Vectors and Matrices in Engineering and Mathematics; *n*-Space

Linear algebra is concerned with two basic kinds of quantities: "vectors" and "matrices." The term "vector" has various meanings in engineering, science, and mathematics, some of which will be discussed in this section. We will begin by reviewing the geometric notion of a vector as it is used in basic physics and engineering, next we will discuss vectors in two-dimensional and three-dimensional coordinate systems, and then we will consider how the notion of a vector can be extended to higher-dimensional spaces. Finally, we will talk a little about matrices, explaining how they arise and how they are related to vectors.

source: Contemporary Linear Algebra, H. Anton / R.C. Busby

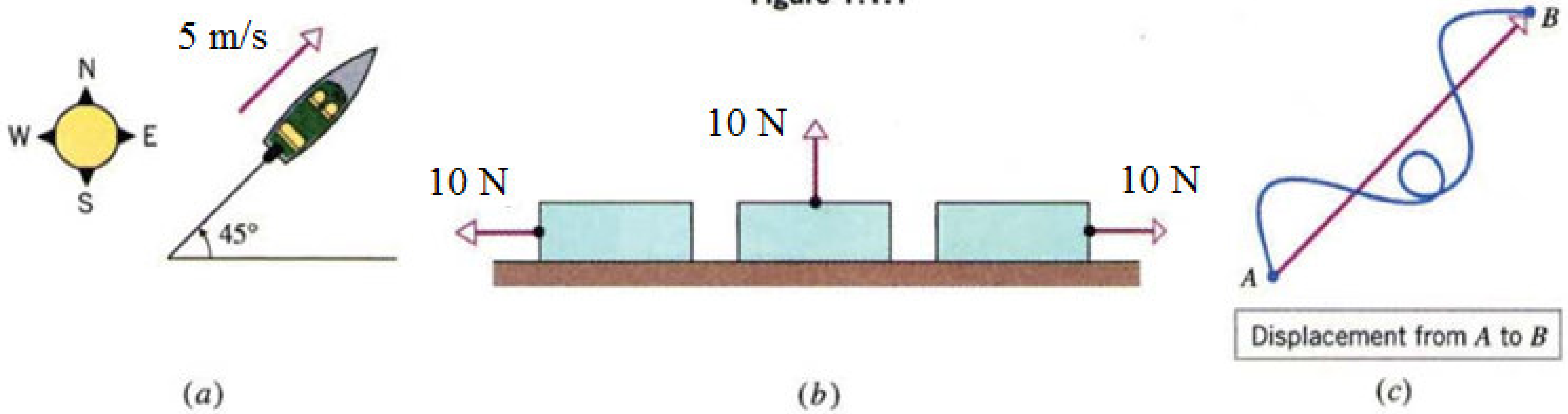
Copyright © 2003 John Wiley & Sons, Inc.

SCALARS AND VECTORS

Engineers and physicists distinguish between two types of physical quantities—***scalars***, which are quantities that can be described by a numerical value alone, and ***vectors***, which require both a numerical value and a direction for their complete description.

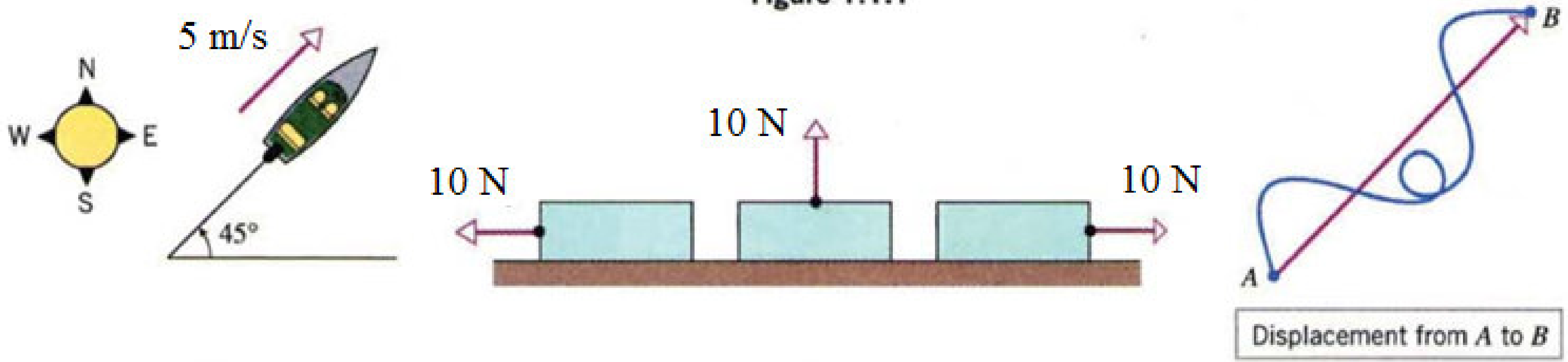
For example, temperature, length, and speed are scalars because they are completely described by a number that tells "how much"—say a temperature of 20°C, a length of 5 cm, or a speed of 10 m/s. In contrast, velocity, force, and displacement are vectors because they involve a direction as well as a numerical value.

