LINEAR ALGEBRA

Richard E. Johnson

Linear Algebra

RICHARD E. JOHNSON

University of New Hampshire

W
P S

PRINDLE, WEBER & SCHMIDT, INCORPORATED

Boston, Massachusetts London Sydney

Library of Congress Catalog Card Number: 67-11932
Printed in the United States of America

LINEAR ALGEBRA

P W S

Preface

Linear algebra is an essential tool of the pure mathematician on the one hand and of the physicist and the theoretical economist on the other. This wide applicability of the subject makes its early mastery by the science student imperative. Like most branches of algebra, linear algebra also offers a convenient language for conveying complex ideas in a simple form. In linear algebra, it is the language of vectors, linear transformations, and matrices.

This book is intended for use in an introductory course on linear algebra. Principally, it is a study of finite-dimensional vector spaces and their associated algebras of linear transformations and matrices. No attempt has been made to show any applications of linear algebra other than to geometry. At every stage in the development of the subject matter, the direct, intuitive approach has been used whenever possible. We believe this approach will give the student the necessary background to appreciate in a later course a more sophisticated development using modules and their duals. The direct approach also has the advantage of being constructive and in this way allows a wide use of examples to illustrate new results and their proofs.

The first chapter has a two-fold purpose; to discuss the theory of finite-dimensional vector spaces and to develop examples of such spaces. Matrices are introduced in Chapter Two as arrays of numbers associated with systems of linear equations. Then, in Chapter Three, matrices are used to represent linear transformations of vector spaces. Determinants are developed in Chapter Four to allow us to compute inverses of matrices and solutions of systems of linear equations. The minimal polynomial of a linear transformation is described in Chapter Five and is used to find simple matrix representations of the transformation. Chapter Six discusses the splitting of a vector space into a direct sum of invariant subspaces relative to a linear transformation. The one-dimensional invariant subspaces yield the characteristic values of the transformation. The general theory of this splitting is described in Chapter Eight. Isometries and symmetric linear transformations of Euclidean spaces are analyzed in Chapter Seven.

The book contains enough material for a three- or four-hour semester course. A possible outline for a three-hour semester course is as follows: Chapter One, 8 days; Chapter Two, 4 days; Chapter Three, 6 days; Chapter Four, 5 days; Chapter Five, 4 days; Chapter Six, 6 days; Chapter Seven, 6 days. Chapter Eight can be omitted without any harm to the course. On the other hand, it makes good reading for the superior student and can be used in this way.

Richard E. Johnson

LINEAR ALGEBRA

P W S

Contents

Chapter One

Vector Spaces

1. FIELDS

Among the common number systems of mathematics are:

$\mathbb{Z}$, the system of integers,
$\mathbb{Q}$, the system of rational numbers,
$\mathbb{R}$, the system of real numbers,
$\mathbb{C}$, the system of complex numbers.

These are increasingly larger sets of numbers,

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Each of these systems is closed under the operations of addition and multiplication, which have the following properties:

1.1 $$a + b = b + a, \quad ab = ba \quad (\textit{Commutative laws}),$$

1.2 $$a + (b + c) = (a + b) + c, \quad a(bc) = (ab)c \quad (\textit{Associative laws}),$$

1.3 $$(a + b)c = ac + bc, \quad c(a + b) = ca + cb \quad (\textit{Distributive law}).$$

There exist numbers 0 and 1 with the following special property:

1.4 $$0 + a = a + 0 = a, \quad 1 \cdot a = a \cdot 1 = a \quad (\textit{Identity elements}).$$

Thus 0 is called the additive identity element and 1 the multiplicative identity element. Each number a has an opposite $-a$, called the negative of a, having the following property:

1.5 $$a + (-a) = (-a) + a = 0 \quad (\textit{Additive inverse}).$$

The five properties above hold for all elements a, b, c in any one of the systems $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. The systems $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ also have the following additional property: each nonzero number a has a reciprocal $1/a$ such that

1.6 $$a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1 \quad (\textit{Multiplicative inverse}).$$

We are now in a position to make the following definition.

1.7. DEFINITION OF A FIELD. An algebraic system composed of a set of elements F and operations of addition and multiplication in F is called a field if and only if the operations have properties 1.1 through 1.6.

Thus, according to this definition, the systems of rational numbers, real numbers, and complex numbers are examples of fields.

There exist fields of quite a different nature from $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. For example, for each prime number p there is a unique field having exactly p elements. This field is called the field of *integers modulo p* and is denoted by $\mathbb{Z}_p$,

$$\mathbb{Z}_p = \{0, 1, 2, \ldots, p-1\}.$$

Addition and multiplication in $\mathbb{Z}_p$ are defined to be the same as addition and multiplication in $\mathbb{Z}$, reduced modulo p. That is, to find $a + b$ and $a \cdot b$ for $a, b \in \mathbb{Z}_p$, we first compute them in $\mathbb{Z}$ and then subtract multiples of p from $a + b$ and $a \cdot b$ until we find a remainder in $\mathbb{Z}_p$. For example, in $\mathbb{Z}$ we have $3 + 4 = 7$; therefore, $3 + 4 = 7 - 5 = 2$ in $\mathbb{Z}_5$, and $3 + 4 = 7 - 7 = 0$ in $\mathbb{Z}_7$. Similarly, $3 \cdot 4 = 12$ in $\mathbb{Z}$; therefore $3 \cdot 4 = 12 - 10 = 2$ in $\mathbb{Z}_5$, and $3 \cdot 4 = 12 - 7 = 5$ in $\mathbb{Z}_7$. It is not hard to show that $\mathbb{Z}_p$ with operations of addition and multiplication so defined is a field.

