CHAPTER 4

Vector Spaces

4.1 Introduction

This chapter introduces the underlying structure of linear algebra, that of a finite-dimensional vector space. The definition of a vector space V, whose elements are called *vectors*, involves an arbitrary field K, whose elements are called *scalars*. The following notation will be used (unless otherwise stated or implied):

V	the given vector space
u, v, w	vectors in V
K	the given number field
$a, b, c,$ or k	scalars in K

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

The reader might suspect that the real line $\mathbf{R}$ has "dimension" one, the cartesian plane $\mathbf{R}^2$ has "dimension" two, and the space $\mathbf{R}^3$ has "dimension" three. This chapter formalizes the notion of "dimension," and this definition will agree with the reader's intuition.

Throughout this text, we will use the following set notation:

$a \in A$	Element a belongs to set A
$a, b \in A$	Elements a and b belong to A
$\forall x \in A$	For every x in A
$\exists x \in A$	There exists an x in A
$A \subseteq B$	A is a subset of B
$A \cap B$	Intersection of A and B
$A \cup B$	Union of A and B
$\emptyset$	Empty set

4.2 Vector Spaces

The following defines the notion of a vector space V where K is the field of scalars.

DEFINITION: Let V be a nonempty set with two operations:

(i) ***Vector Addition:*** This assigns to any $u, v \in V$ a *sum* $u + v$ in V.

(ii) ***Scalar Multiplication:*** This assigns to any $u \in V$, $k \in K$ a *product* $ku \in V$.

Then V is called a *vector space* (over the field K) if the following axioms hold for any vectors $u, v, w \in V$:

[A_1] $(u+v)+w=u+(v+w)$

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

[A_3] For each $u \in V$, there is a vector in V, denoted by $-u$, and called the *negative* of u, such that
$$u+(-u)=(-u)+u=0.$$

[A_4] $u+v=v+u$.

[M_1] $k(u+v)=ku+kv$, for any scalar $k \in K$.

[M_2] $(a+b)u=au+bu$, for any scalars $a, b \in K$.

[M_3] $(ab)u=a(bu)$, for any scalars $a, b \in K$.

[M_4] $1u=u$, for the unit scalar $1 \in K$.

The above axioms naturally split into two sets (as indicated by the labeling of the axioms). The first four are concerned only with the additive structure of V and can be summarized by saying V is a *commutative group under addition*. This means

(a) Any sum $v_1+v_2+\cdots+v_m$ of vectors requires no parentheses and does not depend on the order of the summands.

(b) The zero vector 0 is unique, and the negative $-u$ of a vector u is unique.

(c) (Cancellation Law) If $u+w=v+w$, then $u=v$.

Also, *subtraction* in V is defined by $u-v=u+(-v)$, where $-v$ is the unique negative of v.

On the other hand, the remaining four axioms are concerned with the "action" of the field K of scalars on the vector space V. Using these additional axioms, we prove (Problem 4.2) 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 $k \in K$ and any $u \in V$, $(-k)u=k(-u)=-ku$.

4.3 Examples of Vector Spaces

This section lists important examples of vector spaces that will be used throughout the text.

Space K^n

Let K be an arbitrary field. The notation K^n is frequently used to denote the set of all n-tuples of elements in K. Here K^n is a vector space over K using the following operations:

(i) ***Vector Addition:*** $(a_1, a_2, \ldots, a_n)+(b_1, b_2, \ldots, b_n)=(a_1+b_1,\ a_2+b_2, \ldots,\ a_n+b_n)$

(ii) ***Scalar Multiplication:*** $k(a_1, a_2, \ldots, a_n)=(ka_1, ka_2, \ldots, ka_n)$

The zero vector in K^n is the n-tuple of zeros,
$$0=(0,0,\ldots,0)$$
and the negative of a vector is defined by
$$-(a_1, a_2, \ldots, a_n)=(-a_1, -a_2, \ldots, -a_n)$$

Observe that these are the same as the operations defined for $\mathbf{R}^n$ in Chapter 1. The proof that K^n is a vector space is identical to the proof of Theorem 1.1, which we now regard as stating that $\mathbf{R}^n$ with the operations defined there is a vector space over $\mathbf{R}$.

Polynomial Space $\mathbf{P}(t)$

Let $\mathbf{P}(t)$ denote the set of all polynomials of the form

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s \qquad (s = 1, 2, \ldots)$$

where the coefficients a_i belong to a field K. Then $\mathbf{P}(t)$ is a vector space over K using the following operations:

(i) ***Vector Addition:*** Here $p(t) + q(t)$ in $\mathbf{P}(t)$ is the usual operation of addition of polynomials.

(ii) ***Scalar Multiplication:*** Here $kp(t)$ in $\mathbf{P}(t)$ is the usual operation of the product of a scalar k and a polynomial $p(t)$.

The zero polynomial 0 is the zero vector in $\mathbf{P}(t)$.

Polynomial Space $\mathbf{P}_n(t)$

Let $\mathbf{P}_n(t)$ denote the set of all polynomials $p(t)$ over a field K, where the degree of $p(t)$ is less than or equal to n; that is,

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s$$

where $s \leq n$. Then $\mathbf{P}_n(t)$ is a vector space over K with respect to the usual operations of addition of polynomials and of multiplication of a polynomial by a constant (just like the vector space $\mathbf{P}(t)$ above). We include the zero polynomial 0 as an element of $\mathbf{P}_n(t)$, even though its degree is undefined.

Matrix Space $\mathbf{M}_{m,n}$

The notation $\mathbf{M}_{m,n}$, or simply $\mathbf{M}$, will be used to denote the set of all $m \times n$ matrices with entries in a field K. Then $\mathbf{M}_{m,n}$ is a vector space over K with respect to the usual operations of matrix addition and scalar multiplication of matrices, as indicated by Theorem 2.1.

Function Space $F(X)$

Let X be a nonempty set and let K be an arbitrary field. Let $F(X)$ denote the set of all functions of X into K. [Note that $F(X)$ is nonempty, because X is nonempty.] Then $F(X)$ is a vector space over K with respect to the following operations:

(i) ***Vector Addition:*** The sum of two functions f and g in $F(X)$ is the function $f + g$ in $F(X)$ defined by

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

(ii) ***Scalar Multiplication:*** The product of a scalar $k \in K$ and a function f in $F(X)$ is the function kf in $F(X)$ defined by

$$(kf)(x) = kf(x) \qquad \forall x \in X$$

The zero vector in $F(X)$ is the zero function $\mathbf{0}$, which maps every $x \in X$ into the zero element $0 \in K$;

$$\mathbf{0}(x) = 0 \qquad \forall x \in X$$

Also, for any function f in $F(X)$, negative of f is the function $-f$ in $F(X)$ defined by

$$(-f)(x) = -f(x) \qquad \forall x \in X$$

Fields and Subfields

Suppose a field E is an extension of a field K; that is, suppose E is a field that contains K as a subfield. Then E may be viewed as a vector space over K using the following operations:

(i) ***Vector Addition:*** Here $u + v$ in E is the usual addition in E.

(ii) ***Scalar Multiplication:*** Here ku in E, where $k \in K$ and $u \in E$, is the usual product of k and u as elements of E.

That is, the eight axioms of a vector space are satisfied by E and its subfield K with respect to the above two operations.