Figure 1.1.1



Vectors in two dimensions (2-space) or three dimensions (3-space) can be represented geometrically by *arrows*—the length of the arrow is proportional to the ***magnitude*** (or numerical part) of the vector, and the direction of the arrow indicates the direction of the vector. The *tail* of the arrow is called the ***initial point*** and the *tip* is called the ***terminal point***.

Figure 1.1.1

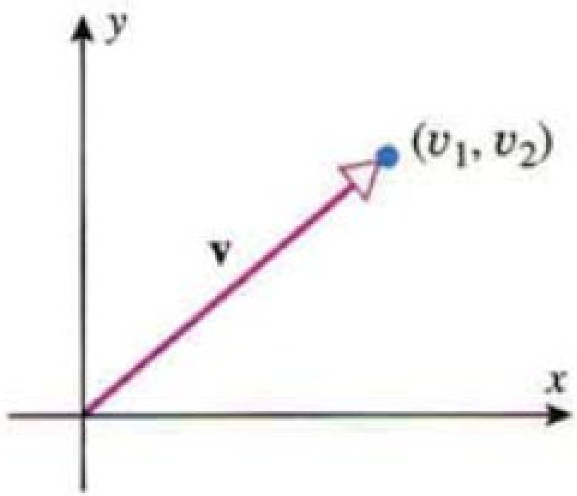


(*a*) (*b*) (*c*)

VECTORS IN COORDINATE SYSTEMS

If a vector **v** in 2-space or 3-space is positioned with its initial point at the origin of a rectangular coordinate system, then the vector is completely determined by the coordinates of its terminal point, and we call these coordinates the ***components*** of **v** relative to the coordinate system.

We will write $\mathbf{v} = (v_1, v_2)$ for the vector **v** in 2-space with components (v_1, v_2) and $\mathbf{v} = (v_1, v_2, v_3)$ for the vector **v** in 3-space with components (v_1, v_2, v_3) (Figure 1.1.13).



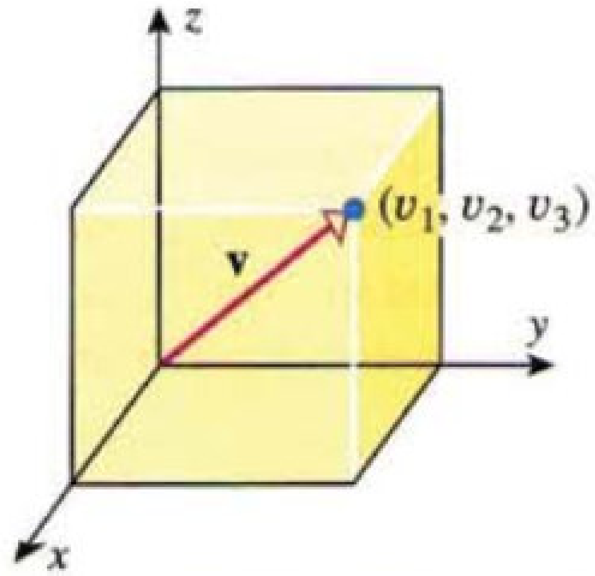


Figure 1.1.13

Algebraically, vectors in 2-space can now be viewed as ordered pairs of real numbers and vectors in 3-space as ordered triples of real numbers. Thus, we will denote the set of all vectors in 2-space by R^2 and the set of all vectors in 3-space by R^3 (the "R" standing for the word "real").