Since each of the fields $\mathbb{Z}_p$, p a prime, is finite, we can give complete addition and multiplication tables for it. For example, the field $\mathbb{Z}_2 = \{0, 1\}$ has the simple tables:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

The field $\mathbb{Z}_3 = \{0, 1, 2\}$ has the following addition and multiplication tables:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Rational numbers and real numbers are ordered in the sense that for any two such numbers, one is greater than or equal to the other. Relative to this relation of "greater than or equal to," $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields as defined below.

1.8. DEFINITION OF AN ORDERED FIELD. A field F is called an ordered field if and only if it has a relation $\geq$ with the following properties:

(1) $a \geq a$ for all $a \in F$ (*Reflexive*).
(2) If $a \geq b$ and $b \geq a$, then $a = b$ (*Antisymmetric*).
(3) If $a \geq b$ and $b \geq c$, then $a \geq c$ (*Transitive*).
(4) For all $a, b \in F$, either $a \geq b$ or $b \geq a$.
(5) If $a \geq b$, then $a + c \geq b + c$ for all $c \in F$.
(6) If $a \geq b$, then $ac \geq bc$ for all $c \geq 0$.

The other order relations $>$, $\leq$, and $<$ are defined as usual. Thus, $a > b$ if and only if $a \geq b$ and $a \neq b$; $a \leq b$ if and only if $b \geq a$; and $a < b$ if and only if $b > a$. The relation $>$ (and, similarly, $<$) has the following properties:

1.9 If $a > b$ and $b > c$, then $a > c$ (*Transitive*).

1.10 If $a > b$, then $a + c > b + c$ for all $c \in F$.

1.11 If $a > b$, then $ac > bc$ for all $c > 0$.

Proof of 1.9: If $a > b$ and $b > c$, then $a \geq b$ and $b \geq c$ so that $a \geq c$ by 1.8(3). We claim $a > c$; for if $a = c$, then $c \geq b$, $b \geq c$, and $b = c$ by 1.8(2), contrary to the assumption that $b > c$.

Proof of 1.10: If $a > b$, then $a + c \geq b + c$ for all $c \in F$. If $a + c = b + c$, then $a = b$ by the additive cancellation law, contrary to the fact that $a > b$. Hence $a + c > b + c$.

Proof of 1.11: If $a > b$ and $c > 0$, then $ac \geq bc$ by 1.8(6). If $ac = bc$, then $a = b$ by the multiplicative cancellation law (since $c \neq 0$). This is contrary to the fact that $a > b$. Hence $ac > bc$.

If we let F be an ordered field and

$$F^+ = \{a \in F \mid a > 0\}, \qquad F^- = \{a \in F \mid a < 0\},$$

then we easily prove the following:

1.12 F^+ is closed under addition and multiplication.

1.13 $F^- = \{-a \mid a \in F^+\}$.

By our remarks above, every ordered field F is partitioned into three nonoverlapping subsets: F^+, the set of *positive elements;* F^-, the set of *negative elements;* and $\{0\}$:

$$F = F^+ \cup F^- \cup \{0\}.$$

Since $(-a)^2 = a^2$ for every nonzero $a \in F$ and either a or $-a$ is in F^+, we have, by 1.12, that

1.14 $$a^2 > 0 \qquad \text{for all nonzero } a \in F.$$

In particular, $1 > 0$ since $1^2 = 1$. In turn, $1 + 1 = 2 > 0$, $1 + 2 = 3 > 0$, and so on, by the closure of F^+ under addition.

A field such as $\mathbb{Z}_p$ is not ordered. For if $\mathbb{Z}_p$ were ordered, then $1 > 0$, $2 > 0$, and so on, up to $(p - 1) + 1 > 0$, contrary to the fact that $(p - 1) + 1 = 0$.

The field $\mathbb{C}$ of complex numbers is also not ordered. Thus it contains an element i such that $i^2 = -1$, and if it were ordered, then both -1 and 1 would be in $\mathbb{C}^+$ by 1.14. However, then $0 = (-1) + 1 \in \mathbb{C}^+$ by 1.12, contrary to the fact that $0 \notin \mathbb{C}^+$.