4.4 Linear Combinations, Spanning Sets

Let V be a vector space over a field K. A vector v in V is a *linear combination* of vectors $u_1, u_2, \ldots, u_m$ in V if there exist scalars $a_1, a_2, \ldots, a_m$ in K such that

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

Alternatively, v is a linear combination of $u_1, u_2, \ldots, u_m$ if there is a solution to the vector equation

$$v = x_1u_1 + x_2u_2 + \cdots + x_mu_m$$

where $x_1, x_2, \ldots, x_m$ are unknown scalars.

Each element of the space X that can be expressed as $\alpha_1x_1 + \alpha_2x_2 + \alpha_3x_3 + \cdots + \alpha_nx_n$ where $\alpha_i \in K$ is called a linear combination of the elements $x_1 + x_2 + \cdots + x_n$ of the vector space X (over the field K)

EXAMPLE 4.1 (Linear Combinations in $\mathbf{R}^n$) Suppose we want to express $v = (3, 7, -4)$ in $\mathbf{R}^3$ as a linear combination of the vectors

$$u_1 = (1, 2, 3), \qquad u_2 = (2, 3, 7), \qquad u_3 = (3, 5, 6)$$

We seek scalars x, y, z such that $v = xu_1 + yu_2 + zu_3$; that is,

$$\begin{bmatrix} 3 \\ 7 \\ -4 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} + z\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix} \qquad \text{or} \qquad \begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 5z &= 7 \\ 3x + 7y + 6z &= -4 \end{aligned}$$

(For notational convenience, we have written the vectors in $\mathbf{R}^3$ as columns, because it is then easier to find the equivalent system of linear equations.) Reducing the system to echelon form yields

$$\begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ y - 3z &= -13 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x + 2y + 3z &= 3 \\ -y - z &= 1 \\ -4z &= -12 \end{aligned}$$

Back-substitution yields the solution $x = 2$, $y = -4$, $z = 3$. Thus, $v = 2u_1 - 4u_2 + 3u_3$.

Remark: Generally speaking, the question of expressing a given vector v in K^n as a linear combination of vectors $u_1, u_2, \ldots, u_m$ in K^n is equivalent to solving a system $AX = B$ of linear equations, where v is the column B of constants, and the u's are the columns of the coefficient matrix A. Such a system may have a unique solution (as above), many solutions, or no solution. The last case—no solution—means that v cannot be written as a linear combination of the u's.

EXAMPLE 4.2 (Linear combinations in $\mathbf{P}(t)$) Suppose we want to express the polynomial $v = 3t^2 + 5t - 5$ as a linear combination of the polynomials

$$p_1 = t^2 + 2t + 1, \qquad p_2 = 2t^2 + 5t + 4, \qquad p_3 = t^2 + 3t + 6$$

We seek scalars x, y, z such that $v = xp_1 + yp_2 + zp_3$; that is,

$$3t^2 + 5t - 5 = x(t^2 + 2t + 1) + y(2t^2 + 5t + 4) + z(t^2 + 3t + 6) \qquad (*)$$

There are two ways to proceed from here.

(1) Expand the right-hand side of (*) obtaining:

$$\begin{aligned} 3t^2 + 5t - 5 &= xt^2 + 2xt + x + 2yt^2 + 5yt + 4y + zt^2 + 3zt + 6z \\ &= (x + 2y + z)t^2 + (2x + 5y + 3z)t + (x + 4y + 6z) \end{aligned}$$

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

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

The system is in triangular form and has a solution. Back-substitution yields the solution $x = 3, y = 1, z = -2$. Thus,

$$v = 3p_1 + p_2 - 2p_3$$

(2) The equation (*) is actually an identity in the variable t; that is, the equation holds for any value of t. We can obtain three equations in the unknowns x, y, z by setting t equal to any three values. For example,

Set $t = 0$ in (1) to obtain: $x + 4y + 6z = -5$

Set $t = 1$ in (1) to obtain: $4x + 11y + 10z = 3$

Set $t = -1$ in (1) to obtain: $y + 4z = -7$

Reducing this system to echelon form and solving by back-substitution again yields the solution $x = 3, y = 1, z = -2$. Thus (again), $v = 3p_1 + p_2 - 2p_3$.

The set of all possible linear combination of the set Z is called the span of the set $Z \subset X$ (Span(S) is the set that contains the linear combinations)

Spanning Sets

Let V be a vector space over K. Vectors $u_1, u_2, \ldots, u_m$ in V are said to *span* V or to form a *spanning set* of V if every v in V is a linear combination of the vectors $u_1, u_2, \ldots, u_m$—that is, if there exist scalars $a_1, a_2, \ldots, a_m$ in K such that

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m$$

The following remarks follow directly from the definition.

Remark 1: Suppose $u_1, u_2, \ldots, u_m$ span V. Then, for any vector w, the set $w, u_1, u_2, \ldots, u_m$ also spans V.

Remark 2: Suppose $u_1, u_2, \ldots, u_m$ span V and suppose u_k is a linear combination of some of the other u's. Then the u's without u_k also span V.

Remark 3: Suppose $u_1, u_2, \ldots, u_m$ span V and suppose one of the u's is the zero vector. Then the u's without the zero vector also span V.

EXAMPLE 4.3 Consider the vector space $V = \mathbf{R}^3$.

(a) We claim that the following vectors form a spanning set of $\mathbf{R}^3$:

$$e_1 = (1, 0, 0), \qquad e_2 = (0, 1, 0), \qquad e_3 = (0, 0, 1)$$

Specifically, if $v = (a, b, c)$ is any vector in $\mathbf{R}^3$, then

$$v = ae_1 + be_2 + ce_3$$

For example, $v = (5, -6, 2) = -5e_1 - 6e_2 + 2e_3$.

(b) We claim that the following vectors also form a spanning set of $\mathbf{R}^3$:

$$w_1 = (1, 1, 1), \qquad w_2 = (1, 1, 0), \qquad w_3 = (1, 0, 0)$$

Specifically, if $v = (a, b, c)$ is any vector in $\mathbf{R}^3$, then (Problem 4.62)

$$v = (a, b, c) = cw_1 + (b - c)w_2 + (a - b)w_3$$

For example, $v = (5, -6, 2) = 2w_1 - 8w_2 + 11w_3$.

(c) One can show (Problem 3.24) that $v = (2, 7, 8)$ cannot be written as a linear combination of the vectors

$$u_1 = (1, 2, 3), \qquad u_2 = (1, 3, 5), \qquad u_3 = (1, 5, 9)$$

Accordingly, u_1, u_2, u_3 do not span $\mathbf{R}^3$.

EXAMPLE 4.4 Consider the vector space $V = \mathbf{P}_n(t)$ consisting of all polynomials of degree $\leq n$.

(a) Clearly every polynomial in $\mathbf{P}_n(t)$ can be expressed as a linear combination of the $n+1$ polynomials

$$1, \quad t, \quad t^2, \quad t^3, \quad \ldots, \quad t^n$$

Thus, these powers of t (where $1 = t^0$) form a spanning set for $\mathbf{P}_n(t)$.

(b) One can also show that, for any scalar c, the following $n+1$ powers of $t-c$,

$$1, \quad t-c, \quad (t-c)^2, \quad (t-c)^3, \quad \ldots, \quad (t-c)^n$$

(where $(t-c)^0 = 1$), also form a spanning set for $\mathbf{P}_n(t)$.

