0$, then we define $f(T)$ in the same way as we did for matrices:

$$f(T) = a_n T^n + \cdots + a_1 T + a_0 I$$

where $I$ is now the identity mapping. We also say that $T$ is a *zero* or *root* of $f(t)$ if $f(T) = 0$. We remark that the relations in Theorem 9.1 hold for operators as they do for matrices; hence any two polynomials in $T$ commute.

Furthermore, if $A$ is a matrix representation of $T$, then $f(A)$ is the matrix representation of $f(T)$. In particular, $f(T) = 0$ if and only if $f(A) = 0$.

## EIGENVALUES AND EIGENVECTORS

90, 85, 55

Let $T: V \to V$ be a linear operator on a vector space $V$ over a field $K$. A scalar $\lambda \in K$ is called an *eigenvalue* of $T$ if there exists a nonzero vector $v \in V$ for which

$$T(v) = \lambda v$$

Every vector satisfying this relation is then called an *eigenvector* of $T$ belonging to the eigenvalue $\lambda$. Note that each scalar multiple $kv$ is such an eigenvector:

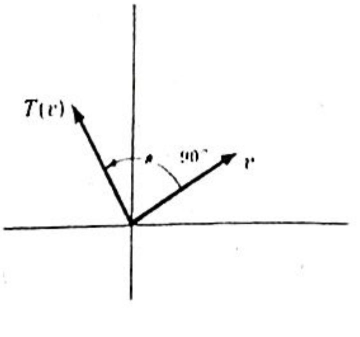
$$T(kv) = k\,T(v) = k(\lambda v) = \lambda(kv)$$

The set of all such vectors is a subspace of $V$ (Problem 9.6) called the *eigenspace* of $\lambda$.

The terms *characteristic value* and *characteristic vector* (or: *proper value* and *proper vector*) are frequently used instead of eigenvalue and eigenvector.

**Example 9.2:** Let $I: V \to V$ be the identity mapping. Then, for every $v \in V$, $I(v) = v = 1v$. Hence 1 is an eigenvalue of $I$, and every vector in $V$ is an eigenvector belonging to 1.

**Example 9.3:** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be the linear operator which rotates each vector $v \in \mathbf{R}^2$ by an angle $\theta = 90°$. Note that no nonzero vector is a multiple of itself. Hence $T$ has no eigenvalues and so no eigenvectors.




**Example 9.4:** Let $D$ be the differential operator on the vector space $V$ of differentiable functions. We have $D(e^{5t}) = 5e^{5t}$. Hence 5 is an eigenvalue of $D$ with eigenvector $e^{5t}$.

If $A$ is an $n$-square matrix over $K$, then an eigenvalue of $A$ means an eigenvalue of $A$ viewed as an operator on $K^n$. That is, $\lambda \in K$ is an eigenvalue of $A$ if, for some nonzero (column) vector $v \in K^n$,

$$Av = \lambda v$$

In this case $v$ is an eigenvector of $A$ belonging to $\lambda$.

**Example 9.5:** Find eigenvalues and associated nonzero eigenvectors of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$.

We seek a scalar $t$ and a nonzero vector $X = \begin{pmatrix} x \\ y \end{pmatrix}$ such that $AX = tX$:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = t\begin{pmatrix} x \\ y \end{pmatrix}$$

The above matrix equation is equivalent to the homogeneous system

$$\begin{cases} x + 2y = tx \\ 3x + 2y = ty \end{cases} \quad \text{or} \quad \begin{cases} (t-1)x - 2y = 0 \\ -3x + (t-2)y = 0 \end{cases} \qquad (1)$$

Recall that the homogeneous system has a nonzero solution if and only if the determinant of the matrix of coefficients is 0:

$$\begin{vmatrix} t-1 & -2 \\ -3 & t-2 \end{vmatrix} = t^2 - 3t - 4 = (t-4)(t+1) = 0$$

Thus $t$ is an eigenvalue of $A$ if and only if $t = 4$ or $t = -1$.

Setting $t = 4$ in *(1)*,

$$\begin{cases} 3x - 2y = 0 \\ -3x + 2y = 0 \end{cases} \quad \text{or simply} \quad 3x - 2y = 0$$

Thus $v = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is a nonzero eigenvector belonging to the eigenvalue $t = 4$, and every other eigenvector belonging to $t = 4$ is a multiple of $v$.

Setting $t = -1$ in *(1)*,

$$\begin{cases} -2x - 2y = 0 \\ -3x - 3y = 0 \end{cases} \quad \text{or simply} \quad x + y = 0$$

Thus $w = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is a nonzero eigenvector belonging to the eigenvalue $t = -1$, and every other eigenvector belonging to $t = -1$ is a multiple of $w$.

The next theorem gives an important characterization of eigenvalues which is frequently used as its definition.

**Theorem 9.2:** Let $T: V \to V$ be a linear operator on a vector space over $K$. Then $\lambda \in K$ is an eigenvalue of $T$ if and only if the operator $\lambda I - T$ is singular. The eigenspace of $\lambda$ is then the kernel of $\lambda I - T$.

*Proof.* $\lambda$ is an eigenvalue of $T$ if and only if there exists a nonzero vector $v$ such that

$$T(v) = \lambda v \quad \text{or} \quad (\lambda I)(v) - T(v) = 0 \quad \text{or} \quad (\lambda I - T)(v) = 0$$

i.e. $\lambda I - T$ is singular. We also have that $v$ is in the eigenspace of $\lambda$ if and only if the above relations hold; hence $v$ is in the kernel of $\lambda I - T$.

We now st te a very useful theorem which we prove (Problem 9.14) by induction:

**Theorem 9.3:** Nonzero eigenvectors belonging to distinct eigenvalues are linearly independent.

**Example 9.6:** Consider the functions $e^{a_1 t}, e^{a_2 t}, \ldots, e^{a_n t}$ where $a_1, \ldots, a_n$ are distinct real numbers. If $D$ is the differential operator then $D(e^{a_k t}) = a_k e^{a_k t}$. Accordingly, $e^{a_1 t}, \ldots, e^{a_n t}$ are eigenvectors of $D$ belonging to the distinct eigenvalues $a_1, \ldots, a_n$, and so, by Theorem 9.3, are linearly independent.

We remark that independent eigenvectors can belong to the same eigenvalue (see Problem 9.7).

## DIAGONALIZATION AND EIGENVECTORS

Let $T: V \to V$ be a linear operator on a vector space $V$ with finite dimension $n$. Note that $T$ can be represented by a diagonal matrix

$$\begin{pmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & k_n \end{pmatrix}$$

if and only if there exists a basis $\{v_1, \ldots, v_n\}$ of $V$ for which

$$\begin{aligned} T(v_1) &= k_1 v_1 \\ T(v_2) &= \quad k_2 v_2 \\ &\cdots\cdots\cdots\cdots \\ T(v_n) &= \qquad k_n v_n \end{aligned}$$

that is, such that the vectors $v_1, \ldots, v_n$ are eigenvectors of $T$ belonging respectively to eigenvalues $k_1, \ldots, k_n$. In other words:

**Theorem 9.4:** A linear operator $T: V \to V$ can be represented by a diagonal matrix $B$ if and only if $V$ has a basis consisting of eigenvectors of $T$. In this case the diagonal elements of $B$ are the corresponding eigenvalues.

We have the following equivalent statement.

**Alternative Form of Theorem 9.4:** An $n$-square matrix $A$ is similar to a diagonal matrix $B$ if and only if $A$ has $n$ linearly independent eigenvectors. In this case the diagonal elements of $B$ are the corresponding eigenvalues.

In the above theorem, if we let $P$ be the matrix whose columns are the $n$ independent eigenvectors of $A$, then $B = P^{-1}AP$.

**Example 9.7:** Consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. By Example 9.5, $A$ has two independent eigenvectors $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Set $P = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, and so $P^{-1} = \begin{pmatrix} 1/5 & 1/5 \\ 3/5 & -2/5 \end{pmatrix}$. Then $A$ is similar to the diagonal matrix

$$B = P^{-1}AP = \begin{pmatrix} 1/5 & 1/5 \\ 3/5 & -2/5 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

As expected, the diagonal elements 4 and $-1$ of the diagonal matrix $B$ are the eigenvalues corresponding to the given eigenvectors.

## CHARACTERISTIC POLYNOMIAL, CAYLEY-HAMILTON THEOREM

Consider an $n$-square matrix $A$ over a field $K$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

The matrix $tI_n - A$, where $I_n$ is the $n$-square identity matrix and $t$ is an indeterminant, is called the *characteristic matrix* of $A$:

$$tI_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & t - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & t - a_{nn} \end{pmatrix}$$

Its determinant

$$\Delta_A(t) = \det(tI_n - A)$$

which is a polynomial in $t$, is called the *characteristic polynomial* of $A$. We also call

$$\Delta_A(t) = \det(tI_n - A) = 0$$

the *characteristic equation* of $A$.

Now each term in the determinant contains one and only one entry from each row and from each column; hence the above characteristic polynomial is of the form

$$\Delta_A(t) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$
$$+ \text{ terms with at most } n-2 \text{ factors of the form } t - a_{ii}$$

Accordingly,

$$\Delta_A(t) = t^n - (a_{11} + a_{22} + \cdots + a_{nn})t^{n-1} + \text{ terms of lower degree}$$

Recall that the trace of $A$ is the sum of its diagonal elements. Thus the characteristic polynomial $\Delta_A(t) = \det(tI_n - A)$ of $A$ is a monic polynomial of degree $n$, and the coefficient of $t^{n-1}$ is the negative of the trace of $A$. (A polynomial is *monic* if its leading coefficient is 1.)

Furthermore, if we set $t = 0$ in $\Delta_A(t)$, we obtain

$$\Delta_A(0) = |-A| = (-1)^n |A|$$

But $\Delta_A(0)$ is the constant term of the polynomial $\Delta_A(t)$. Thus the constant term of the characteristic polynomial of the matrix $A$ is $(-1)^n|A|$ where $n$ is the order of $A$.