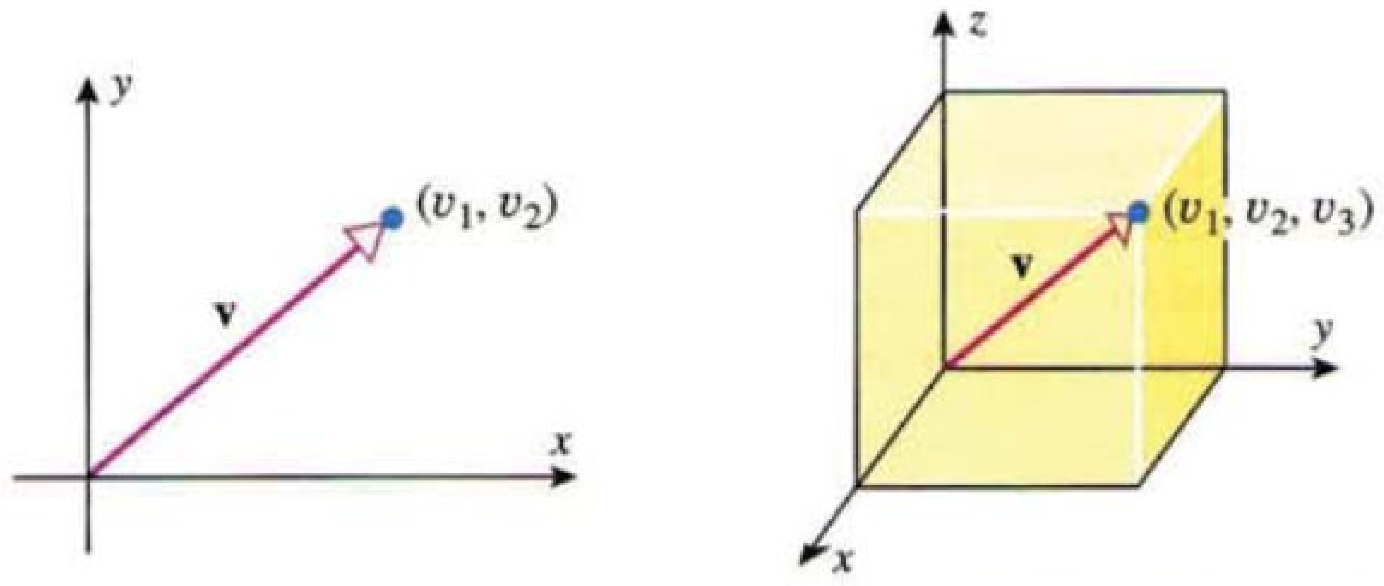


Figure 1.1.13

VECTORS IN R^n

Experimental Data—A scientist performs an experiment and makes n numerical measurements each time the experiment is performed. The result of each experiment can be regarded as a vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in R^n in which $y_1, y_2, \ldots, y_n$ are the measured values.

Graphical Images—One way in which color images are created on computer screens is by assigning each pixel (an addressable point on the screen) three numbers that describe the ***hue***, ***saturation***, and ***brightness*** of the pixel. Thus, a complete color image can be viewed as a set of 5-tuples of the form $\mathbf{v} = (x, y, h, s, b)$ in which x and y are the screen coordinates of a pixel and h, s, and b are its hue, saturation, and brightness.

Mechanical Systems—Suppose that six particles move along the same coordinate line so that at time t their coordinates are $x_1, x_2, \ldots, x_6$ and their velocities are $v_1, v_2, \ldots, v_6$, respectively. This information can be represented by the vector

$$\mathbf{v} = (x_1, x_2, x_3, x_4, x_5, x_6, v_1, v_2, v_3, v_4, v_5, v_6, t)$$

in R^{13}. This vector is called the ***state*** of the particle system at time t.

→ Vectors, as above, are quantities composed of an ordered sequence of real numbers.

VECTORS IN R^n

Definition 1.1.2 If n is a positive integer, then an ***ordered n-tuple*** is a sequence of n real numbers $(v_1, v_2, \ldots, v_n)$. The set of all ordered n-tuples is called ***n-space*** and is denoted by R^n.

REMARK You can think of the numbers in an n-tuple $(v_1, v_2, \ldots, v_n)$ as either the coordinates of a *generalized point* or the components of a *generalized vector*, depending on the geometric image you want to bring to mind—the choice makes no difference mathematically, since it is the algebraic properties of n-tuples that are of concern.

We will denote n-tuples using the vector notation $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, and we will write $\mathbf{0} = (0, 0, \ldots, 0)$ for the n-tuple whose components are all zero. We will call this the ***zero vector*** or sometimes the ***origin*** of R^n.

We will sometimes refer to R^1, R^2, and R^3 as ***visible space*** and R^4, R^5, … as ***higher-dimensional spaces***.

Definition 1.1.3 Vectors $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ in R^n are said to be ***equivalent*** (also called ***equal***) if

$$v_1 = w_1, \quad v_2 = w_2, \ldots, \quad v_n = w_n$$

We indicate this by writing $\mathbf{v} = \mathbf{w}$.