EXAMPLE 4.5 Consider the vector space $\mathbf{M} = \mathbf{M}_{2,2}$ consisting of all 2×2 matrices, and consider the following four matrices in $\mathbf{M}$:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then clearly any matrix A in $\mathbf{M}$ can be written as a linear combination of the four matrices. For example,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Accordingly, the four matrices $E_{11}, E_{12}, E_{21}, E_{22}$ span $\mathbf{M}$.

4.5 Subspaces

This section introduces the important notion of a subspace.

DEFINITION: Let V be a vector space over a field K and let W be a subset of V. Then W is 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.

The way in which one shows that any set W is a vector space is to show that W satisfies the eight axioms of a vector space. However, if W is a subset of a vector space V, then some of the axioms automatically hold in W, because they already hold in V. Simple criteria for identifying subspaces follow.

THEOREM 4.2: Suppose W is a subset of a vector space V. Then W is a subspace of V if the following two conditions hold:

(a) The zero vector 0 belongs to W.

(b) For every $u, v \in W, k \in K$: (i) The sum $u + v \in W$. (ii) The multiple $ku \in W$.

Property (i) in (b) states that W is *closed under vector addition*, and property (ii) in (b) states that W is *closed under scalar multiplication*. Both properties may be combined into the following equivalent single statement:

(b') For every $u, v \in W, a, b \in K$, the linear combination $au + bv \in W$.

Now let V be any vector space. Then V automatically contains two subspaces: the set $\{0\}$ consisting of the zero vector alone and the whole space V itself. These are sometimes called the *trivial* subspaces of V. Examples of nontrivial subspaces follow.

EXAMPLE 4.6 Consider the vector space $V = \mathbf{R}^3$.

(a) Let U consist of all vectors in $\mathbf{R}^3$ whose entries are equal; that is,

$$U = \{(a, b, c) : a = b = c\}$$

For example, $(1,1,1)$, $(-3,-3,-3)$, $(7,7,7)$, $(-2,-2,-2)$ are vectors in U. Geometrically, U is the line through the origin O and the point $(1,1,1)$ as shown in Fig. 4-1(a). Clearly $0 = (0,0,0)$ belongs to U, because

EXAMPLE 4.4 Consider the vector space $V = \mathbf{P}_n(t)$ consisting of all polynomials of degree $\leq n$.

(a) Clearly every polynomial in $\mathbf{P}_n(t)$ can be expressed as a linear combination of the $n + 1$ polynomials

$$1, \quad t, \quad t^2, \quad t^3, \quad \ldots, \quad t^n$$

Thus, these powers of t (where $1 = t^0$) form a spanning set for $\mathbf{P}_n(t)$.

(b) One can also show that, for any scalar c, the following $n+1$ powers of $t - c$,

$$1, \quad t-c, \quad (t-c)^2, \quad (t-c)^3, \quad \ldots, \quad (t-c)^n$$

(where $(t-c)^0 = 1$), also form a spanning set for $\mathbf{P}_n(t)$.

EXAMPLE 4.5 Consider the vector space $\mathbf{M} = \mathbf{M}_{2,2}$ consisting of all 2×2 matrices, and consider the following four matrices in $\mathbf{M}$:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then clearly any matrix A in $\mathbf{M}$ can be written as a linear combination of the four matrices. For example,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Accordingly, the four matrices $E_{11}, E_{12}, E_{21}, E_{22}$ span $\mathbf{M}$.

4.5 Subspaces

This section introduces the important notion of a subspace.

DEFINITION: Let V be a vector space over a field K and let W be a subset of V. Then W is 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.

The way in which one shows that any set W is a vector space is to show that W satisfies the eight axioms of a vector space. However, if W is a subset of a vector space V, then some of the axioms automatically hold in W, because they already hold in V. Simple criteria for identifying subspaces follow.

THEOREM 4.2: Suppose W is a subset of a vector space V. Then W is a subspace of V if the following two conditions hold:

(a) The zero vector 0 belongs to W.

(b) For every $u, v \in W, k \in K$: (i) The sum $u + v \in W$. (ii) The multiple $ku \in W$.

Property (i) in (b) states that W is *closed under vector addition*, and property (ii) in (b) states that W is *closed under scalar multiplication*. Both properties may be combined into the following equivalent single statement:

(b') For every $u, v \in W, a, b \in K$, the linear combination $au + bv \in W$.

Now let V be any vector space. Then V automatically contains two subspaces: the set $\{0\}$ consisting of the zero vector alone and the whole space V itself. These are sometimes called the *trivial* subspaces of V. Examples of nontrivial subspaces follow.

EXAMPLE 4.6 Consider the vector space $V = \mathbf{R}^3$.

(a) Let U consist of all vectors in $\mathbf{R}^3$ whose entries are equal; that is,

$$U = \{(a, b, c) : a = b = c\}$$

For example, $(1,1,1)$, $(-3,-3,-3)$, $(7,7,7)$, $(-2,-2,-2)$ are vectors in U. Geometrically, U is the line through the origin O and the point $(1,1,1)$ as shown in Fig. 4-1(a). Clearly $0 = (0,0,0)$ belongs to U, because

EXAMPLE 4.4 Consider the vector space $V = \mathbf{P}_n(t)$ consisting of all polynomials of degree $\leq n$.

(a) Clearly every polynomial in $\mathbf{P}_n(t)$ can be expressed as a linear combination of the $n+1$ polynomials

$$1, \quad t, \quad t^2, \quad t^3, \quad \ldots, \quad t^n$$

Thus, these powers of t (where $1 = t^0$) form a spanning set for $\mathbf{P}_n(t)$.

(b) One can also show that, for any scalar c, the following $n+1$ powers of $t - c$,

$$1, \quad t-c, \quad (t-c)^2, \quad (t-c)^3, \quad \ldots, \quad (t-c)^n$$

(where $(t-c)^0 = 1$), also form a spanning set for $\mathbf{P}_n(t)$.

EXAMPLE 4.5 Consider the vector space $\mathbf{M} = \mathbf{M}_{2,2}$ consisting of all 2×2 matrices, and consider the following four matrices in $\mathbf{M}$:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then clearly any matrix A in $\mathbf{M}$ can be written as a linear combination of the four matrices. For example,

$$A = \begin{bmatrix} 5 & -6 \\ 7 & 8 \end{bmatrix} = 5E_{11} - 6E_{12} + 7E_{21} + 8E_{22}$$

Accordingly, the four matrices $E_{11}, E_{12}, E_{21}, E_{22}$ span $\mathbf{M}$.

4.5 Subspaces

This section introduces the important notion of a subspace.

DEFINITION: Let V be a vector space over a field K and let W be a subset of V. Then W is 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.

The way in which one shows that any set W is a vector space is to show that W satisfies the eight axioms of a vector space. However, if W is a subset of a vector space V, then some of the axioms automatically hold in W, because they already hold in V. Simple criteria for identifying subspaces follow.

THEOREM 4.2: Suppose W is a subset of a vector space V. Then W is a subspace of V if the following two conditions hold:

(a) The zero vector 0 belongs to W.

(b) For every $u, v \in W, k \in K$: (i) The sum $u + v \in W$. (ii) The multiple $ku \in W$.

Property (i) in (b) states that W is *closed under vector addition*, and property (ii) in (b) states that W is *closed under scalar multiplication*. Both properties may be combined into the following equivalent single statement:

(b') For every $u, v \in W, a, b \in K$, the linear combination $au + bv \in W$.

Now let V be any vector space. Then V automatically contains two subspaces: the set $\{0\}$ consisting of the zero vector alone and the whole space V itself. These are sometimes called the *trivial* subspaces of V. Examples of nontrivial subspaces follow.