**Example 9.8:** The characteristic polynomial of the matrix $A = \begin{pmatrix} 1 & 3 & 0 \\ -2 & 2 & -1 \\ 4 & 0 & -2 \end{pmatrix}$ is

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -3 & 0 \\ 2 & t-2 & 1 \\ -4 & 0 & t+2 \end{vmatrix} = t^3 - t^2 + 2t + 28$$

As expected, $\Delta(t)$ is a monic polynomial of degree 3.

We now state one of the most important theorems in linear algebra.

**Cayley-Hamilton Theorem 9.5:** Every matrix is a zero of its characteristic polynomial.

**Example 9.9:** The characteristic polynomial of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ is

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -2 \\ -3 & t-2 \end{vmatrix} = t^2 - 3t - 4$$

As expected from the Cayley-Hamilton theorem, $A$ is a zero of $\Delta(t)$:

$$\Delta(A) = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} - 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The next theorem shows the intimate relationship between characteristic polynomials and eigenvalues.

**Theorem 9.6:** Let $A$ be an $n$-square matrix over a field $K$. A scalar $\lambda \in K$ is an eigenvalue of $A$ if and only if $\lambda$ is a root of the characteristic polynomial $\Delta(t)$ of $A$.

*Proof.* By Theorem 9.2, $\lambda$ is an eigenvalue of $A$ if and only if $\lambda I - A$ is singular. Furthermore, by Theorem 8.4, $\lambda I - A$ is singular if and only if $|\lambda I - A| = 0$, i.e. $\lambda$ is a root of $\Delta(t)$. Thus the theorem is proved.

Using Theorems 9.3, 9.4 and 9.6, we obtain

**Corollary 9.7:** If the characteristic polynomial $\Delta(t)$ of an $n$-square matrix $A$ is a product of distinct linear factors:

$$\Delta(t) = (t-a_1)(t-a_2)\cdots(t-a_n)$$

i.e. if $a_1, \ldots, a_n$ are distinct roots of $\Delta(t)$, then $A$ is similar to a diagonal matrix whose diagonal elements are the $a_i$.

Furthermore, using the Fundamental Theorem of Algebra (every polynomial over $\mathbf{C}$ has a root) and the above theorem, we obtain

**Corollary 9.8:** Let $A$ be an $n$-square matrix over the complex field $\mathbf{C}$. Then $A$ has at least one eigenvalue.

**Example 9.10:** Let $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix}$. Its characteristic polynomial is

$$\Delta(t) = \begin{vmatrix} t-3 & 0 & 0 \\ 0 & t-2 & 5 \\ 0 & -1 & t+2 \end{vmatrix} = (t-3)(t^2+1)$$

We consider two cases:

(i) $A$ is a matrix over the real field $\mathbf{R}$. Then $A$ has only the one eigenvalue 3. Since 3 has only one independent eigenvector, $A$ is not diagonalizable.

(ii) $A$ is a matrix over the complex field $\mathbf{C}$. Then $A$ has three distinct eigenvalues: 3, $i$ and $-i$. Thus there exists an invertible matrix $P$ over the complex field $\mathbf{C}$ for which

$$P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$$

i.e. $A$ is diagonalizable.

Now suppose $A$ and $B$ are similar matrices, say $B = P^{-1}AP$ where $P$ is invertible. We show that $A$ and $B$ have the same characteristic polynomial. Using $tI = P^{-1}tIP$,

$$\begin{aligned} |tI - B| &= |tI - P^{-1}AP| = |P^{-1}tIP - P^{-1}AP| \\ &= |P^{-1}(tI - A)P| = |P^{-1}|\,|tI - A|\,|P| \end{aligned}$$

Since determinants are scalars and commute, and since $|P^{-1}|\,|P| = 1$, we finally obtain

$$|tI - B| = |tI - A|$$

Thus we have proved

**Theorem 9.9:** Similar matrices have the same characteristic polynomial.

## MINIMUM POLYNOMIAL

Let $A$ be an $n$-square matrix over a field $K$. Observe that there are nonzero polynomials $f(t)$ for which $f(A) = 0$; for example, the characteristic polynomial of $A$. Among these polynomials we consider those of lowest degree and from them we select one whose leading coefficient is 1, i.e. which is monic. Such a polynomial $m(t)$ exists and is unique (Problem 9.25); we call it the *minimum polynomial* of $A$.

**Theorem 9.10:** The minimum polynomial $m(t)$ of $A$ divides every polynomial which has $A$ as a zero. In particular, $m(t)$ divides the characteristic polynomial $\Delta(t)$ of $A$.

There is an even stronger relationship between $m(t)$ and $\Delta(t)$.

**Theorem 9.11:** The characteristic and minimum polynomials of a matrix $A$ have the same irreducible factors.

This theorem does not say that $m(t) = \Delta(t)$, only that any irreducible factor of one must divide the other. In particular, since a linear factor is irreducible, $m(t)$ and $\Delta(t)$ have the same linear factors, hence they have the same roots. Thus from Theorem 9.6 we obtain

**Theorem 9.12:** A scalar $\lambda$ is an eigenvalue for a matrix $A$ if and only if $\lambda$ is a root of the minimum polynomial of $A$.

**Example 9.11:** Find the minimum polynomial $m(t)$ of the matrix $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$

The characteristic polynomial of $A$ is $\Delta(t) = |tI - A| = (t-2)^3(t-5)$. By Theorem 9.11, both $t-2$ and $t-5$ must be factors of $m(t)$. But by Theorem 9.10, $m(t)$ must divide $\Delta(t)$; hence $m(t)$ must be one of the following three polynomials

$$m_1(t) = (t-2)(t-5), \quad m_2(t) = (t-2)^2(t-5), \quad m_3(t) = (t-2)^3(t-5)$$

We know from the Cayley-Hamilton theorem that $m_3(A) = \Delta(A) = 0$. The reader can verify that $m_1(A) \neq 0$ but $m_2(A) = 0$. Accordingly, $m_2(t) = (t-2)^2(t-5)$ is the minimum polynomial of $A$.

**Example 9.12:** Let $A$ be a 3 by 3 matrix over the real field $\mathbf{R}$. We show that $A$ cannot be a zero of the polynomial $f(t) = t^2 + 1$. By the Cayley-Hamilton theorem, $A$ is a zero of its characteristic polynomial $\Delta(t)$. Note that $\Delta(t)$ is of degree 3, hence it has at least one real root.

Now suppose $A$ is a zero of $f(t)$. Since $f(t)$ is irreducible over $\mathbf{R}$, $f(t)$ must be the minimal polynomial of $A$. But $f(t)$ has no real root. This contradicts the fact that the characteristic and minimal polynomials have the same roots. Thus $A$ is not a zero of $f(t)$.

The reader can verify that the following 3 by 3 matrix over the complex field $\mathbf{C}$ is a zero of $f(t)$:

$$\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & i \end{pmatrix}$$

## CHARACTERISTIC AND MINIMUM POLYNOMIALS OF LINEAR OPERATORS

Now suppose $T: V \to V$ is a linear operator on a vector space $V$ with finite dimension. We define the *characteristic polynomial* $\Delta(t)$ of $T$ to be the characteristic polynomial of any matrix representation of $T$. By Theorem 9.9, $\Delta(t)$ is independent of the particular basis in which the matrix representation is computed. Note that the degree of $\Delta(t)$ is equal to the dimension of $V$. We have theorems for $T$ which are similar to the ones we had for matrices:

**Theorem 9.5′:** $T$ is a zero of its characteristic polynomial.

**Theorem 9.6′:** The scalar $\lambda \in K$ is an eigenvalue of $T$ if and only if $\lambda$ is a root of the characteristic polynomial of $T$.

The *algebraic multiplicity* of an eigenvalue $\lambda \in K$ of $T$ is defined to be the multiplicity of $\lambda$ as a root of the characteristic polynomial of $T$. The *geometric multiplicity* of the eigenvalue $\lambda$ is defined to be the dimension of its eigenspace.

**Theorem 9.13:** The geometric multiplicity of an eigenvalue $\lambda$ does not exceed its algebraic multiplicity.

**Example 9.13:** Let $V$ be the vector space of functions which has $\{\sin\theta, \cos\theta\}$ as a basis, and let $D$ be the differential operator on $V$. Then

$$D(\sin\theta) = \cos\theta = 0(\sin\theta) + 1(\cos\theta)$$
$$D(\cos\theta) = -\sin\theta = -1(\sin\theta) + 0(\cos\theta)$$

The matrix $A$ of $D$ in the above basis is therefore $A = [D] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Thus

$$\det(tI - A) = \begin{vmatrix} t & -1 \\ 1 & t \end{vmatrix} = t^2 + 1$$

and the characteristic polynomial of $D$ is $\Delta(t) = t^2 + 1$.

On the other hand, the *minimum polynomial* $m(t)$ of the operator $T$ is defined independently of the theory of matrices, as the polynomial of lowest degree and leading coefficient 1 which has $T$ as a zero. However, for any polynomial $f(t)$,

$$f(T) = 0 \quad \text{if and only if} \quad f(A) = 0$$

where $A$ is any matrix representation of $T$. Accordingly, $T$ and $A$ have the same minimum polynomial. We remark that all the theorems in this chapter on the minimum polynomial of a matrix also hold for the minimum polynomial of the operator $T$.

## Solved Problems

### POLYNOMIALS OF MATRICES AND LINEAR OPERATORS

**9.1.** Find $f(A)$ where $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$ and $f(t) = t^2 - 3t + 7$.

$$f(A) = A^2 - 3A + 7I = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 12 & 9 \end{pmatrix}$$

**9.2.** Show that $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ is a zero of $f(t) = t^2 - 4t - 5$.

$$f(A) = A^2 - 4A - 5I = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}^2 - 4\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

**9.3.** Let $V$ be the vector space of functions which has $\{\sin\theta, \cos\theta\}$ as a basis, and let $D$ be the differential operator on $V$. Show that $D$ is a zero of $f(t) = t^2 + 1$.