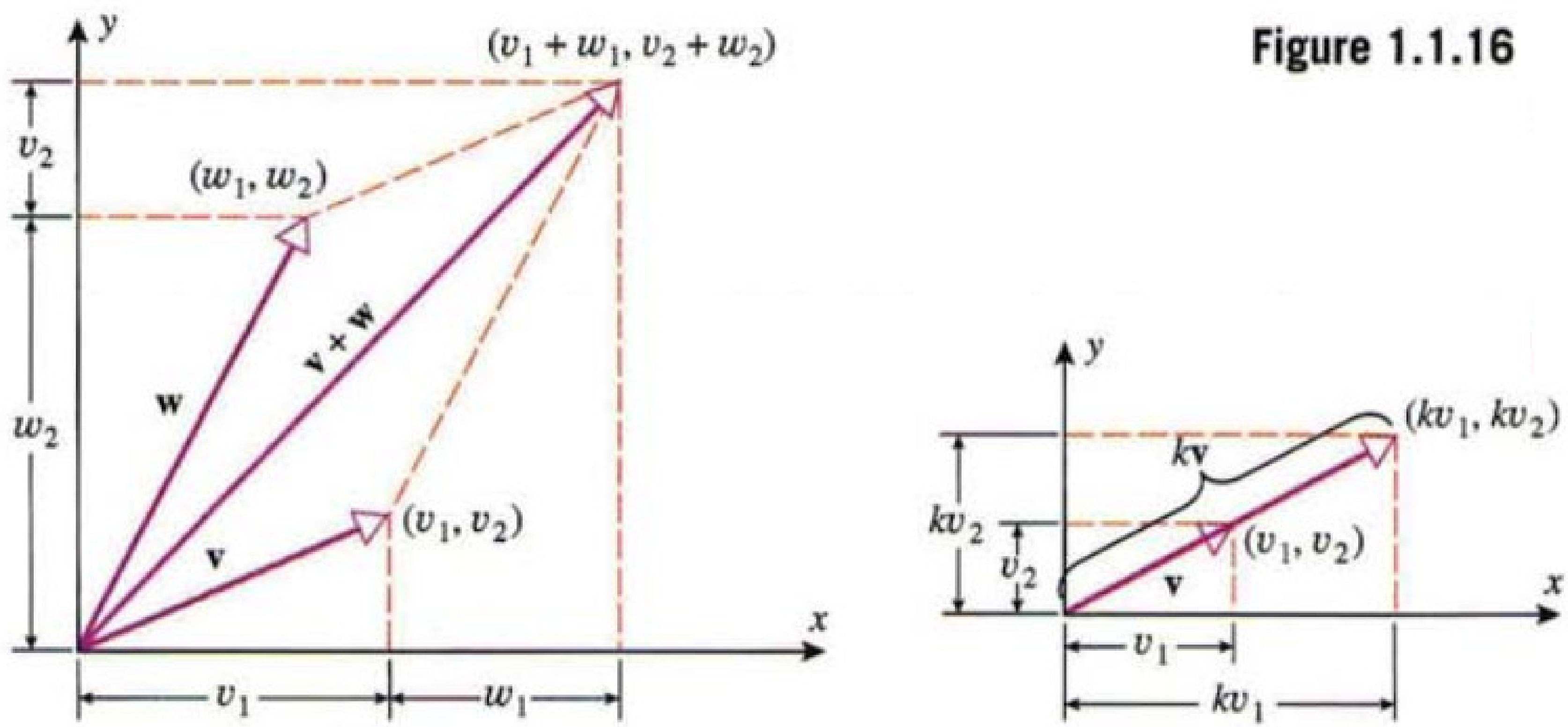
Definition 1.1.4 If $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ are vectors in R^n, and if k is any scalar, then we define

$$\mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \tag{10}$$

$$k\mathbf{v} = (kv_1, kv_2, \ldots, kv_n) \tag{11}$$

$$-\mathbf{v} = (-v_1, -v_2, \ldots, -v_n) \tag{12}$$

$$\mathbf{w} - \mathbf{v} = \mathbf{w} + (-\mathbf{v}) = (w_1 - v_1, w_2 - v_2, \ldots, w_n - v_n) \tag{13}$$



y
x
$(v_1 + w_1, v_2 + w_2)$
(w_1, w_2)
(v_1, v_2)
w
v + w
v
v_2
w_2
v_1
w_1
y
x
(kv_1, kv_2)
$k\mathbf{v}$
(v_1, v_2)
v
kv_2
v_2
v_1
kv_1