EXAMPLE 4.6 Consider the vector space $V = \mathbf{R}^3$.

(a) Let U consist of all vectors in $\mathbf{R}^3$ whose entries are equal; that is,

$$U = \{(a, b, c) : a = b = c\}$$

For example, $(1, 1, 1)$, $(-3, -3, -3)$, $(7, 7, 7)$, $(-2, -2, -2)$ are vectors in U. Geometrically, U is the line through the origin O and the point $(1, 1, 1)$ as shown in Fig. 4-1(a). Clearly $0 = (0, 0, 0)$ belongs to U, because all entries in 0 are equal. Further, suppose u and v are arbitrary vectors in U, say, $u = (a, a, a)$ and $v = (b, b, b)$. Then, for any scalar $k \in \mathbf{R}$, the following are also vectors in U:

$$u + v = (a + b,\ a + b,\ a + b) \qquad \text{and} \qquad ku = (ka,\ ka,\ ka)$$

Thus, U is a subspace of $\mathbf{R}^3$.

(b) Let W be any plane in $\mathbf{R}^3$ passing through the origin, as pictured in Fig. 4-1(b). Then $0 = (0, 0, 0)$ belongs to W, because we assumed W passes through, the origin O. Further, suppose u and v are vectors in W. Then u and v may be viewed as arrows in the plane W emanating from the origin O, as in Fig. 4-1(b). The sum $u + v$ and any multiple ku of u also lie in the plane W. Thus, W is a subspace of $\mathbf{R}^3$.

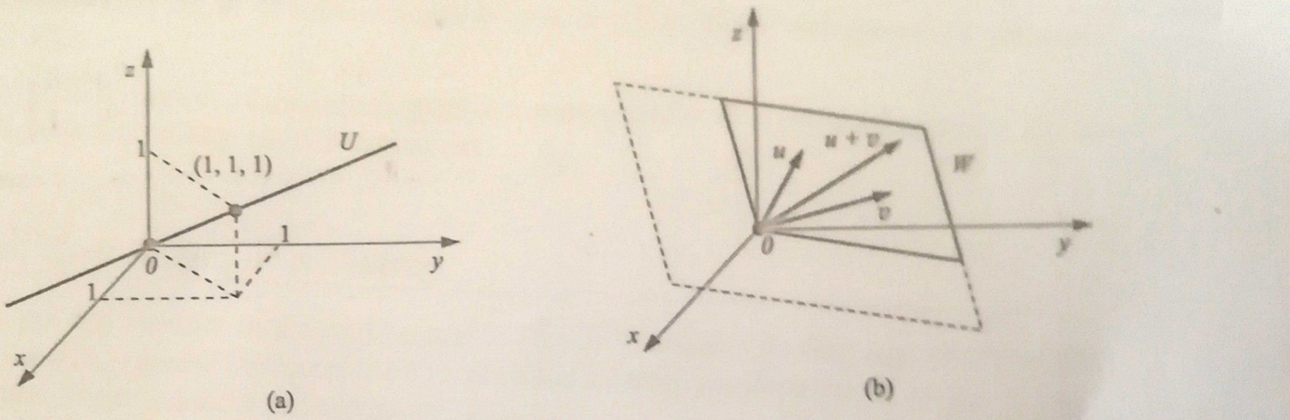


Figure 4-1

EXAMPLE 4.7

(a) Let $V = \mathbf{M}_{n,n}$, the vector space of $n \times n$ matrices. Let W_1 be the subset of all (upper) triangular matrices and let W_2 be the subset of all symmetric matrices. Then W_1 is a subspace of V, because W_1 contains the zero matrix 0 and W_1 is closed under matrix addition and scalar multiplication; that is, the sum and scalar multiple of such triangular matrices are also triangular. Similarly, W_2 is a subspace of V.

(b) Let $V = \mathbf{P}(t)$, the vector space $\mathbf{P}(t)$ of polynomials. Then the space $\mathbf{P}_n(t)$ of polynomials of degree at most n may be viewed as a subspace of $\mathbf{P}(t)$. Let $\mathbf{Q}(t)$ be the collection of polynomials with only even powers of t. For example, the following are polynomials in $\mathbf{Q}(t)$:

$$p_1 = 3 + 4t^2 - 5t^6 \qquad \text{and} \qquad p_2 = 6 - 7t^4 + 9t^6 + 3t^{12}$$

(We assume that any constant $k = kt^0$ is an even power of t.) Then $\mathbf{Q}(t)$ is a subspace of $\mathbf{P}(t)$.

(c) Let V be the vector space of real-valued functions. Then the collection W_1 of continuous functions and the collection W_2 of differentiable functions are subspaces of V.

Intersection of Subspaces

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 $0 \in W$, because U and W are subspaces; whence $0 \in U \cap W$. Now suppose u and v belong to the intersection $U \cap W$. Then $u, v \in U$ and $u, v \in W$. Further, because U and W are subspaces, for any scalars $a, b \in K$,

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

Thus, $au + bv \in U \cap W$. Therefore, $U \cap W$ is a subspace of V.

The above result generalizes as follows.

THEOREM 4.3: The intersection of any number of subspaces of a vector space V is a subspace of V.

Solution Space of a Homogeneous System

Consider a system $AX = B$ of linear equations in n unknowns. Then every solution u may be viewed as a vector in K^n. Thus, the solution set of such a system is a subset of K^n. Now suppose the system is homogeneous; that is, suppose the system has the form $AX = 0$. Let W be its solution set. Because $A0 = 0$, the zero vector $0 \in W$. Moreover, suppose u and v belong to W. Then u and v are solutions of $AX = 0$, or, in other words, $Au = 0$ and $Av = 0$. Therefore, for any scalars a and b, we have

$$A(au + bv) = aAu + bAv = a0 + b0 = 0 + 0 = 0$$

Thus, $au + bv$ belongs to W, because it is a solution of $AX = 0$. Accordingly, W is a subspace of K^n.
We state the above result formally.

THEOREM 4.4: The solution set W of a homogeneous system $AX = 0$ in n unknowns is a subspace of K^n.

We emphasize that the solution set of a nonhomogeneous system $AX = B$ is not a subspace of K^n. In fact, the zero vector 0 does not belong to its solution set.

4.6 Linear Spans, Row Space of a Matrix

Suppose $u_1, u_2, \dots, u_m$ are any vectors in a vector space V. Recall (Section 4.4) that any vector of the form $a_1u_1 + a_2u_2 + \cdots + a_mu_m$, where the a_i are scalars, is called a *linear combination* of $u_1, u_2, \dots, u_m$. The collection of all such linear combinations, denoted by

$$\text{span}(u_1, u_2, \dots, u_m) \qquad \text{or} \qquad \text{span}(u_i)$$

is called the *linear span* of $u_1, u_2, \dots, u_m$.

Clearly the zero vector 0 belongs to $\text{span}(u_i)$, because

$$0 = 0u_1 + 0u_2 + \cdots + 0u_m$$

Furthermore, suppose v and v' belong to $\text{span}(u_i)$, say,

$$v = a_1u_1 + a_2u_2 + \cdots + a_mu_m \qquad \text{and} \qquad v' = b_1u_1 + b_2u_2 + \cdots + b_mu_m$$

Then,

$$v + v' = (a_1 + b_1)u_1 + (a_2 + b_2)u_2 + \cdots + (a_m + b_m)u_m$$

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

$$kv = ka_1u_1 + ka_2u_2 + \cdots + ka_mu_m$$

Thus, $v + v'$ and kv also belong to $\text{span}(u_i)$. Accordingly, $\text{span}(u_i)$ is a subspace of V.
More generally, for any subset S of V, $\text{span}(S)$ consists of all linear combinations of vectors in S or, when $S = \phi$, $\text{span}(S) = \{0\}$. Thus, in particular, S is a spanning set (Section 4.4) of $\text{span}(S)$.
The following theorem, which was partially proved above, holds.