Apply $f(D)$ to each basis vector:

$$f(D)(\sin\theta) = (D^2 + I)(\sin\theta) = D^2(\sin\theta) + I(\sin\theta) = -\sin\theta + \sin\theta = 0$$
$$f(D)(\cos\theta) = (D^2 + I)(\cos\theta) = D^2(\cos\theta) + I(\cos\theta) = -\cos\theta + \cos\theta = 0$$

Since each basis vector is mapped into 0, every vector $v \in V$ is also mapped into 0 by $f(D)$. Thus $f(D) = 0$.

This result is expected since, by Example 9.13, $f(t)$ is the characteristic polynomial of $D$.

9.4. Let $A$ be a matrix representation of an operator $T$. Show that $f(A)$ is the matrix representation of $f(T)$, for any polynomial $f(t)$.

Let $\phi$ be the mapping $T \mapsto A$, i.e. which sends the operator $T$ into its matrix representation $A$. We need to prove that $\phi(f(T)) = f(A)$. Suppose $f(t) = a_n t^n + \cdots + a_1 t + a_0$. The proof is by induction on $n$, the degree of $f(t)$.

Suppose $n = 0$. Recall that $\phi(I') = I$ where $I'$ is the identity mapping and $I$ is the identity matrix. Thus

$$\phi(f(T)) = \phi(a_0 I') = a_0 \phi(I') = a_0 I = f(A)$$

and so the theorem holds for $n = 0$.

Now assume the theorem holds for polynomials of degree less than $n$. Then since $\phi$ is an algebra isomorphism,

$$\begin{aligned} \phi(f(T)) &= \phi(a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') \\ &= a_n \phi(T)\, \phi(T^{n-1}) + \phi(a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I') \\ &= a_n A A^{n-1} + (a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I) = f(A) \end{aligned}$$

and the theorem is proved.

**9.5.** Prove Theorem 9.1: Let $f$ and $g$ be polynomials over $K$. Let $A$ be a square matrix over $K$. Then: (i) $(f+g)(A) = f(A) + g(A)$; (ii) $(fg)(A) = f(A)\,g(A)$; and (iii) $(kf)(A) = k\,f(A)$ where $k \in K$.

Suppose $f = a_n t^n + \cdots + a_1 t + a_0$ and $g = b_m t^m + \cdots + b_1 t + b_0$. Then by definition,

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I \quad \text{and} \quad g(A) = b_m A^m + \cdots + b_1 A + b_0 I$$

(i) Suppose $m \le n$ and let $b_i = 0$ if $i > m$. Then

$$f + g = (a_n + b_n)t^n + \cdots + (a_1 + b_1)t + (a_0 + b_0)$$

Hence

$$\begin{aligned} (f+g)(A) &= (a_n + b_n)A^n + \cdots + (a_1 + b_1)A + (a_0 + b_0)I \\ &= a_n A^n + b_n A^n + \cdots + a_1 A + b_1 A + a_0 I + b_0 I = f(A) + g(A) \end{aligned}$$

(ii) By definition, $fg = c_{n+m} t^{n+m} + \cdots + c_1 t + c_0 = \sum_{k=0}^{n+m} c_k t^k$ where $c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0 = \sum_{i=0}^{k} a_i b_{k-i}$. Hence $(fg)(A) = \sum_{k=0}^{n+m} c_k A^k$ and

$$f(A)\,g(A) = \left(\sum_{i=0}^{n} a_i A^i\right)\left(\sum_{j=0}^{m} b_j A^j\right) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j A^{i+j} = \sum_{k=0}^{n+m} c_k A^k = (fg)(A)$$

(iii) By definition, $kf = k a_n t^n + \cdots + k a_1 t + k a_0$, and so

$$(kf)(A) = k a_n A^n + \cdots + k a_1 A + k a_0 I = k(a_n A^n + \cdots + a_1 A + a_0 I) = k\,f(A)$$

## EIGENVALUES AND EIGENVECTORS

**9.6.** Let $\lambda$ be an eigenvalue of an operator $T: V \to V$. Let $V_\lambda$ denote the set of all eigenvectors of $T$ belonging to the eigenvalue $\lambda$ (called the *eigenspace* of $\lambda$). Show that $V_\lambda$ is a subspace of $V$.

Suppose $v, w \in V_\lambda$; that is, $T(v) = \lambda v$ and $T(w) = \lambda w$. Then for any scalars $a, b \in K$,

$$T(av + bw) = a\,T(v) + b\,T(w) = a(\lambda v) + b(\lambda w) = \lambda(av + bw)$$

Thus $av + bw$ is an eigenvector belonging to $\lambda$, i.e. $av + bw \in V_\lambda$. Hence $V_\lambda$ is a subspace of $V$.

**9.7.** Let $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. (i) Find all eigenvalues of $A$ and the corresponding eigenvectors. (ii) Find an invertible matrix $P$ such that $P^{-1}AP$ is diagonal.

(i) Form the characteristic matrix $tI - A$ of $A$:

$$tI - A = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} t-1 & -4 \\ -2 & t-3 \end{pmatrix} \tag{1}$$

The characteristic polynomial $\Delta(t)$ of $A$ is its determinant:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & -4 \\ -2 & t-3 \end{vmatrix} = t^2 - 4t - 5 = (t-5)(t+1)$$

The roots of $\Delta(t)$ are 5 and $-1$, and so these numbers are the eigenvalues of $A$.

We obtain the eigenvectors belonging to the eigenvalue 5. First substitute $t = 5$ into the characteristic matrix (1) to obtain the matrix $\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}$. The eigenvectors belonging to 5 form the solution of the homogeneous system determined by the above matrix, i.e.,

$$\begin{pmatrix} 4 & -4 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} 4x - 4y = 0 \\ -2x + 2y = 0 \end{cases} \quad \text{or} \quad x - y = 0$$

(In other words, the eigenvectors belonging to 5 form the kernel of the operator $tI - A$ for $t = 5$.) The above system has only one independent solution; for example, $x = 1$, $y = 1$. Thus $v = (1, 1)$ is an eigenvector which generates the eigenspace of 5, i.e. every eigenvector belonging to 5 is a multiple of $v$.

We obtain the eigenvectors belonging to the eigenvalue $-1$. Substitute $t = -1$ into (1) to obtain the homogeneous system

$$\begin{pmatrix} -2 & -4 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -2x - 4y = 0 \\ -2x - 4y = 0 \end{cases} \quad \text{or} \quad x + 2y = 0$$

The system has only one independent solution; for example, $x = 2$, $y = -1$. Thus $w = (2, -1)$ is an eigenvector which generates the eigenspace of $-1$.

(ii) Let $P$ be the matrix whose columns are the above eigenvectors: $P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$. Then $B = P^{-1}AP$ is the diagonal matrix whose diagonal entries are the respective eigenvalues:

$$B = P^{-1}AP = \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & -1/3 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

(*Remark.* Here $P$ is the transition matrix from the usual basis of $\mathbf{R}^2$ to the basis of eigenvectors $\{v, w\}$. Hence $B$ is the matrix representation of the operator $A$ in this new basis.)

**9.8.** For each matrix, find all eigenvalues and a basis of each eigenspace:

$$\text{(i)} \quad A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \qquad \text{(ii)} \quad B = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$$

Which matrix can be diagonalized, and why?

(i) Form the characteristic matrix $tI - A$ and compute its determinant to obtain the characteristic polynomial $\Delta(t)$ of $A$:

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-1 & 3 & -3 \\ -3 & t+5 & -3 \\ -6 & 6 & t-4 \end{vmatrix} = (t+2)^2(t-4)$$

The roots of $\Delta(t)$ are $-2$ and 4; hence these numbers are the eigenvalues of $A$.

We find a basis of the eigenspace of the eigenvalue $-2$. Substitute $t = -2$ into the characteristic matrix $tI - A$ to obtain the homogeneous system

$$\begin{pmatrix} -3 & 3 & -3 \\ -3 & 3 & -3 \\ -6 & 6 & -6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -3x + 3y - 3z = 0 \\ -3x + 3y - 3z = 0 \\ -6x + 6y - 6z = 0 \end{cases} \quad \text{or} \quad x - y + z = 0$$

The system has two independent solutions, e.g. $x = 1, y = 1, z = 0$ and $x = 1, y = 0, z = -1$. Thus $u = (1, 1, 0)$ and $v = (1, 0, -1)$ are independent eigenvectors which generate the eigenspace of $-2$. That is, $u$ and $v$ form a basis of the eigenspace of $-2$. This means that every eigenvector belonging to $-2$ is a linear combination of $u$ and $v$.

We find a basis of the eigenspace of the eigenvalue 4. Substitute $t = 4$ into the characteristic matrix $tI - A$ to obtain the homogeneous system

$$\begin{pmatrix} 3 & 3 & -3 \\ -3 & 9 & -3 \\ -6 & 6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} 3x + 3y - 3z = 0 \\ -3x + 9y - 3z = 0 \\ -6x + 6y \qquad\;\; = 0 \end{cases} \quad \text{or} \quad \begin{cases} x + y - z = 0 \\ \quad 2y - z = 0 \end{cases}$$

The system has only one free variable; hence any particular nonzero solution, e.g. $x = 1, y = 1, z = 2$, generates its solution space. Thus $w = (1, 1, 2)$ is an eigenvector which generates, and so forms a basis, of the eigenspace of 4.

* Since $A$ has three linearly independent eigenvectors, $A$ is diagonalizable. In fact, let $P$ be the matrix whose columns are the three independent eigenvectors:

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}. \quad \text{Then} \quad P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

As expected, the diagonal elements of $P^{-1}AP$ are the eigenvalues of $A$ corresponding to the columns of $P$.

(ii)
$$\Delta(t) = |tI - B| = \begin{vmatrix} t+3 & -1 & 1 \\ 7 & t-5 & 1 \\ 6 & -6 & t+2 \end{vmatrix} = (t+2)^2(t-4)$$

The eigenvalues of $B$ are therefore $-2$ and 4.