Figure 1.1.16

Theorem 1.1.5 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *and if* k *and* l *are scalars, then:*

(*a*) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(*b*) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

(*c*) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

(*d*) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(*e*) $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$

(*f*) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

(*g*) $k(l\mathbf{u}) = (kl)\mathbf{u}$

(*h*) $1\mathbf{u} = \mathbf{u}$

Theorem 1.1.6 *If* $\mathbf{v}$ *is a vector in* R^n *and* k *is a scalar, then:*

(*a*) $0\mathbf{v} = \mathbf{0}$

(*b*) $k\mathbf{0} = \mathbf{0}$

(*c*) $(-1)\mathbf{v} = -\mathbf{v}$

PARALLEL AND COLLINEAR VECTORS

Definition 1.1.7 Two vectors in R^n are said to be ***parallel*** or, alternatively, ***collinear*** if at least one of the vectors is a scalar multiple of the other. If one of the vectors is a positive scalar multiple of the other, then the vectors are said to have the ***same direction***, and if one of them is a negative scalar multiple of the other, then the vectors are said to have ***opposite directions***.

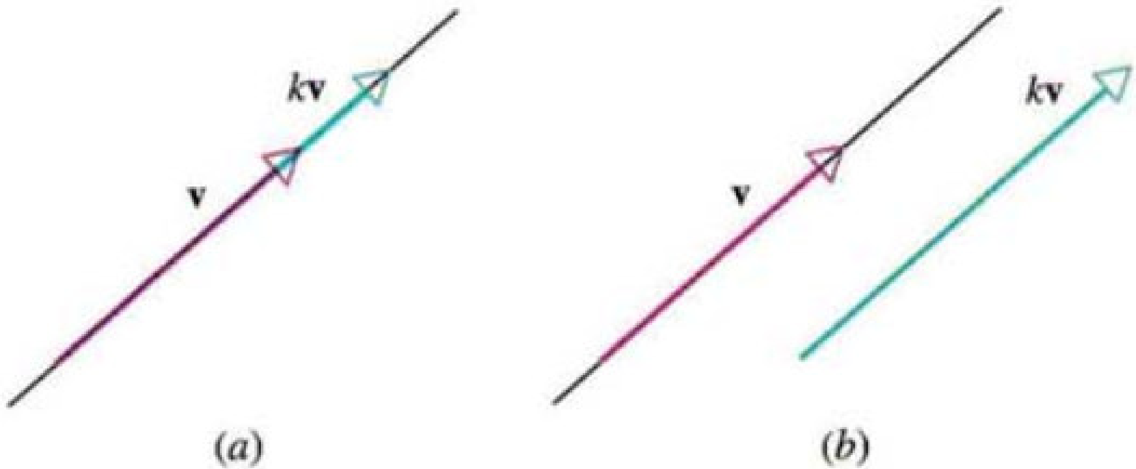


Figure 1.1.18

LINEAR COMBINATIONS

Definition 1.1.8 A vector $\mathbf{w}$ in R^n is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in R^n if $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \tag{14}$$

The scalars $c_1, c_2, \ldots, c_k$ are called the ***coefficients*** in the linear combination. In the case where $k = 1$, Formula (14) becomes $\mathbf{w} = c_1\mathbf{v}_1$, so to say that $\mathbf{w}$ is a linear combination of $\mathbf{v}_1$ is the same as saying that $\mathbf{w}$ is a scalar multiple of $\mathbf{v}_1$.

SUBSPACES OF R^n

In general, if W is a nonempty set of vectors in R^n, then we say that W is ***closed under scalar multiplication*** if any scalar multiple of a vector in W is also in W, and we say that W is ***closed under addition*** if the sum of any two vectors in W is also in W. We also make the following definition to describe sets that have these two closure properties.

Definition 3.4.1 A nonempty set of vectors in R^n is called a ***subspace*** of R^n if it is closed under scalar multiplication and addition.

Theorem 3.4.2 *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ are vectors in R^n, then the set of all linear combinations*

$$\mathbf{x} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_s\mathbf{v}_s \tag{3}$$

is a subspace of R^n.

The subspace W of R^n whose vectors satisfy (3) is called the ***span*** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ and is denoted by

$$W = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\} \tag{4}$$

We also say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ ***span*** W. The scalars in (3) are called ***parameters***,

LINEAR INDEPENDENCE

Definition 3.4.5 A nonempty set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ in R^n is said to be ***linearly independent*** if the only scalars $c_1, c_2, \ldots, c_s$ that satisfy the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0} \tag{9}$$

are $c_1 = 0, c_2 = 0, \ldots, c_s = 0$. If there are scalars, not all zero, that satisfy this equation, then the set is said to be ***linearly dependent***.