THEOREM 4.5: Let S be a subset of a vector space V.

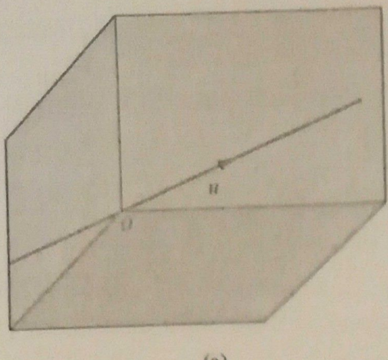
(i) Then $\text{span}(S)$ is a subspace of V that contains S.

(ii) If W is a subspace of V containing S, then $\text{span}(S) \subseteq W$.

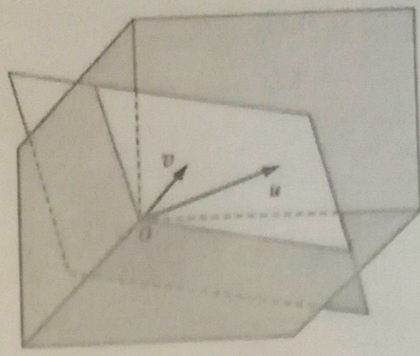
Condition (ii) in theorem 4.5 may be interpreted as saying that $\text{span}(S)$ is the "smallest" subspace of V containing S.

EXAMPLE 4.8 Consider the vector space $V = \mathbf{R}^3$.

(a) Let u be any nonzero vector in $\mathbf{R}^3$. Then $\text{span}(u)$ consists of all scalar multiples of u. Geometrically, $\text{span}(u)$ is the line through the origin O and the endpoint of u, as shown in Fig. 4-2(a).



(a)



(b)

Figure 4-2

(b) Let u and v be vectors in $\mathbf{R}^3$ that are not multiples of each other. Then $\text{span}(u, v)$ is the plane through the origin O and the endpoints of u and v as shown in Fig. 4-2(b).

(c) Consider the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ in $\mathbf{R}^3$. Recall [Example 4.1(a)] that every vector in $\mathbf{R}^3$ is a linear combination of e_1, e_2, e_3. That is, e_1, e_2, e_3 form a spanning set of $\mathbf{R}^3$. Accordingly, $\text{span}(e_1, e_2, e_3) = \mathbf{R}^3$.

Row Space of a Matrix

Let $A = [a_{ij}]$ be an arbitrary $m \times n$ matrix over a field K. The rows of A,

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \qquad R_2 = (a_{21}, a_{22}, \ldots, a_{2n}), \qquad \ldots, \qquad R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

may be viewed as vectors in K^n; hence, they span a subspace of K^n called the *row space* of A and denoted by rowsp(A). That is,

$$\text{rowsp}(A) = \text{span}(R_1, R_2, \ldots, R_m)$$

Analagously, the columns of A may be viewed as vectors in K^m called the *column space* of A and denoted by colsp(A). Observe that $\text{colsp}(A) = \text{rowsp}(A^T)$.

Recall that matrices A and B are row equivalent, written $A \sim B$, if B can be obtained from A by a sequence of elementary row operations. Now suppose M is the matrix obtained by applying one of the following elementary row operations on a matrix A:

(1) Interchange R_i and R_j, (2) Replace R_i by kR_i, (3) Replace R_j by $kR_i + R_j$

Then each row of M is a row of A or a linear combination of rows of A. Hence, the row space of M is contained in the row space of A. On the other hand, we can apply the inverse elementary row operation on M to obtain A; hence, the row space of A is contained in the row space of M. Accordingly, A and M have the same row space. This will be true each time we apply an elementary row operation. Thus, we have proved the following theorem.

THEOREM 4.6: Row equivalent matrices have the same row space.

We are now able to prove (Problems 4.45–4.47) basic results on row equivalence (which first appeared as Theorems 3.7 and 3.8 in Chapter 3).

THEOREM 4.7: Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are row equivalent echelon matrices with respective pivot entries

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r} \qquad \text{and} \qquad b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$$

Then A and B have the same number of nonzero rows—that is, $r = s$—and their pivot entries are in the same positions—that is, $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$.

THEOREM 4.8: Suppose A and B are row canonical matrices. Then A and B have the same row space if and only if they have the same nonzero rows.

COROLLARY 4.9: Every matrix A is row equivalent to a unique matrix in row canonical form.

We apply the above results in the next example.

EXAMPLE 4.9 Consider the following two sets of vectors in $\mathbf{R}^4$:

$$u_1 = (1,2,-1,3), \qquad u_2 = (2,4,1,-2), \qquad u_3 = (3,6,3,-7)$$
$$w_1 = (1,2,-4,11), \qquad w_2 = (2,4,-5,14)$$

Let $U = \text{span}(u_i)$ and $W = \text{span}(w_i)$. There are two ways to show that $U = W$.

(a) Show that each u_i is a linear combination of w_1 and w_2, and show that each w_i is a linear combination of u_1, u_2, u_3. Observe that we have to show that six systems of linear equations are consistent.

(b) Form the matrix A whose rows are u_1, u_2, u_3 and row reduce A to row canonical form, and form the matrix B whose rows are w_1 and w_2 and row reduce B to row canonical form:

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{bmatrix}$$

Because the nonzero rows of the matrices in row canonical form are identical, the row spaces of A and B are equal. Therefore, $U = W$.

Clearly, the method in (b) is more efficient than the method in (a).

4.7 Linear Dependence and Independence

Let V be a vector space over a field K. The following defines the notion of linear dependence and independence of vectors over K. (One usually suppresses mentioning K when the field is understood.) This concept plays an essential role in the theory of linear algebra and in mathematics in general.

DEFINITION: We say that the vectors $v_1, v_2, \ldots, v_m$ in V are *linearly dependent* if there exist scalars $a_1, a_2, \ldots, a_m$ in K, not all of them 0, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0$$

Otherwise, we say that the vectors are *linearly independent*.

The above definition may be restated as follows. Consider the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m = 0 \qquad (*)$$

where the x's are unknown scalars. This equation always has the *zero solution* $x_1 = 0$, $x_2 = 0, \ldots, x_m = 0$. Suppose this is the only solution; that is, suppose we can show:

$$x_1 v_1 + x_2 v_2 + \cdots + x_m v_m = 0 \quad \text{implies} \quad x_1 = 0, \quad x_2 = 0, \quad \ldots, \quad x_m = 0$$

Then the vectors $v_1, v_2, \ldots, v_m$ are linearly independent, On the other hand, suppose the equation (*) has a nonzero solution; then the vectors are linearly dependent.

A set $S = \{v_1, v_2, \ldots, v_m\}$ of vectors in V is linearly dependent or independent according to whether the vectors $v_1, v_2, \ldots, v_m$ are linearly dependent or independent.

An infinite set S of vectors is linearly dependent or independent according to whether there do or do not exist vectors $v_1, v_2, \ldots, v_k$ in S that are linearly dependent.