We find a basis of the eigenspace of the eigenvalue $-2$. Substitute $t = -2$ into $tI - B$ to obtain the homogeneous system

$$\begin{pmatrix} 1 & -1 & 1 \\ 7 & -7 & 1 \\ 6 & -6 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} x - y + z = 0 \\ 7x - 7y + z = 0 \\ 6x - 6y \qquad = 0 \end{cases} \quad \text{or} \quad \begin{cases} x - y + z = 0 \\ x - y \qquad = 0 \end{cases}$$

The system has only one independent solution, e.g. $x = 1, y = 1, z = 0$. Thus $u = (1, 1, 0)$ forms a basis of the eigenspace of $-2$.

We find a basis of the eigenspace of the eigenvalue 4. Substitute $t = 4$ into $tI - B$ to obtain the homogeneous system

$$\begin{pmatrix} 7 & -1 & 1 \\ 7 & -1 & 1 \\ 6 & -6 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} 7x - y + z = 0 \\ 7x - y + z = 0 \\ 6x - 6y + 6z = 0 \end{cases} \quad \text{or} \quad \begin{cases} 7x - y + z = 0 \\ \qquad\qquad x = 0 \end{cases}$$

The system has only one independent solution, e.g. $x = 0, y = 1, z = 1$. Thus $v = (0, 1, 1)$ forms a basis of the eigenspace of 4.

Observe that $B$ is not similar to a diagonal matrix since $B$ has only two independent eigenvectors. Furthermore, since $A$ can be diagonalized but $B$ cannot, $A$ and $B$ are not similar matrices, even though they have the same characteristic polynomial.

**9.9.** Let $A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$. Find all eigenvalues and the corresponding eigenvectors of $A$ and $B$ viewed as matrices over (i) the real field **R**, (ii) the complex field **C**.

(i)
$$\Delta_A(t) = |tI - A| = \begin{vmatrix} t-3 & 1 \\ -1 & t-1 \end{vmatrix} = t^2 - 4t + 4 = (t-2)^2$$

Hence only 2 is an eigenvalue. Put $t = 2$ into $tI - A$ and obtain the homogeneous system

$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -x + y = 0 \\ -x + y = 0 \end{cases} \quad \text{or} \quad x - y = 0$$

The system has only one independent solution, e.g. $x = 1$, $y = 1$. Thus $v = (1, 1)$ is an eigenvector which generates the eigenspace of 2, i.e. every eigenvector belonging to 2 is a multiple of $v$.

We also have

$$\Delta_B(t) = |tI - B| = \begin{vmatrix} t-1 & 1 \\ -2 & t+1 \end{vmatrix} = t^2 + 1$$

Since $t^2 + 1$ has no solution in **R**, $B$ has no eigenvalue as a matrix over **R**.

(ii) Since $\Delta_A(t) = (t-2)^2$ has only the real root 2, the results are the same as in (i). That is, 2 is an eigenvalue of $A$, and $v = (1, 1)$ is an eigenvector which generates the eigenspace of 2, i.e. every eigenvector of 2 is a (complex) multiple of $v$.

The characteristic matrix of $B$ is $\Delta_B(t) = |tI - B| = t^2 + 1$. Hence $i$ and $-i$ are the eigenvalues of $B$.

We find the eigenvectors associated with $t = i$. Substitute $t = i$ in $tI - B$ to obtain the homogeneous system

$$\begin{pmatrix} i-1 & 1 \\ -2 & i+1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} (i-1)x + y = 0 \\ -2x + (i+1)y = 0 \end{cases} \quad \text{or} \quad (i-1)x + y = 0$$

The system has only one independent solution, e.g. $x = 1$, $y = 1 - i$. Thus $w = (1, 1-i)$ is an eigenvector which generates the eigenspace of $i$.

Now substitute $t = -i$ into $tI - B$ to obtain the homogeneous system

$$\begin{pmatrix} -i-1 & 1 \\ -2 & -i-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} (-i-1)x + y = 0 \\ -2x + (-i-1)y = 0 \end{cases} \quad \text{or} \quad (-i-1)x + y = 0$$

The system has only one independent solution, e.g. $x = 1$, $y = 1 + i$. Thus $w' = (1, 1+i)$ is an eigenvector which generates the eigenspace of $-i$.

**9.10.** Find all eigenvalues and a basis of each eigenspace of the operator $T: \mathbf{R}^3 \to \mathbf{R}^3$ defined by $T(x, y, z) = (2x + y,\ y - z,\ 2y + 4z)$.

First find a matrix representation of $T$, say relative to the usual basis of $\mathbf{R}^3$:

$$A = [T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

The characteristic polynomial $\Delta(t)$ of $T$ is then

$$\Delta(t) = |tI - A| = \begin{vmatrix} t-2 & -1 & 0 \\ 0 & t-1 & 1 \\ 0 & -2 & t-4 \end{vmatrix} = (t-2)^2(t-3)$$

Thus 2 and 3 are the eigenvalues of $T$.

We find a basis of the eigenspace of the eigenvalue 2. Substitute $t = 2$ into $tI - A$ to obtain the homogeneous system

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -y = 0 \\ y + z = 0 \\ -2y - 2z = 0 \end{cases} \quad \text{or} \quad \begin{cases} y = 0 \\ y + z = 0 \end{cases}$$

The system has only one independent solution, e.g. $x = 1$, $y = 0$, $z = 0$. Thus $u = (1, 0, 0)$ forms a basis of the eigenspace of 2.

We find a basis of the eigenspace of the eigenvalue 3. Substitute $t = 3$ into $tI - A$ to obtain the homogeneous system

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \\ -2y - z = 0 \end{cases} \quad \text{or} \quad \begin{cases} x - y = 0 \\ 2y + z = 0 \end{cases}$$

The system has only one independent solution, e.g. $x = 1$, $y = 1$, $z = -2$. Thus $v = (1, 1, -2)$ forms a basis of the eigenspace of 3.

Observe that $T$ is not diagonalizable, since $T$ has only two linearly independent eigenvectors.

**9.11.** Show that 0 is an eigenvalue of $T$ if and only if $T$ is singular.

We have that 0 is an eigenvalue of $T$ if and only if there exists a nonzero vector $v$ such that $T(v) = 0v = 0$, i.e. that $T$ is singular.

**9.12.** Let $A$ and $B$ be $n$-square matrices. Show that $AB$ and $BA$ have the same eigenvalues.

By Problem 9.11 and the fact that the product of nonsingular matrices is nonsingular, the following statements are equivalent: (i) 0 is an eigenvalue of $AB$, (ii) $AB$ is singular, (iii) $A$ or $B$ is singular, (iv) $BA$ is singular, (v) 0 is an eigenvalue of $BA$.

Now suppose $\lambda$ is a nonzero eigenvalue of $AB$. Then there exists a nonzero vector $v$ such that $ABv = \lambda v$. Set $w = Bv$. Since $\lambda \neq 0$ and $v \neq 0$,

$$Aw = ABv = \lambda v \neq 0 \quad \text{and so} \quad w \neq 0$$

But $w$ is an eigenvector of $BA$ belonging to the eigenvalue $\lambda$ since

$$BAw = BABv = B\lambda v = \lambda Bv = \lambda w$$

Hence $\lambda$ is an eigenvalue of $BA$. Similarly, any nonzero eigenvalue of $BA$ is also an eigenvalue of $AB$.

Thus $AB$ and $BA$ have the same eigenvalues.

**9.13.** Suppose $\lambda$ is an eigenvalue of an invertible operator $T$. Show that $\lambda^{-1}$ is an eigenvalue of $T^{-1}$.

Since $T$ is invertible, it is also nonsingular; hence by Problem 9.11, $\lambda \neq 0$.

By definition of an eigenvalue, there exists a nonzero vector $v$ for which $T(v) = \lambda v$. Applying $T^{-1}$ to both sides, we obtain $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$. Hence $T^{-1}(v) = \lambda^{-1} v$; that is, $\lambda^{-1}$ is an eigenvalue of $T^{-1}$.

**9.14.** Prove Theorem 9.3: Let $v_1, \ldots, v_n$ be nonzero eigenvectors of an operator $T: V \to V$ belonging to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $v_1, \ldots, v_n$ are linearly independent.

The proof is by induction on $n$. If $n = 1$, then $v_1$ is linearly independent since $v_1 \neq 0$. Assume $n > 1$. Suppose

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0 \tag{1}$$

where the $a_i$ are scalars. Applying $T$ to the above relation, we obtain by linearity

$$a_1T(v_1) + a_2T(v_2) + \cdots + a_nT(v_n) = T(0) = 0$$

But by hypothesis $T(v_i) = \lambda_i v_i$; hence

$$a_1\lambda_1v_1 + a_2\lambda_2v_2 + \cdots + a_n\lambda_nv_n = 0 \tag{2}$$

On the other hand, multiplying (1) by $\lambda_n$,

$$a_1\lambda_n v_1 + a_2\lambda_n v_2 + \cdots + a_n\lambda_n v_n = 0 \tag{3}$$

Now subtracting (3) from (2),

$$a_1(\lambda_1 - \lambda_n)v_1 + a_2(\lambda_2 - \lambda_n)v_2 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} = 0$$

By induction, each of the above coefficients is 0. Since the $\lambda_i$ are distinct, $\lambda_i - \lambda_n \neq 0$ for $i \neq n$. Hence $a_1 = \cdots = a_{n-1} = 0$. Substituting this into (1) we get $a_n v_n = 0$, and hence $a_n = 0$. Thus the $v_i$ are linearly independent.

## CHARACTERISTIC POLYNOMIAL, CAYLEY-HAMILTON THEOREM

**9.15.** Consider a triangular matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Find its characteristic polynomial $\Delta(t)$ and its eigenvalues.

Since $A$ is triangular and $tI$ is diagonal, $tI - A$ is also triangular with diagonal elements $t - a_{ii}$:

$$tI - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ 0 & t - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & t - a_{nn} \end{pmatrix}$$

Then $\Delta(t) = |tI - A|$ is the product of the diagonal elements $t - a_{ii}$:

$$\Delta(t) = (t - a_{11})(t - a_{22})\cdots(t - a_{nn})$$

Hence the eigenvalues of $A$ are $a_{11}, a_{22}, \dots, a_{nn}$, i.e. its diagonal elements.