REMARK Strictly speaking, the terms "linearly dependent" and "linearly independent" apply to nonempty finite *sets* of vectors; however, we will also find it convenient to apply them to the vectors themselves.

Thus, we will say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ are linearly independent or dependent in accordance with whether the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ is linearly independent or dependent.

Also, if S is a set with infinitely many vectors, then we will say that S is linearly independent if every finite subset is linearly independent and is linearly dependent if some finite subset is linearly dependent.

ALTERNATIVE NOTATIONS FOR VECTORS

Up to now we have been writing vectors in R^n using the notation

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) \tag{15}$$

We call this the ***comma-delimited*** form. However, a vector in R^n is essentially just a list of n numbers (the components) in a definite order, so any notation that displays the components of the vector in their correct order is a valid alternative to the comma-delimited notation. For example, the vector in (15) might be written as

$$\mathbf{v} = [v_1 \quad v_2 \quad \cdots \quad v_n] \tag{16}$$

which is called ***row-vector*** form, or as

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{17}$$

which is called ***column-vector*** form. The choice of notation is often a matter of taste or convenience, but sometimes the nature of the problem under consideration will suggest a particular notation. All three notations will be used in this text.

MATRICES

we define a ***matrix*** to be a rectangular array of numbers, called the ***entries*** of the matrix.

If a matrix has m rows and n columns, then it is said to have ***size*** $m \times n$, where the number of rows is always written first. Thus, for example, the matrix in (18) has size 4×7.

$$\begin{bmatrix} 2 & 1 & 2 & 0 & 3 & 0 & 1 \\ 2 & 0 & 1 & 3 & 1 & 0 & 1 \\ 1 & 3 & 0 & 0 & 1 & 0 & 1 \\ 1 & 2 & 4 & 1 & 0 & 0 & 2 \end{bmatrix} \qquad (18)$$

A matrix with one row is called a ***row vector***, and a matrix with one column is called a ***column vector***.

You can also think of a matrix as a list of row vectors or column vectors.

source: Contemporary Linear Algebra, H. Anton / R.C. Busby

Copyright © 2003 John Wiley & Sons, Inc.

VECTOR ALGEBRA

There are essentially three different ways to introduce vector algebra: *geometrically*, *analytically*, and *axiomatically*. In the geometric approach, vectors are represented by directed line segments, or arrows. Algebraic operations on vectors, such as addition, subtraction, and multiplication by real numbers, are defined and studied by geometric methods.

In the analytic approach, vectors and vector operations are described entirely in terms of *numbers*, called *components*. Properties of the vector operations are then deduced from corresponding properties of numbers. The analytic description of vectors arises naturally from the geometric description as soon as a coordinate system is introduced.

In the axiomatic approach, no attempt is made to describe the nature of a vector or of the algebraic operations on vectors. Instead, vectors and vector operations are thought of as *undefined concepts* of which we know nothing except that they satisfy a certain set of axioms. Such an algebraic system, with appropriate axioms, is called a *linear space* or a *linear vector space*. The algebra of directed line segments and the algebra of vectors described by components are merely two examples of linear spaces.

source: Calculus (Vol. I), 2nd. ed., T.M. Apostol

Copyright ©1967 John Wiley & Sons, Inc.

The study of vector algebra from the axiomatic point of view is perhaps the most mathematically satisfactory approach to use since it furnishes a description of vectors that is free of coordinate systems and free of any particular geometric representation.

In this, we base our treatment on the analytic approach, and we also use directed line segments to interpret many of the results geometrically.

The vector space of *n*-tuples of real numbers

One can consider an *n-tuple* of real numbers

$$(a_1, a_2, \ldots, a_n)$$

for any integer $n \geq 1$. Such an n-tuple is called an *n-dimensional point* or an *n-dimensional vector*, the individual numbers $a_1, a_2, \ldots, a_n$ being referred to as *coordinates* or *components* of the vector. The collection of all n-dimensional vectors is called *the vector space of n-tuples*, or simply *n-space*. We denote this space by V_n.

To convert V_n into an algebraic system, we introduce *equality* of vectors and two vector operations called *addition* and *multiplication by scalars*. The word "scalar" is used here as a synonym for "real number."