Warning: The set $S = \{v_1, v_2, \ldots, v_m\}$ above represents a *list* or, in other words, a finite sequence of vectors where the vectors are ordered and repetition is permitted.

The following remarks follow directly from the above definition.

Remark 1: Suppose 0 is one of the vectors $v_1, v_2, \ldots, v_m$, say $v_1 = 0$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$:

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

Remark 2: Suppose v is a nonzero vector. Then v, by itself, is linearly independent, because

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

Remark 3: Suppose two of the vectors $v_1, v_2, \ldots, v_m$ are equal or one is a scalar multiple of the other, say $v_1 = kv_2$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$:

$$v_1 - kv_2 + 0v_3 + \cdots + 0v_m = 0$$

Remark 4: Two vectors v_1 and v_2 are linearly dependent if and only if one of them is a multiple of the other.

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

Remark 6: If a set S of vectors is linearly independent, then any subset of S is linearly independent. Alternatively, if S contains a linearly dependent subset, then S is linearly dependent.

EXAMPLE 4.10

(a) Let $u = (1,1,0)$, $v = (1,3,2)$, $w = (4,9,5)$. Then u, v, w are linearly dependent, because

$$3u + 5v - 2w = 3(1,1,0) + 5(1,3,2) - 2(4,9,5) = (0,0,0) = 0$$

(b) We show that the vectors $u = (1,2,3)$, $v = (2,5,7)$, $w = (1,3,5)$ are linearly independent. We form the vector equation $xu + yv + zw = 0$, where x, y, z are unknown scalars. This yields

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

Back-substitution yields $x = 0$, $y = 0$, $z = 0$. We have shown that

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

Accordingly, u, v, w are linearly independent.

(c) Let V be the vector space of functions from **R** into **R**. We show that the functions $f(t) = \sin t$, $g(t) = e^t$, $h(t) = t^2$ are linearly independent. We form the vector (function) equation $xf + yg + zh = 0$, where x, y, z are unknown scalars. This function equation means that, for every value of t,

$$x \sin t + ye^t + zt^2 = 0$$

Thus, in this equation, we choose appropriate values of t to easily get $x = 0$, $y = 0$, $z = 0$. For example,

(i)	Substitute $t = 0$	to obtain $x(0) + y(1) + z(0) = 0$	or	$y = 0$
(ii)	Substitute $t = \pi$	to obtain $x(0) + 0(e^{\pi}) + z(\pi^2) = 0$	or	$z = 0$
(iii)	Substitute $t = \pi/2$	to obtain $x(1) + 0(e^{\pi/2}) + 0(\pi^2/4) = 0$	or	$x = 0$

We have shown

$$xf + yg + zf = 0 \qquad \text{implies} \qquad x = 0, \quad y = 0, \quad z = 0$$

Accordingly, u, v, w are linearly independent.

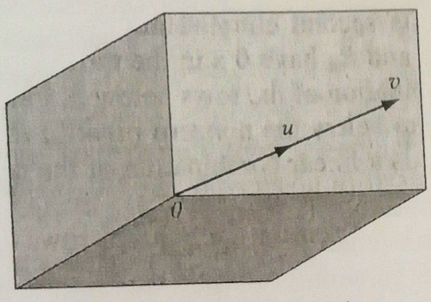
Linear Dependence in $\mathbf{R}^3$

Linear dependence in the vector space $V = \mathbf{R}^3$ can be described geometrically as follows:

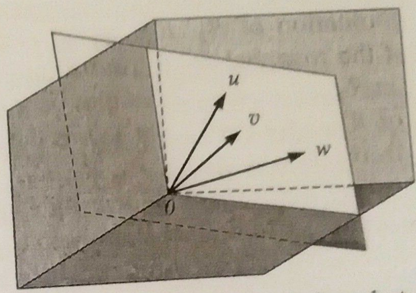
(a) Any two vectors u and v in $\mathbf{R}^3$ are linearly dependent if and only if they lie on the same line through the origin O, as shown in Fig. 4-3(a).

(b) Any three vectors u, v, w in $\mathbf{R}^3$ are linearly dependent if and only if they lie on the same plane through the origin O, as shown in Fig. 4-3(b).

Later, we will be able to show that any four or more vectors in $\mathbf{R}^3$ are automatically linearly dependent.



(a) u and v are linearly dependent.



(b) u, v, and w are linearly dependent.

Figure 4-3

Linear Dependence and Linear Combinations

The notions of linear dependence and linear combinations are closely related. Specifically, for more than one vector, we show that the vectors $v_1, v_2, \ldots, v_m$ are linearly dependent if and only if one of them is a linear combination of the others.

Suppose, say, v_i is a linear combination of the others,

$$v_i = a_1 v_1 + \cdots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \cdots + a_m v_m$$

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

$$a_1 v_1 + \cdots + a_{i-1} v_{i-1} - v_i + a_{i+1} v_{i+1} + \cdots + a_m v_m = 0$$

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

$$b_1 v_1 + \cdots + b_j v_j + \cdots + b_m v_m = 0, \qquad \text{where} \qquad b_j \neq 0$$

Then we can solve for v_j obtaining

$$v_j = b_j^{-1} b_1 v_1 - \cdots - b_j^{-1} b_{j-1} v_{j-1} - b_j^{-1} b_{j+1} v_{j+1} - \cdots - b_j^{-1} b_m v_m$$

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

We now state a slightly stronger statement than the one above. This result has many important consequences.

LEMMA 4.10: Suppose two or more nonzero vectors $v_1, v_2, \ldots, v_m$ are linearly dependent. Then one of the vectors is a linear combination of the preceding vectors; that is, there exists $k > 1$ such that

$$v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1}$$

Linear Dependence and Echelon Matrices

Consider the following echelon matrix A, whose pivots have been circled:

$$A = \begin{bmatrix} 0 & \textcircled{2} & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & \textcircled{4} & 3 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & \textcircled{7} & 8 & 9 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{6} & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe that the rows R_2, R_3, R_4 have 0's in the second column below the nonzero pivot in R_1, and hence any linear combination of R_2, R_3, R_4 must have 0 as its second entry. Thus, R_1 cannot be a linear combination of the rows below it. Similarly, the rows R_3 and R_4 have 0's in the third column below the nonzero pivot in R_2, and hence R_2 cannot be a linear combination of the rows below it. Finally, R_3 cannot be a multiple of R_4, because R_4 has a 0 in the fifth column below the nonzero pivot in R_3. Viewing the nonzero rows from the bottom up, R_4, R_3, R_2, R_1, no row is a linear combination of the preceding rows. Thus, the rows are linearly independent by Lemma 4.10.

The argument used with the above echelon matrix A can be used for the nonzero rows of any echelon matrix. Thus, we have the following very useful result.

THEOREM 4.11: The nonzero rows of a matrix in echelon form are linearly independent.

4.8 Basis and Dimension

First we state two equivalent ways to define a basis of a vector space V. (The equivalence is proved in Problem 4.28.)

DEFINITION A: A set $S = \{u_1, u_2, \ldots, u_n\}$ of vectors is a *basis* of V if it has the following two properties: (1) S is linearly independent. (2) S spans V.

DEFINITION B: A set $S = \{u_1, u_2, \ldots, u_n\}$ of vectors is a *basis* of V if every $v \in V$ can be written uniquely as a linear combination of the basis vectors.

The following is a fundamental result in linear algebra.

THEOREM 4.12: Let V be a vector space such that one basis has m elements and another basis has n elements. Then $m = n$.