**9.16.** Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$. Is $A$ similar to a diagonal matrix? If so, find one such matrix.

Since $A$ is triangular, the eigenvalues of $A$ are the diagonal elements 1, 2 and 3. Since they are distinct, $A$ is similar to a diagonal matrix whose diagonal elements are 1, 2 and 3; for example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

**9.17.** For each matrix find a polynomial having the matrix as a root:

(i) $A = \begin{pmatrix} 2 & 5 \\ 1 & -3 \end{pmatrix}$, (ii) $B = \begin{pmatrix} 2 & -3 \\ 7 & -4 \end{pmatrix}$, (iii) $C = \begin{pmatrix} 1 & 4 & -3 \\ 0 & 3 & 1 \\ 0 & 2 & -1 \end{pmatrix}$.

By the Cayley-Hamilton theorem every matrix is a root of its characteristic polynomial. Therefore we find the characteristic polynomial $\Delta(t)$ in each case.

(i) $\Delta(t) = |tI - A| = \begin{vmatrix} t - 2 & -5 \\ -1 & t + 3 \end{vmatrix} = t^2 + t - 11$

(ii) $\Delta(t) = |tI - B| = \begin{vmatrix} t - 2 & 3 \\ -7 & t + 4 \end{vmatrix} = t^2 + 2t + 13$

(iii) $\Delta(t) = |tI - C| = \begin{vmatrix} t - 1 & -4 & 3 \\ 0 & t - 3 & -1 \\ 0 & -2 & t + 1 \end{vmatrix} = (t - 1)(t^2 - 2t - 5)$

v.v.9

**9.18.** Prove the Cayley-Hamilton Theorem 9.5: Every matrix is a zero of its characteristic polynomial.

Let $A$ be an arbitrary $n$-square matrix and let $\Delta(t)$ be its characteristic polynomial; say,

$$\Delta(t) = |tI - A| = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Now let $B(t)$ denote the classical adjoint of the matrix $tI - A$. The elements of $B(t)$ are cofactors of the matrix $tI - A$ and hence are polynomials in $t$ of degree not exceeding $n-1$. Thus

$$B(t) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

where the $B_i$ are $n$-square matrices over $K$ which are independent of $t$. By the fundamental property of the classical adjoint (Theorem 8.8),

$$(tI - A)B(t) = |tI - A|I$$

or

$$(tI - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Removing parentheses and equating the coefficients of corresponding powers of $t$,

$$\begin{aligned} B_{n-1} &= I \\ B_{n-2} - AB_{n-1} &= a_{n-1}I \\ B_{n-3} - AB_{n-2} &= a_{n-2}I \\ &\cdots\cdots \\ B_0 - AB_1 &= a_1 I \\ -AB_0 &= a_0 I \end{aligned}$$

Multiplying the above matrix equations by $A^n, A^{n-1}, \ldots, A, I$ respectively,

$$\begin{aligned} A^nB_{n-1} &= A^n \\ A^{n-1}B_{n-2} - A^nB_{n-1} &= a_{n-1}A^{n-1} \\ A^{n-2}B_{n-3} - A^{n-1}B_{n-2} &= a_{n-2}A^{n-2} \\ &\cdots\cdots \\ AB_0 - A^2B_1 &= a_1 A \\ -AB_0 &= a_0 I \end{aligned}$$

Adding the above matrix equations,

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I$$

In other words, $\Delta(A) = 0$. That is, $A$ is a zero of its characteristic polynomial.

**9.19.** Show that a matrix $A$ and its transpose $A^t$ have the same characteristic polynomial.

v.v.9

By the transpose operation, $(tI - A)^t = tI^t - A^t = tI - A^t$. Since a matrix and its transpose have the same determinant, $|tI - A| = |(tI - A)^t| = |tI - A^t|$. Hence $A$ and $A^t$ have the same characteristic polynomial.

**9.20.** Suppose $M = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ where $A_1$ and $A_2$ are square matrices. Show that the characteristic polynomial of $M$ is the product of the characteristic polynomials of $A_1$ and $A_2$. Generalize.

$tI - M = \begin{pmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{pmatrix}$. Hence by Problem 8.70, $|tI - M| = \begin{vmatrix} tI - A_1 & -B \\ 0 & tI - A_2 \end{vmatrix} =$ $|tI - A||tI - B|$, as required.

By induction, the characteristic polynomial of the triangular block matrix

$$M = \begin{pmatrix} A_1 & B & \cdots & C \\ 0 & A_2 & \cdots & D \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_n \end{pmatrix}$$

where the $A_i$ are square matrices, is the product of the characteristic polynomials of the $A_i$.

## MINIMUM POLYNOMIAL

**9.21.** Find the minimum polynomial $m(t)$ of $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}$.

The characteristic polynomial of $A$ is

$$\Delta(t) = \begin{vmatrix} t-2 & -1 & 0 & 0 \\ 0 & t-2 & 0 & 0 \\ 0 & 0 & t-1 & -1 \\ 0 & 0 & 2 & t-4 \end{vmatrix} = \begin{vmatrix} t-2 & -1 \\ 0 & t-2 \end{vmatrix} \begin{vmatrix} t-1 & -1 \\ 2 & t-4 \end{vmatrix} = (t-3)(t-2)^3$$

The minimum polynomial $m(t)$ must divide $\Delta(t)$. Also, each irreducible factor of $\Delta(t)$, i.e. $t-2$ and $t-3$, must be a factor of $m(t)$. Thus $m(t)$ is exactly one of the following:

$$f(t) = (t-3)(t-2), \quad g(t) = (t-3)(t-2)^2, \quad h(t) = (t-3)(t-2)^3$$

We have

$$f(A) = (A-3I)(A-2I) = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 2 \end{pmatrix} \neq 0$$

$$g(A) = (A-3I)(A-2I)^2 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 2 \end{pmatrix}^2 = 0$$

Thus $g(t) = (t-3)(t-2)^2$ is the minimum polynomial of $A$.

*Remark.* We know that $h(A) = \Delta(A) = 0$ by the Cayley-Hamilton theorem. However, the degree of $g(t)$ is less than the degree of $h(t)$; hence $g(t)$, and not $h(t)$, is the minimum polynomial of $A$.

**9.22.** Find the minimal polynomial $m(t)$ of each matrix (where $a \neq 0$):

(i) $A = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}$, (ii) $B = \begin{pmatrix} \lambda & a & 0 \\ 0 & \lambda & a \\ 0 & 0 & \lambda \end{pmatrix}$, (iii) $C = \begin{pmatrix} \lambda & a & 0 & 0 \\ 0 & \lambda & a & 0 \\ 0 & 0 & \lambda & a \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.

(i) The characteristic polynomial of $A$ is $\Delta(t) = (t-\lambda)^2$. We find $A - \lambda I \neq 0$; hence $m(t) = \Delta(t) = (t-\lambda)^2$.

(ii) The characteristic polynomial of $B$ is $\Delta(t) = (t-\lambda)^3$. (Note $m(t)$ is exactly one of $t-\lambda$, $(t-\lambda)^2$ or $(t-\lambda)^3$.) We find $(B-\lambda I)^2 \neq 0$; thus $m(t) = \Delta(t) = (t-\lambda)^3$.

(iii) The characteristic polynomial of $C$ is $\Delta(t) = (t-\lambda)^4$. We find $(C-\lambda I)^3 \neq 0$; hence $m(t) = \Delta(t) = (t-\lambda)^4$.

**9.23.** Let $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ where $A$ and $B$ are square matrices. Show that the minimum polynomial $m(t)$ of $M$ is the least common multiple of the minimum polynomials $g(t)$ and $h(t)$ of $A$ and $B$ respectively. Generalize.

Since $m(t)$ is the minimum polynomial of $M$, $m(M) = \begin{pmatrix} m(A) & 0 \\ 0 & m(B) \end{pmatrix} = 0$ and hence $m(A) = 0$ and $m(B) = 0$. Since $g(t)$ is the minimum polynomial of $A$, $g(t)$ divides $m(t)$. Similarly, $h(t)$ divides $m(t)$. Thus $m(t)$ is a multiple of $g(t)$ and $h(t)$.

Now let $f(t)$ be another multiple of $g(t)$ and $h(t)$; then $f(M) = \begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$. But $m(t)$ is the minimum polynomial of $M$; hence $m(t)$ divides $f(t)$. Thus $m(t)$ is the least common multiple of $g(t)$ and $h(t)$.

We then have, by induction, that the minimum polynomial of

$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$$

where the $A_i$ are square matrices, is the least common multiple of the minimum polynomials of the $A_i$.

**9.24.** Find the minimum polynomial $m(t)$ of

$$M = \begin{pmatrix} 2 & 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

Let $A = \begin{pmatrix} 2 & 8 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$, $D = (5)$. The minimum polynomials of $A$, $C$ and $D$ are $(t-2)^2$, $t^2$ and $t-5$ respectively. The characteristic polynomial of $B$ is

$$|tI - B| = \begin{vmatrix} t-4 & -2 \\ -1 & t-3 \end{vmatrix} = t^2 - 7t + 10 = (t-2)(t-5)$$

and so it is also the minimum polynomial of $B$.

Observe that $M = \begin{pmatrix} A & 0 & 0 & 0 \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & D \end{pmatrix}$. Thus $m(t)$ is the least common multiple of the minimum polynomials of $A$, $B$, $C$ and $D$. Accordingly, $m(t) = t^2(t-2)^2(t-5)$.

**9.25.** Show that the minimum polynomial of a matrix (operator) $A$ exists and is unique.

By the Cayley-Hamilton theorem, $A$ is a zero of some nonzero polynomial (see also Problem 9.31). Let $n$ be the lowest degree for which a polynomial $f(t)$ exists such that $f(A) = 0$. Dividing $f(t)$ by its leading coefficient, we obtain a monic polynomial $m(t)$ of degree $n$ which has $A$ as a zero. Suppose $m'(t)$ is another monic polynomial of degree $n$ for which $m'(A) = 0$. Then the difference $m(t) - m'(t)$ is a nonzero polynomial of degree less than $n$ which has $A$ as a zero. This contradicts the original assumption on $n$; hence $m(t)$ is a unique minimum polynomial.