DEFINITION. *Two vectors A and B in V_n are called equal whenever they agree in their respective components. That is, if $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$, the vector equation $A = B$ means exactly the same as the n scalar equations*

$$a_1 = b_1, \qquad a_2 = b_2, \quad \ldots, \quad a_n = b_n .$$

The sum $A + B$ is defined to be the vector obtained by adding corresponding components:

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

If c is a scalar, we define cA or Ac to be the vector obtained by multiplying each component of A by c:

$$cA = (ca_1, ca_2, \ldots, ca_n) .$$

The dot product

DEFINITION. *If* $A = (a_1, \ldots, a_n)$ *and* $B = (b_1, \ldots, b_n)$ *are two vectors in* V_n, *their dot product is denoted by* $A \cdot B$ *and is defined by the equation*

$$A \cdot B = \sum_{k=1}^{n} a_k b_k .$$

Thus, to compute $A \cdot B$ we multiply corresponding components of A and B and then add all the products. This multiplication has the following algebraic properties.

THEOREM 12.2. *For all vectors* A, B, C *in* V_n *and all scalars* c, *we have the following properties:*

(a) $A \cdot B = B \cdot A$ (*commutative law*),
(b) $A \cdot (B + C) = A \cdot B + A \cdot C$ (*distributive law*),
(c) $c(A \cdot B) = (cA) \cdot B = A \cdot (cB)$ (*homogeneity*),
(d) $A \cdot A > 0 \quad \text{if} \quad A \neq O$ (*positivity*),
(e) $A \cdot A = 0 \quad \text{if} \quad A = O$.

Length or norm of a vector

DEFINITION. *If A is a vector in* V_n, *its length or norm is denoted by* $\|A\|$ *and is defined by the equation*

$$\|A\| = (A \cdot A)^{1/2}.$$

The fundamental properties of the dot product lead to corresponding properties of norms.

THEOREM 12.4. *If A is a vector in* V_n *and if c is a scalar, we have the following properties:*

(a) $\|A\| > 0 \quad \text{if} \quad A \neq O$ (*positivity*),

(b) $\|A\| = 0 \quad \text{if} \quad A = O$,

(c) $\|cA\| = |c| \, \|A\|$ (*homogeneity*).

Orthogonality of vectors

DEFINITION. *Two vectors A and B in* V_n *are called perpendicular or orthogonal if* $A \cdot B = 0$.

Angle between vectors in n-space

DEFINITION. *Let A and B be two vectors in V_n. If both A and B are nonzero, the angle θ between A and B is defined by the equation*

$$\theta = \arccos \frac{A \cdot B}{\|A\| \, \|B\|} .$$

Note: The arc cosine function restricts θ to the interval $0 \leq \theta \leq \pi$. Note also that $\theta = \frac{1}{2}\pi$ when $A \cdot B = 0$.

The unit coordinate vectors

DEFINITION. *In* V_n, *the n vectors* $E_1 = (1, 0, \ldots, 0)$, $E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, 0, \ldots, 0, 1)$ *are called the unit coordinate vectors. It is understood that the kth component of* E_k *is* 1 *and all other components are* 0.

The name "unit vector" comes from the fact that each vector E_k has length 1. Note that these vectors are mutually orthogonal, that is, the dot product of any two distinct vectors is zero,

$$E_k \cdot E_j = 0 \qquad \text{if} \quad k \neq j .$$

THEOREM 12.6. *Every vector* $X = (x_1, \ldots, x_n)$ *in* V_n *can be expressed in the form*

$$X = x_1E_1 + \cdots + x_nE_n = \sum_{k=1}^{n} x_kE_k .$$

Moreover, this representation is unique. That is, if

$$X = \sum_{k=1}^{n} x_kE_k \qquad \textit{and} \qquad X = \sum_{k=1}^{n} y_kE_k ,$$

then $x_k = y_k$ *for each* $k = 1, 2, \ldots, n$.

A sum of the type $\sum c_i A_i$ is called a *linear combination* of the vectors $A_1, \ldots, A_n$. Theorem 12.6 tells us that every vector in V_n can be expressed as a linear combination of the unit coordinate vectors.

We describe this by saying that the unit coordinate vectors $E_1, \ldots, E_n$ *span* the space V_n. We also say they span V_n *uniquely* because each representation of a vector as a linear combination of $E_1, \ldots, E_n$ is unique. Some collections of vectors other than $E_1, \ldots, E_n$ also span V_n uniquely.