A vector space V is said to be of *finite dimension* n or *n-dimensional*, written

$$\dim V = n$$

if V has a basis with n elements. Theorem 4.12 tells us that all bases of V have the same number of elements, so this definition is well defined.

The vector space $\{0\}$ is defined to have dimension 0.

Suppose a vector space V does not have a finite basis. Then V is said to be of *infinite dimension* or to be *infinite-dimensional*.

The above fundamental Theorem 4.12 is a consequence of the following "replacement lemma" (proved in Problem 4.35).

LEMMA 4.13: Suppose $\{v_1, v_2, \ldots, v_n\}$ spans V, and suppose $\{w_1, w_2, \ldots, w_m\}$ is linearly independent. Then $m \le n$, and V is spanned by a set of the form

$$\{w_1, w_2, \ldots, w_m, \; v_{i_1}, v_{i_2}, \ldots, v_{i_{n-m}}\}$$

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

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

Examples of Bases

This subsection presents important examples of bases of some of the main vector spaces appearing in this text.

(a) Vector space K^n: Consider the following n vectors in K^n:

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

These vectors are linearly independent. (For example, they form a matrix in echelon form.) Furthermore, any vector $u = (a_1, a_2, \ldots, a_n)$ in K^n can be written as a linear combination of the above vectors. Specifically,

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n$$

Accordingly, the vectors form a basis of K^n called the *usual* or *standard* basis of K^n. Thus (as one might expect), K^n has dimension n. In particular, any other basis of K^n has n elements.

(b) Vector space $\mathbf{M} = \mathbf{M}_{r,s}$ of all $r \times s$ matrices: The following six matrices form a basis of the vector space $\mathbf{M}_{2,3}$ of all 2×3 matrices over K:

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

More generally, in the vector space $\mathbf{M} = \mathbf{M}_{r,s}$ of all $r \times s$ matrices, let E_{ij} be the matrix with ij-entry 1 and 0's elsewhere. Then all such matrices form a basis of $\mathbf{M}_{r,s}$ called the *usual* or *standard* basis of $\mathbf{M}_{r,s}$. Accordingly, $\dim \mathbf{M}_{r,s} = rs$.

(c) Vector space $\mathbf{P}_n(t)$ of all polynomials of degree $\leq n$: The set $S = \{1, t, t^2, t^3, \ldots, t^n\}$ of $n+1$ polynomials is a basis of $\mathbf{P}_n(t)$. Specifically, any polynomial $f(t)$ of degree $\leq n$ can be expessed as a linear combination of these powers of t, and one can show that these polynomials are linearly independent. Therefore, $\dim \mathbf{P}_n(t) = n+1$.

(d) Vector space $\mathbf{P}(t)$ of all polynomials: Consider any finite set $S = \{f_1(t), f_2(t), \ldots, f_m(t)\}$ of polynomials in $\mathbf{P}(t)$, and let m denote the largest of the degrees of the polynomials. Then any polynomial $g(t)$ of degree exceeding m cannot be expressed as a linear combination of the elements of S. Thus, S cannot be a basis of $\mathbf{P}(t)$. This means that the dimension of $\mathbf{P}(t)$ is infinite. We note that the infinite set $S' = \{1, t, t^2, t^3, \ldots\}$, consisting of all the powers of t, spans $\mathbf{P}(t)$ and is linearly independent. Accordingly, S' is an infinite basis of $\mathbf{P}(t)$.

Theorems on Bases

The following three theorems (proved in Problems 4.37, 4.38, and 4.39) will be used frequently.

THEOREM 4.14: Let V be a vector space of finite dimension n. Then:

(i) Any $n+1$ or more vectors in V are linearly dependent.

(ii) Any linearly independent set $S = \{u_1, u_2, \ldots, u_n\}$ with n elements is a basis of V.

(iii) Any spanning set $T = \{v_1, v_2, \ldots, v_n\}$ of V with n elements is a basis of V.

THEOREM 4.15: Suppose S spans a vector space V. Then:

(i) Any maximum number of linearly independent vectors in S form a basis of V.

(ii) Suppose one deletes from S every vector that is a linear combination of preceding vectors in S. Then the remaining vectors form a basis of V.

THEOREM 4.16: Let V be a vector space of finite dimension and let $S = \{u_1, u_2, \ldots, u_r\}$ be a set of linearly independent vectors in V. Then S is part of a basis of V; that is, S may be extended to a basis of V.

EXAMPLE 4.11

(a) The following four vectors in $\mathbf{R}^4$ form a matrix in echelon form:

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

Thus, the vectors are linearly independent, and, because $\dim \mathbf{R}^4 = 4$, the four vectors form a basis of $\mathbf{R}^4$.

(b) The following $n+1$ polynomials in $\mathbf{P}_n(t)$ are of increasing degree:

$$1,\ t-1,\ (t-1)^2,\ \ldots,\ (t-1)^n$$

Therefore, no polynomial is a linear combination of preceding polynomials; hence, the polynomials are linear independent. Furthermore, they form a basis of $\mathbf{P}_n(t)$, because $\dim \mathbf{P}_n(t) = n+1$.

(c) Consider any four vectors in $\mathbf{R}^3$, say

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

By Theorem 4.14(i), the four vectors must be linearly dependent, because they come from the three-dimensional vector space $\mathbf{R}^3$.

Dimension and Subspaces

The following theorem (proved in Problem 4.40) gives the basic relationship between the dimension of a vector space and the dimension of a subspace.

THEOREM 4.17: 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 4.12 Let W be a subspace of the real space $\mathbf{R}^3$. Note that $\dim \mathbf{R}^3 = 3$. Theorem 4.17 tells us that the dimension of W can only be 0, 1, 2, or 3. The following cases apply:

(a) If $\dim W = 0$, then $W = \{0\}$, a point.

(b) If $\dim W = 1$, then W is a line through the origin 0.

(c) If $\dim W = 2$, then W is a plane through the origin 0.

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

4.9 Application to Matrices, Rank of a Matrix

Let A be any $m \times n$ matrix over a field K. Recall that the rows of A may be viewed as vectors in K^n and that the row space of A, written rowsp(A), is the subspace of K^n spanned by the rows of A. The following definition applies.

DEFINITION: The *rank* of a matrix A, written rank(A), is equal to the maximum number of linearly independent rows of A or, equivalently, the dimension of the row space of A.

Recall, on the other hand, that the columns of an $m \times n$ matrix A may be viewed as vectors in K^m and that the column space of A, written colsp(A), is the subspace of K^m spanned by the columns of A. Although m may not be equal to n—that is, the rows and columns of A may belong to different vector spaces—we have the following fundamental result.

THEOREM 4.18: The maximum number of linearly independent rows of any matrix A is equal to the maximum number of linearly independent columns of A. Thus, the dimension of the row space of A is equal to the dimension of the column space of A.

Accordingly, one could restate the above definition of the rank of A using columns instead of rows.

Basis-Finding Problems

This subsection shows how an echelon form of any matrix A gives us the solution to certain problems about A itself. Specifically, let A and B be the following matrices, where the echelon matrix B (whose pivots are circled) is an echelon form of A:

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 5 & 5 & 6 & 4 & 5 \\ 3 & 7 & 6 & 11 & 6 & 9 \\ 1 & 5 & 10 & 8 & 9 & 9 \\ 2 & 6 & 8 & 11 & 9 & 12 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \textcircled{1} & 2 & 1 & 3 & 1 & 2 \\ 0 & \textcircled{1} & 3 & 1 & 2 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We solve the following four problems about the matrix A, where $C_1, C_2, \ldots, C_6$ denote its columns:

(a) Find a basis of the row space of A.