**9.26.** Prove Theorem 9.10. The minimum polynomial $m(t)$ of a matrix (operator) $A$ divides every polynomial which has $A$ as a zero. In particular, $m(t)$ divides the characteristic polynomial of $A$.

Suppose $f(t)$ is a polynomial for which $f(A) = 0$. By the division algorithm there exist polynomials $q(t)$ and $r(t)$ for which $f(t) = m(t)\,q(t) + r(t)$ and $r(t) = 0$ or $\deg r(t) < \deg m(t)$. Substituting $t = A$ in this equation, and using that $f(A) = 0$ and $m(A) = 0$, we obtain $r(A) = 0$. If $r(t) \neq 0$, then $r(t)$ is a polynomial of degree less than $m(t)$ which has $A$ as a zero; this contradicts the definition of the minimum polynomial. Thus $r(t) = 0$ and so $f(t) = m(t)\,q(t)$, i.e. $m(t)$ divides $f(t)$.

86, 90

**9.27.** Let $m(t)$ be the minimum polynomial of an $n$-square matrix $A$. Show that the characteristic polynomial of $A$ divides $(m(t))^n$.

Suppose $m(t) = t^r + c_1 t^{r-1} + \cdots + c_{r-1}t + c_r$. Consider the following matrices:

$$\begin{aligned} B_0 &= I \\ B_1 &= A + c_1 I \\ B_2 &= A^2 + c_1 A + c_2 I \\ &\cdots\cdots\cdots\cdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I \end{aligned}$$

Then

$$\begin{aligned} B_0 &= I \\ B_1 - AB_0 &= c_1 I \\ B_2 - AB_1 &= c_2 I \\ &\cdots\cdots\cdots \\ B_{r-1} - AB_{r-2} &= c_{r-1} I \end{aligned}$$

Also,

$$\begin{aligned} -AB_{r-1} &= c_r I - (A^r + c_1 A^{r-1} + \cdots + c_{r-1}A + c_r I) \\ &= c_r I - m(A) \\ &= c_r I \end{aligned}$$

Set

$$B(t) = t^{r-1}B_0 + t^{r-2}B_1 + \cdots + tB_{r-2} + B_{r-1}$$

Then

$$\begin{aligned} (tI - A)\cdot B(t) &= (t^r B_0 + t^{r-1}B_1 + \cdots + tB_{r-1}) - (t^{r-1}AB_0 + t^{r-2}AB_1 + \cdots + AB_{r-1}) \\ &= t^r B_0 + t^{r-1}(B_1 - AB_0) + t^{r-2}(B_2 - AB_1) + \cdots + t(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= t^r I + c_1 t^{r-1} I + c_2 t^{r-2} I + \cdots + c_{r-1} tI + c_r I \\ &= m(t) I \end{aligned}$$

The determinant of both sides gives $|tI - A|\,|B(t)| = |m(t)\,I| = (m(t))^n$. Since $|B(t)|$ is a polynomial, $|tI - A|$ divides $(m(t))^n$; that is, the characteristic polynomial of $A$ divides $(m(t))^n$.

**9.28.** Prove Theorem 9.11: The characteristic polynomial $\Delta(t)$ and the minimum polynomial $m(t)$ of a matrix $A$ have the same irreducible factors.

Suppose $f(t)$ is an irreducible polynomial. If $f(t)$ divides $m(t)$ then, since $m(t)$ divides $\Delta(t)$, $f(t)$ divides $\Delta(t)$. On the other hand, if $f(t)$ divides $\Delta(t)$ then, by the preceding problem, $f(t)$ divides $(m(t))^n$. But $f(t)$ is irreducible; hence $f(t)$ also divides $m(t)$. Thus $m(t)$ and $\Delta(t)$ have the same irreducible factors.

**9.29.** Let $T$ be a linear operator on a vector space $V$ of finite dimension. Show that $T$ is invertible if and only if the constant term of the minimal (characteristic) polynomial of $T$ is not zero.

Suppose the minimal (characteristic) polynomial of $T$ is $f(t) = t^r + a_{n-1}t^{r-1} + \cdots + a_1 t + a_0$. Each of the following statements is equivalent to the succeeding one by preceding results: (i) $T$ is invertible; (ii) $T$ is nonsingular; (iii) 0 is not an eigenvalue of $T$; (iv) 0 is not a root of $m(t)$; (v) the constant term $a_0$ is not zero. Thus the theorem is proved.

**9.30.** **Suppose $\dim V = n$. Let $T: V \to V$ be an invertible operator. Show that $T^{-1}$ is equal to a polynomial in $T$ of degree not exceeding $n$.**

Let $m(t)$ be the minimal polynomial of $T$. Then $m(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$, where $r \le n$. Since $T$ is invertible, $a_0 \ne 0$. We have

$$m(T) = T^r + a_{r-1}T^{r-1} + \cdots + a_1 T + a_0 I = 0$$

Hence

$$-\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)T = I \quad \text{and} \quad T^{-1} = -\frac{1}{a_0}(T^{r-1} + a_{r-1}T^{r-2} + \cdots + a_1 I)$$

## MISCELLANEOUS PROBLEMS

**9.31.** **Let $T$ be a linear operator on a vector space $V$ of dimension $n$. Without using the Cayley-Hamilton theorem, show that $T$ is a zero of a nonzero polynomial.**

Let $N = n^2$. Consider the following $N+1$ operators on $V$: $I, T, T^2, \ldots, T^N$. Recall that the vector space $A(V)$ of operators on $V$ has dimension $N = n^2$. Thus the above $N+1$ operators are linearly dependent. Hence there exist scalars $a_0, a_1, \ldots, a_N$ for which $a_N T^N + \cdots + a_1 T + a_0 I = 0$. Accordingly, $T$ is a zero of the polynomial $f(t) = a_N t^N + \cdots + a_1 t + a_0$.

**9.32.** **Prove Theorem 9.13: Let $\lambda$ be an eigenvalue of an operator $T: V \to V$. The geometric multiplicity of $\lambda$ does not exceed its algebraic multiplicity.**

Suppose the geometric multiplicity of $\lambda$ is $r$. Then $\lambda$ contains $r$ linearly independent eigenvectors $v_1, \ldots, v_r$. Extend the set $\{v_i\}$ to a basis of $V$: $\{v_1, \ldots, v_r, w_1, \ldots, w_s\}$. We have

$$\begin{aligned}
T(v_1) &= \lambda v_1 \\
T(v_2) &= \qquad \lambda v_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(v_r) &= \qquad\qquad \lambda v_r \\
T(w_1) &= a_{11}v_1 + \cdots + a_{1r}v_r + b_{11}w_1 + \cdots + b_{1s}w_s \\
T(w_2) &= a_{21}v_1 + \cdots + a_{2r}v_r + b_{21}w_1 + \cdots + b_{2s}w_s \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
T(w_s) &= a_{s1}v_1 + \cdots + a_{sr}v_r + b_{s1}w_1 + \cdots + b_{ss}w_s
\end{aligned}$$

The matrix of $T$ in the above basis is

$$M = \left(\begin{array}{cccc|cccc}
\lambda & 0 & \ldots & 0 & a_{11} & a_{21} & \ldots & a_{s1} \\
0 & \lambda & \ldots & 0 & a_{12} & a_{22} & \ldots & a_{s2} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & \lambda & a_{1r} & a_{2r} & \ldots & a_{sr} \\
\hline
0 & 0 & \ldots & 0 & b_{11} & b_{21} & \ldots & b_{s1} \\
0 & 0 & \ldots & 0 & b_{12} & b_{22} & \ldots & b_{s2} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 0 & b_{1s} & b_{2s} & \ldots & b_{ss}
\end{array}\right) = \left(\begin{array}{c|c} \lambda I_r & A \\ \hline 0 & B \end{array}\right)$$

where $A = (a_{ij})^t$ and $B = (b_{ij})^t$.

By Problem 9.20 the characteristic polynomial of $\lambda I_r$, which is $(t-\lambda)^r$, must divide the characteristic polynomial of $M$ and hence $T$. Thus the algebraic multiplicity of $\lambda$ for the operator $T$ is at least $r$, as required.

**9.33.** **Show that $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.**

The characteristic polynomial of $A$ is $\Delta(t) = (t-1)^2$; hence 1 is the only eigenvalue of $A$. We find a basis of the eigenspace of the eigenvalue 1. Substitute $t = 1$ into the matrix $tI - A$ to obtain the homogeneous system

$$\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -y = 0 \\ 0 = 0 \end{cases} \quad \text{or} \quad y = 0$$

The system has only one independent solution, e.g. $x = 1$, $y = 0$. Hence $u = (1, 0)$ forms a basis of the eigenspace of 1.

Since $A$ has at most one independent eigenvector, $A$ cannot be diagonalized.

**9.34.** Let $F$ be an extension of a field $K$. Let $A$ be an $n$-square matrix over $K$. Note that $A$ may also be viewed as a matrix $\hat{A}$ over $F$. Clearly $|tI - A| = |tI - \hat{A}|$, that is, $A$ and $\hat{A}$ have the same characteristic polynomial. Show that $A$ and $\hat{A}$ also have the same minimum polynomial.

Let $m(t)$ and $m'(t)$ be the minimum polynomials of $A$ and $\hat{A}$ respectively. Now $m'(t)$ divides every polynomial over $F$ which has $A$ as a zero. Since $m(t)$ has $A$ as a zero and since $m(t)$ may be viewed as a polynomial over $F$, $m'(t)$ divides $m(t)$. We show now that $m(t)$ divides $m'(t)$.