When vectors are expressed as linear combinations of the unit coordinate vectors, algebraic manipulations involving vectors can be performed by treating the sums $\sum x_k E_k$ according to the usual rules of algebra. The various components can be recognized at any stage in the calculation by collecting the coefficients of the unit coordinate vectors. For example, to add two vectors, say $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$, we write

$$A = \sum_{k=1}^{n} a_k E_k , \qquad B = \sum_{k=1}^{n} b_k E_k ,$$

and apply the linearity property of finite sums to obtain

$$A + B = \sum_{k=1}^{n} a_k E_k + \sum_{k=1}^{n} b_k E_k = \sum_{k=1}^{n} (a_k + b_k) E_k .$$

The coefficient of E_k on the right is the kth component of the sum $A + B$.

LINEAR SPACES

Briefly, a linear space is a set of elements of any kind on which certain operations (called *addition* and *multiplication by numbers*) can be performed.

In defining a linear space, we do not specify the nature of the elements nor do we tell how the operations are to be performed on them.

Instead, we require that the operations have certain properties which we take as axioms for a linear space. We turn now to a detailed description of these axioms.

The definition of a linear space

Let V denote a nonempty set of objects, called *elements*. The set V is called a linear space if it satisfies the following ten axioms which we list in three groups.

Closure axioms

AXIOM 1. CLOSURE UNDER ADDITION. *For every pair of elements* x *and* y *in* V *there corresponds a unique element in* V *called the sum of* x *and* y, *denoted by* $x + y$.

AXIOM 2. CLOSURE UNDER MULTIPLICATION BY REAL NUMBERS. *For every* x *in* V *and every real number* a *there corresponds an element in* V *called the product of* a *and* x, *denoted by* ax.

Axioms for addition

AXIOM 3. COMMUTATIVE LAW. *For all x and y in V, we have* $x + y = y + x$.

AXIOM 4. ASSOCIATIVE LAW. *For all x, y, and z in V, we have* $(x + y) + z = x + (y + z)$.

AXIOM 5. EXISTENCE OF ZERO ELEMENT. *There is an element in V, denoted by O, such that*

$$x + O = x \qquad \text{for all } x \text{ in } V .$$

AXIOM 6. EXISTENCE OF NEGATIVES. *For every x in V, the element* $(-1)x$ *has the property*

$$x + (-1)x = O .$$

Axioms for multiplication by numbers

AXIOM 7. ASSOCIATIVE LAW. *For every x in V and all real numbers a and b, we have*

$$a(bx) = (ab)x .$$

AXIOM 8. DISTRIBUTIVE LAW FOR ADDITION IN V. *For all x and y in V and all real a, we have*

$$a(x + y) = ax + ay .$$

AXIOM 9. DISTRIBUTIVE LAW FOR ADDITION OF NUMBERS. *For all x in V and all real a and b, we have*

$$(a + b)x = ax + bx .$$

AXIOM 10. EXISTENCE OF IDENTITY. *For every x in V, we have $1x = x$.*

Linear spaces, as defined above, are sometimes called *real* linear spaces to emphasize the fact that we are multiplying the elements of V by real numbers. If *real number* is replaced by *complex number* in Axioms 2, 7, 8, and 9, the resulting structure is called a *complex linear space*.

Sometimes a linear space is referred to as a *linear vector space* or simply a *vector space*; the numbers used as multipliers are also called *scalars*. A real linear space has real numbers as scalars; a complex linear space has complex numbers as scalars.

Examples of linear spaces

EXAMPLE 1. Let $V = \mathbf{R}$, the set of all real numbers, and let $x + y$ and ax be ordinary addition and multiplication of real numbers.

EXAMPLE 2. Let $V = \mathbf{C}$, the set of all complex numbers, define $x + y$ to be ordinary addition of complex numbers, and define ax to be multiplication of the complex number x by the real number a. Even though the elements of V are complex numbers, this is a real linear space because the scalars are real.

EXAMPLE 3. Let $V = V_n$, the vector space of all n-tuples of real numbers, with addition and multiplication by scalars defined in the usual way in terms of components.

The following examples are called *function spaces*. The elements of V are real-valued functions, with addition of two functions f and g defined in the usual way:

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

for every real x in the intersection of the domains of f and g. Multiplication of a function f by a real scalar a is defined as follows: af is that function whose value at each x in the domain of f is $af(x)$. The zero element is the function whose values are everywhere zero. The reader can easily verify that each of the following sets is a function space.

EXAMPLE 5. The set of all functions defined on a given interval.

EXAMPLE 6. The set of all polynomials.

EXAMPLE 9. The set of all functions differentiable at a given point.

EXAMPLE 10. The set of all functions integrable on a given interval.