(b) Find each column C_k of A that is a linear combination of preceding columns of A.

(c) Find a basis of the column space of A.

(d) Find the rank of A.

(a) We are given that A and B are row equivalent, so they have the same row space. Moreover, B is in echelon form, so its nonzero rows are linearly independent and hence form a basis of the row space of B. Thus, they also form a basis of the row space of A. That is,

$$\text{basis of rowsp}(A): \quad (1,2,1,3,1,2), \quad (0,1,3,1,2,1), \quad (0,0,0,1,1,2)$$

(b) Let $M_k = [C_1, C_2, \ldots, C_k]$, the submatrix of A consisting of the first k columns of A. Then M_{k-1} and M_k are, respectively, the coefficient matrix and augmented matrix of the vector equation

$$x_1C_1 + x_2C_2 + \cdots + x_{k-1}C_{k-1} = C_k$$

Theorem 3.9 tells us that the system has a solution, or, equivalently, C_k is a linear combination of the preceding columns of A if and only if $\text{rank}(M_k) = \text{rank}(M_{k-1})$, where $\text{rank}(M_k)$ means the number of pivots in an echelon form of M_k. Now the first k column of the echelon matrix B is also an echelon form of M_k. Accordingly,

$$\text{rank}(M_2) = \text{rank}(M_3) = 2 \qquad \text{and} \qquad \text{rank}(M_4) = \text{rank}(M_5) = \text{rank}(M_6) = 3$$

Thus, C_3, C_5, C_6 are each a linear combination of the preceding columns of A.

(c) The fact that the remaining columns C_1, C_2, C_4 are not linear combinations of their respective preceding columns also tells us that they are linearly independent. Thus, they form a basis of the column space of A. That is,

$$\text{basis of colsp}(A): \quad [1,2,3,1,2]^T, \quad [2,5,7,5,6]^T, \quad [3,6,11,8,11]^T$$

Observe that C_1, C_2, C_4 may also be characterized as those columns of A that contain the pivots in any echelon form of A.

(d) Here we see that three possible definitions of the rank of A yield the same value.

(i) There are three pivots in B, which is an echelon form of A.

(ii) The three pivots in B correspond to the nonzero rows of B, which form a basis of the row space of A.

(iii) The three pivots in B correspond to the columns of A, which form a basis of the column space of A.

Thus, $\text{rank}(A) = 3$.

Application to Finding a Basis for $W = \text{span}(u_1, u_2, \ldots, u_r)$

Frequently, we are given a list $S = \{u_1, u_2, \ldots, u_r\}$ of vectors in K^n and we want to find a basis for the subspace W of K^n spanned by the given vectors—that is, a basis of

$$W = \text{span}(S) = \text{span}(u_1, u_2, \ldots, u_r)$$

The following two algorithms, which are essentially described in the above subsection, find such a basis (and hence the dimension) of W.

Algorithm 4.1 (Row space algorithm)

Step 1. Form the matrix M whose *rows* are the given vectors.

Step 2. Row reduce M to echelon form.

Step 3. Output the nonzero rows of the echelon matrix.

Sometimes we want to find a basis that only comes from the original given vectors. The next algorithm accomplishes this task.

Algorithm 4.2 (Casting-out algorithm)

Step 1. Form the matrix M whose *columns* are the given vectors.

Step 2. Row reduce M to echelon form.

Step 3. For each column C_k in the echelon matrix without a pivot, delete (cast out) the vector u_k from the list S of given vectors.

Step 4. Output the remaining vectors in S (which correspond to columns with pivots).

We emphasize that in the first algorithm we form a matrix whose rows are the given vectors, whereas in the second algorithm we form a matrix whose columns are the given vectors.

EXAMPLE 4.13 Let W be the subspace of $\mathbf{R}^5$ spanned by the following vectors:

$$u_1 = (1,2,1,3,2), \qquad u_2 = (1,3,3,5,3), \qquad u_3 = (3,8,7,13,8)$$
$$u_4 = (1,4,6,9,7), \qquad u_5 = (5,13,13,25,19)$$

Find a basis of W consisting of the original given vectors, and find dim W.

Form the matrix M whose columns are the given vectors, and reduce M to echelon form:

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 2 & 3 & 8 & 4 & 13 \\ 1 & 3 & 7 & 6 & 13 \\ 3 & 5 & 13 & 9 & 25 \\ 2 & 3 & 8 & 7 & 19 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The pivots in the echelon matrix appear in columns C_1, C_2, C_4. Accordingly, we "cast out" the vectors u_3 and u_5 from the original five vectors. The remaining vectors u_1, u_2, u_4, which correspond to the columns in the echelon matrix with pivots, form a basis of W. Thus, in particular, dim $W = 3$.

Remark: The justification of the casting-out algorithm is essentially described above, but we repeat it again here for emphasis. The fact that column C_3 in the echelon matrix in Example 4.13 does not have a pivot means that the vector equation

$$xu_1 + yu_2 = u_3$$

has a solution, and hence u_3 is a linear combination of u_1 and u_2. Similarly, the fact that C_5 does not have a pivot means that u_5 is a linear combination of the preceding vectors. We have deleted each vector in the original spanning set that is a linear combination of preceding vectors. Thus, the remaining vectors are linearly independent and form a basis of W.

EXAMPLE 4.15 Consider the vector space $V = \mathbf{R}^3$.

(a) Let U be the xy-plane and let W be the yz-plane; that is,

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

Then $\mathbf{R}^3 = U + W$, because 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, because such sums are not unique. For example,

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

(b) Let U be the xy-plane and let W be the z-axis; that is,

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

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, $\mathbf{R}^3 = U \oplus W$.

General Direct Sums

The notion of a direct sum is extended to more than one factor in the obvious way. That is, V is the *direct sum* of subspaces $W_1, W_2, \ldots, W_r$, written

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

if every vector $v \in V$ can be written in one and only one way as

$$v = w_1 + w_2 + \cdots + w_r$$

where $w_1 \in W_1, w_2 \in W_2, \ldots, w_r \in W_r$.

The following theorems hold.

THEOREM 4.22: Suppose $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$. Also, for each k, suppose S_k is a linearly independent subset of W_k. Then

(a) The union $S = \bigcup_k S_k$ is linearly independent in V.

(b) If each S_k is a basis of W_k, then $\bigcup_k S_k$ is a basis of V.

(c) $\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_r$.

THEOREM 4.23: Suppose $V = W_1 + W_2 + \cdots + W_r$ and $\dim V = \sum_k \dim W_k$. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$.

4.11 Coordinates

Let V be an n-dimensional vector space over K with basis $S = \{u_1, u_2, \ldots, u_n\}$. Then any vector $v \in V$ can be expressed uniquely as a linear combination of the basis vectors in S, say

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

These n scalars $a_1, a_2, \ldots, a_n$ are called the *coordinates* of v relative to the basis S, and they form a vector $[a_1, a_2, \ldots, a_n]$ in K^n called the *coordinate vector* of v relative to S. We denote this vector by $[v]_S$, or simply $[v]$, when S is understood. Thus,

$$[v]_S = [a_1, a_2, \ldots, a_n]$$

For notational convenience, brackets $[\ldots]$, rather than parentheses $(\ldots)$, are used to denote the coordinate vector.