Since $m'(t)$ is a polynomial over $F$ which is an extension of $K$, we may write

$$m'(t) = f_1(t)\,b_1 + f_2(t)\,b_2 + \cdots + f_n(t)\,b_n$$

where $f_i(t)$ are polynomials over $K$, and $b_1, \ldots, b_n$ belong to $F$ and are linearly independent over $K$. We have

$$m'(A) = f_1(A)\,b_1 + f_2(A)\,b_2 + \cdots + f_n(A)\,b_n = 0 \tag{1}$$

Let $a_{ij}^{(k)}$ denote the $ij$-entry of $f_k(A)$. The above matrix equation implies that, for each pair $(i, j)$,

$$a_{ij}^{(1)}b_1 + a_{ij}^{(2)}b_2 + \cdots + a_{ij}^{(n)}b_n = 0$$

Since the $b_i$ are linearly independent over $K$ and since the $a_{ij}^{(k)} \in K$, every $a_{ij}^{(k)} = 0$. Then

$$f_1(A) = 0,\; f_2(A) = 0,\; \ldots,\; f_n(A) = 0$$

Since the $f_i(t)$ are polynomials over $K$ which have $A$ as a zero and since $m(t)$ is the minimum polynomial of $A$ as a matrix over $K$, $m(t)$ divides each of the $f_i(t)$. Accordingly, by (1), $m(t)$ must also divide $m'(t)$. But monic polynomials which divide each other are necessarily equal. That is, $m(t) = m'(t)$, as required.

**9.35.** Let $\{v_1, \ldots, v_n\}$ be a basis of $V$. Let $T: V \to V$ be an operator for which $T(v_1) = 0$, $T(v_2) = a_{21}v_1$, $T(v_3) = a_{31}v_1 + a_{32}v_2$, $\ldots$, $T(v_n) = a_{n1}v_1 + \cdots + a_{n,n-1}v_{n-1}$. Show that $T^n = 0$.

It suffices to show that

$$T^j(v_j) = 0 \tag{*}$$

for $j = 1, \ldots, n$. For then it follows that

$$T^n(v_j) = T^{n-j}(T^j(v_j)) = T^{n-j}(0) = 0, \quad \text{for } j = 1, \ldots, n$$

and, since $\{v_1, \ldots, v_n\}$ is a basis, $T^n = 0$.

We prove (*) by induction on $j$. The case $j = 1$ is true by hypothesis. The inductive step follows (for $j = 2, \ldots, n$) from

$$\begin{aligned} T^j(v_j) &= T^{j-1}(T(v_j)) = T^{j-1}(a_{j1}v_1 + \cdots + a_{j,j-1}v_{j-1}) \\ &= a_{j1}T^{j-1}(v_1) + \cdots + a_{j,j-1}T^{j-1}(v_{j-1}) \\ &= a_{j1}0 + \cdots + a_{j,j-1}0 = 0 \end{aligned}$$

*Remark.* Observe that the matrix representation of $T$ in the above basis is triangular with diagonal elements 0:

$$\begin{pmatrix} 0 & a_{21} & a_{31} & \ldots & a_{n1} \\ 0 & 0 & a_{32} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{n,n-1} \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

## Supplementary Problems

### POLYNOMIALS OF MATRICES AND LINEAR OPERATORS

**9.36.** Let $f(t) = 2t^2 - 5t + 6$ and $g(t) = t^3 - 2t^2 + t + 3$. Find $f(A)$, $g(A)$, $f(B)$ and $g(B)$ where $A = \begin{pmatrix} 2 & -3 \\ 5 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

**9.37.** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $T(x, y) = (x + y, 2x)$. Let $f(t) = t^2 - 2t + 3$. Find $f(T)(x, y)$.

**9.38.** Let $V$ be the vector space of polynomials $v(x) = ax^2 + bx + c$. Let $D: V \to V$ be the differential operator. Let $f(t) = t^2 + 2t - 5$. Find $f(D)(v(x))$.

**9.39.** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Find $A^2, A^3, A^n$.

**9.40.** Let $B = \begin{pmatrix} 8 & 12 & 0 \\ 0 & 8 & 12 \\ 0 & 0 & 8 \end{pmatrix}$. Find a real matrix $A$ such that $B = A^3$.

**9.41.** Consider a diagonal matrix $M$ and a triangular matrix $N$:

$$M = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} a_1 & b & \dots & c \\ 0 & a_2 & \dots & d \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix}$$

Show that, for any polynomial $f(t)$, $f(M)$ and $f(N)$ are of the form

$$f(M) = \begin{pmatrix} f(a_1) & 0 & \dots & 0 \\ 0 & f(a_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(a_n) \end{pmatrix} \quad \text{and} \quad f(N) = \begin{pmatrix} f(a_1) & x & \dots & y \\ 0 & f(a_2) & \dots & z \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(a_n) \end{pmatrix}$$

**9.42.** Consider a block diagonal matrix $M$ and a block triangular matrix $N$:

$$M = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} A_1 & B & \dots & C \\ 0 & A_2 & \dots & D \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{pmatrix}$$

where the $A_i$ are square matrices. Show that, for any polynomial $f(t)$, $f(M)$ and $f(N)$ are of the form

$$f(M) = \begin{pmatrix} f(A_1) & 0 & \dots & 0 \\ 0 & f(A_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(A_n) \end{pmatrix} \quad \text{and} \quad f(N) = \begin{pmatrix} f(A_1) & X & \dots & Y \\ 0 & f(A_2) & \dots & Z \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(A_n) \end{pmatrix}$$

**9.43.** Show that for any square matrix (or operator) $A$, $(P^{-1}AP)^n = P^{-1}A^nP$ where $P$ is invertible. More generally, show that $f(P^{-1}AP) = P^{-1}f(A)P$ for any polynomial $f(t)$.

**9.44.** Let $f(t)$ be any polynomial. Show that: (i) $f(A^t) = (f(A))^t$; (ii) if $A$ is symmetric, i.e. $A^t = A$, then $f(A)$ is symmetric.

### EIGENVALUES AND EIGENVECTORS

**9.45.** For each matrix, find all eigenvalues and linearly independent eigenvectors:

$$\text{(i)}\ A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}, \quad \text{(ii)}\ B = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}, \quad \text{(iii)}\ C = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}$$

Find invertible matrices $P_1$, $P_2$ and $P_3$ such that $P_1^{-1}AP_1$, $P_2^{-1}BP_2$ and $P_3^{-1}CP_3$ are diagonal.

**9.46.** For each matrix, find all eigenvalues and a basis for each eigenspace:

(i) $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}$, (ii) $B = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}$, (iii) $C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

When possible, find invertible matrices $P_1$, $P_2$ and $P_3$ such that $P_1^{-1}AP_1$, $P_2^{-1}BP_2$ and $P_3^{-1}CP_3$ are diagonal.

**9.47.** Consider $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & -1 \\ 13 & -3 \end{pmatrix}$ as matrices over the real field $\mathbf{R}$. Find all eigenvalues and linearly independent eigenvectors.

**9.48.** Consider $A$ and $B$ in the preceding problem as matrices over the complex field $\mathbf{C}$. Find all eigenvalues and linearly independent eigenvectors.

**9.49.** For each of the following operators $T: \mathbf{R}^2 \to \mathbf{R}^2$, find all eigenvalues and a basis for each eigenspace: (i) $T(x, y) = (3x + 3y, x + 5y)$, (ii) $T(x, y) = (y, x)$, (iii) $T(x, y) = (y, -x)$.

87, 85 **9.50.** For each of the following operators $T: \mathbf{R}^3 \to \mathbf{R}^3$, find all eigenvalues and a basis for each eigenspace: (i) $T(x, y, z) = (x + y + z, 2y + z, 2y + 3z)$; (ii) $T(x, y, z) = (x + y, y + z, -2y - z)$; (iii) $T(x, y, z) = (x - y, 2x + 3y + 2z, x + y + 2z)$.

**9.51.** For each of the following matrices over the complex field $\mathbf{C}$, find all eigenvalues and linearly independent eigenvectors:

(i) $\begin{pmatrix} 1 & i \\ 0 & i \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & -3i \\ i & -1 \end{pmatrix}$, (iv) $\begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$.

**9.52.** Suppose $v$ is an eigenvector of operators $S$ and $T$. Show that $v$ is also an eigenvector of the operator $aS + bT$ where $a$ and $b$ are any scalars.

**9.53.** Suppose $v$ is an eigenvector of an operator $T$ belonging to the eigenvalue $\lambda$. Show that for $n > 0$, $v$ is also an eigenvector of $T^n$ belonging to $\lambda^n$.

**9.54.** Suppose $\lambda$ is an eigenvalue of an operator $T$. Show that $f(\lambda)$ is an eigenvalue of $f(T)$.

**9.55.** Show that similar matrices have the same eigenvalues.

**9.56.** Show that matrices $A$ and $A^t$ have the same eigenvalues. Give an example where $A$ and $A^t$ have different eigenvectors.

**9.57.** Let $S$ and $T$ be linear operators such that $ST = TS$. Let $\lambda$ be an eigenvalue of $T$ and let $W$ be its eigenspace. Show that $W$ is invariant under $S$, i.e. $S(W) \subset W$.

**9.58.** Let $V$ be a vector space of finite dimension over the complex field $\mathbf{C}$. Let $W \neq \{0\}$ be a subspace of $V$ invariant under a linear operator $T: V \to V$. Show that $W$ contains a nonzero eigenvector of $T$.

**9.59.** Let $A$ be an $n$-square matrix over $K$. Let $v_1, \ldots, v_n \in K^n$ be linearly independent eigenvectors of $A$ belonging to the eigenvalues $\lambda_1, \ldots, \lambda_n$ respectively. Let $P$ be the matrix whose columns are the vectors $v_1, \ldots, v_n$. Show that $P^{-1}AP$ is the diagonal matrix whose diagonal elements are the eigenvalues $\lambda_1, \ldots, \lambda_n$.

## CHARACTERISTIC AND MINIMUM POLYNOMIALS

**9.60.** For each matrix, find a polynomial for which the matrix is a root:

(i) $A = \begin{pmatrix} 3 & -7 \\ 4 & 5 \end{pmatrix}$, (ii) $B = \begin{pmatrix} 5 & -1 \\ 8 & 3 \end{pmatrix}$, (iii) $C = \begin{pmatrix} 2 & 3 & -2 \\ 0 & 5 & 4 \\ 1 & 0 & -1 \end{pmatrix}$.

**9.61.** Consider the $n$-square matrix

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

Show that $f(t) = (t-\lambda)^n$ is both the characteristic and minimum polynomial of $A$.

**9.62.** Find the characteristic and minimum polynomials of each matrix:

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} \lambda & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

**9.63.** Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Show that $A$ and $B$ have different characteristic polynomials (and so are not similar), but have the same minimum polynomial. Thus nonsimilar matrices may have the same minimum polynomial.

**9.64.** The mapping $T: V \to V$ defined by $T(v) = kv$ is called the *scalar mapping* belonging to $k \in K$. Show that $T$ is the scalar mapping belonging to $k \in K$ if and only if the minimal polynomial of $T$ is $m(t) = t - k$.

**9.65.** Let $A$ be an $n$-square matrix for which $A^k = 0$ for some $k > n$. Show that $A^n = 0$.

**9.66.** Show that a matrix $A$ and its transpose $A^t$ have the same minimum polynomial.

**9.67.** Suppose $f(t)$ is an irreducible monic polynomial for which $f(T) = 0$ where $T$ is a linear operator $T: V \to V$. Show that $f(t)$ is the minimal polynomial of $T$.

**9.68.** Consider a block matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Show that $tI - M = \begin{pmatrix} tI - A & -B \\ -C & tI - D \end{pmatrix}$ is the characteristic matrix of $M$.

**9.69.** Let $T$ be a linear operator on a vector space $V$ of finite dimension. Let $W$ be a subspace of $V$ invariant under $T$, i.e. $T(W) \subset W$. Let $T_W: W \to W$ be the restriction of $T$ to $W$. (i) Show that the characteristic polynomial of $T_W$ divides the characteristic polynomial of $T$. (ii) Show that the minimum polynomial of $T_W$ divides the minimum polynomial of $T$.

**9.70.** Let $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Show that the characteristic polynomial of $A$ is

$$\Delta(t) = t^3 - (a_{11}+a_{22}+a_{33})t^2 + \left( \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \right) t - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

**9.71.** Let $A$ be an $n$-square matrix. The determinant of the matrix of order $n-m$ obtained by deleting the rows and columns passing through $m$ diagonal elements of $A$ is called a *principal minor* of degree $n-m$. Show that the coefficient of $t^m$ in the characteristic polynomial $\Delta(t) = |tI - A|$ is the sum of all principal minors of $A$ of degree $n-m$ multiplied by $(-1)^{n-m}$. (Observe that the preceding problem is a special case of this result.)

**9.72.** Consider an arbitrary monic polynomial $f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. The following $n$-square matrix $A$ is called the *companion matrix* of $f(t)$:

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

Show that $f(t)$ is the minimum polynomial of $A$.

**9.73.** Find a matrix $A$ whose minimum polynomial is (i) $t^3 - 5t^2 + 6t + 8$, (ii) $t^4 - 5t^3 - 2t^2 + 7t + 4$.

## DIAGONALIZATION

**9.74.** Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a matrix over the real field **R**. Find necessary and sufficient conditions on $a, b, c$ and $d$ so that $A$ is diagonalizable, i.e. has two linearly independent eigenvectors.

**9.75.** Repeat the preceding problem for the case that $A$ is a matrix over the complex field **C**.

**9.76.** Show that a matrix (operator) is diagonalizable if and only if its minimal polynomial is a product of distinct linear factors.

**9.77.** Let $A$ and $B$ be $n$-square matrices over $K$ such that (i) $AB = BA$ and (ii) $A$ and $B$ are both diagonalizable. Show that $A$ and $B$ can be simultaneously diagonalized, i.e. there exists a basis of $K^n$ in which both $A$ and $B$ are represented by diagonal matrices. (See Problem 9.57.)

**9.78.** Let $E: V \to V$ be a projection operator, i.e. $E^2 = E$. Show that $E$ is diagonalizable and, in fact, can be represented by the diagonal matrix $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where $r$ is the rank of $E$.

# Answers to Supplementary Problems

**9.36.** $f(A) = \begin{pmatrix} -26 & -3 \\ 5 & -27 \end{pmatrix}$, $g(A) = \begin{pmatrix} -40 & 39 \\ -65 & -27 \end{pmatrix}$, $f(B) = \begin{pmatrix} 3 & 6 \\ 0 & 9 \end{pmatrix}$, $g(B) = \begin{pmatrix} 3 & 12 \\ 0 & 15 \end{pmatrix}$.

**9.37.** $f(T)(x, y) = (4x - y, -2x + 5y)$.

**9.38.** $f(D)(v(x)) = -5ax^2 + (4a - 5b)x + (2a + 2b - 5c)$.

**9.39.** $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

**9.40.** *Hint.* Let $A = \begin{pmatrix} 2 & a & b \\ 0 & 2 & c \\ 0 & 0 & 2 \end{pmatrix}$. Set $B = A^3$ and then obtain conditions on $a$, $b$ and $c$.

a b

**9.44.** (ii) Using (i), e have $(f(A))^t = f(A^t) = f(A)$.

**9.45.** (i) $\lambda_1 = 1$, $u = (2, -1)$; $\lambda_2 = 4$, $v = (1, 1)$.

(ii) $\lambda_1 = 1$, $u = (2, -3)$; $\lambda_2 = 6$, $v = (1, 1)$.

(iii) $\lambda = 4$, $u = (1, 1)$.

Let $P_1 = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ and $P_2 = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$. $P_3$ does not exist since $C$ has only one independent eigenvector, and so cannot be diagonalized.

**9.46.** (i) $\lambda_1 = 2$, $u = (1, -1, 0)$, $v = (1, 0, -1)$; $\lambda_2 = 6$, $w = (1, 2, 1)$.
(ii) $\lambda_1 = 3$, $u = (1, 1, 0)$, $v = (1, 0, 1)$; $\lambda_2 = 1$, $w = (2, -1, 1)$.
(iii) $\lambda = 1$, $u = (1, 0, 0)$, $v = (0, 0, 1)$.

Let $P_1 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{pmatrix}$ and $P_2 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$. $P_3$ does not exist since $C$ has at most two linearly independent eigenvectors, and so cannot be diagonalized.

**9.47.** (i) $\lambda = 3$, $u = (1, -1)$; (ii) $B$ has no eigenvalues (in **R**).

**9.48.** (i) $\lambda = 3$, $u = (1, -1)$. (ii) $\lambda_1 = 2i$, $u = (1, 3 - 2i)$; $\lambda_2 = -2i$, $v = (1, 3 + 2i)$.

**9.49.** (i) $\lambda_1 = 2$, $u = (3, -1)$; $\lambda_2 = 6$, $v = (1, 1)$. (ii) $\lambda_1 = 1$, $u = (2, 1)$; $\lambda_2 = -1$, $v = (2, -1)$. (iii) There are no eigenvalues (in **R**).

**9.50.** (i) $\lambda_1 = 1$, $u = (1, 0, 0)$; $\lambda_2 = 4$, $v = (1, 1, 2)$.
(ii) $\lambda = 1$, $u = (1, 0, 0)$. There are no other eigenvalues (in **R**).
(iii) $\lambda_1 = 1$, $u = (1, 0, -1)$; $\lambda_2 = 2$, $v = (2, -2, -1)$; $\lambda_3 = 3$, $w = (1, -2, -1)$.

**9.51.** (i) $\lambda_1 = 1$, $u = (1, 0)$; $\lambda_2 = i$, $v = (1, 1 + i)$. (ii) $\lambda = 1$, $u = (1, 0)$. (iii) $\lambda_1 = 2$, $u = (3, i)$; $\lambda_2 = -2$, $v = (1, -i)$. (iv) $\lambda_1 = i$, $u = (2, 1 - i)$; $\lambda_2 = -i$, $v = (2, 1 + i)$.

**9.56.** Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $\lambda = 1$ is the only eigenvalue and $v = (1, 0)$ generates the eigenspace of $\lambda = 1$. On the other hand, for $A^t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\lambda = 1$ is still the only eigenvalue, but $w = (0, 1)$ generates the eigenspace of $\lambda = 1$.

**9.57.** Let $v \in W$ and so $T(v) = \lambda v$. Then $T(Sv) = S(Tv) = S(\lambda v) = \lambda(Sv)$, that is, $Sv$ is an eigenvector of $T$ belonging to the eigenvalue $\lambda$. In other words, $Sv \in W$ and thus $S(W) \subset W$.

**9.58.** Let $\hat{T} : W \to W$ be the restriction of $T$ to $W$. The characteristic polynomial of $\hat{T}$ is a polynomial over the complex field **C** which, by the fundamental theorem of algebra, has a root $\lambda$. Then $\lambda$ is an eigenvalue of $\hat{T}$, and so $\hat{T}$ has a nonzero eigenvector in $W$ which is also an eigenvector of $T$.

**9.59.** Suppose $T(v) = \lambda v$. Then $(kT)(v) = kT(v) = k(\lambda v) = (k\lambda)v$.

**9.60.** (i) $f(t) = t^2 - 8t + 43$, (ii) $g(t) = t^2 - 8t + 23$, (iii) $h(t) = t^3 - 6t^2 + 5t - 12$.

**9.62.** (i) $\Delta(t) = (t - 2)^3(t - 7)^2$; $m(t) = (t - 2)^2(t - 7)$. (ii) $\Delta(t) = (t - 3)^5$; $m(t) = (t - 3)^3$. (iii) $\Delta(t) = (t - \lambda)^5$; $m(t) = t - \lambda$.

**9.73.** Use the result of Problem 9.72. (i) $A = \begin{pmatrix} 0 & 0 & -8 \\ 1 & 0 & -6 \\ 0 & 1 & 5 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{pmatrix}$

**9.77.** *Hint.* Use the result of Problem 9.57.