If F is an ordered field and $A \subset F$, $A \neq \varnothing$, the empty set, then an element b of F is called an *upper bound* of set A if and only if $x \leq b$ for all $x \in A$. Similarly, $c \in F$ is called a *lower bound* of set A if and only if $c \leq x$ for all $x \in A$. If b is an upper bound of A, whereas no element of F smaller than b is an upper bound of A, then b is called a *least upper bound* (l.u.b.) of set A. By definition, if a set has a l.u.b., then the l.u.b. is unique. However, a set need not have a l.u.b. The *greatest lower bound* (g.l.b.) of a set is defined similarly.

An ordered field F is called *complete* if and only if every subset of F which has an upper bound has a l.u.b. It is easily demonstrated that every subset of a complete field which has a lower bound has a g.l.b. It may be proved that the field $\mathbb{R}$ of real numbers is uniquely characterized by the following statement.

1.15. The field $\mathbb{R}$ is a complete ordered field.

This characterization of $\mathbb{R}$ allows us to show, for example, that every $a \in \mathbb{R}^+$ has a unique nth root $\sqrt[n]{a} \in \mathbb{R}^+$ for every integer $n > 1$. Thus it can be shown that

$$\sqrt[n]{a} = \text{l.u.b. } \{x \in \mathbb{R}^+ \mid x^n \leq a\}.$$

The field $\mathbb{C}$ of complex numbers is given by

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}, \qquad \text{where} \quad i^2 = -1.$$

An interesting property of $\mathbb{R}$ is that it can be used as a set of coordinates for the points on a line L. When this is done in the usual way, each point on L is assigned a unique real number as its coordinate, and each real number is the coordinate of a unique point on L. Furthermore, the order in $\mathbb{R}$ is preserved on L; i.e., if point B is between points A and C on L, then the coordinate of B is between the coordinates of A and C. An arrowhead is placed on L to indicate the direction of increasing coordinates, as shown in Fig. 1.1.

Figure 1.1

We shall call a line L having $\mathbb{R}$ as a coordinate system a *coordinate line* or a *coordinate axis*. The point on L with coordinate 0 is called the *origin*.

The numbers are assigned in a regular way on a coordinate line so that distances may be easily computed as follows.

1.16. DISTANCE FORMULA ON A LINE. If points A and B on a coordinate line have respective coordinates a and b, then the distance $d(A, B)$ between A and B is given by

$$d(A, B) = |b - a|.$$

At times it is convenient to use directed distances on a coordinate line L. If points A and B on L have respective coordinates a and b, then the *directed distance* from A to B is defined to be $b - a$. Thus, by definition, the directed distance from A to B is simply $d(A, B)$ if the direction from A to B is the direction of L, zero if $A = B$, and $-d(A, B)$ if the direction from A to B is opposite to the direction of L.

The set of all ordered pairs of real numbers is denoted by $\mathbb{R}^2$. Thus, we have

$$\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}.$$

We can use $\mathbb{R}^2$ as a set of coordinates for the points in a plane. This is usually done as shown in Fig. 1.2. Thus, two perpendicular coordinate axes are chosen in the plane so that they intersect at their origins. One of the coordinate axes is labeled the *x-axis* and the other the *y-axis*. Each point P has unique coordinates (a, b) in $\mathbb{R}^2$, where the coordinate of the foot of the perpendicular drawn from P to the x-axis is a and from P to the y-axis is b. Also, each ordered pair (a, b) in $\mathbb{R}^2$ are the coordinates of a unique point in the plane chosen in the obvious way. If point P has coordinate (a, b) in $\mathbb{R}^2$, then we call a the *x-coordinate* and b the *y-coordinate* of P. We shall call a plane having $\mathbb{R}^2$ as a set of coordinates in the manner described above a *rectangular coordinate plane* or a *Cartesian plane*.

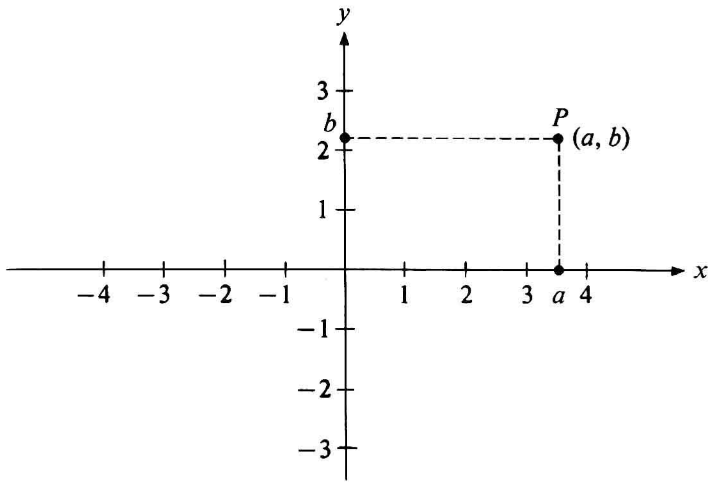

Figure 1.2

If the x-axis and the y-axis in a Cartesian plane have the same scale, then we can find the distance between any two points in the plane as follows.

1.17. DISTANCE FORMULA IN A PLANE. If points A and B in a coordinate plane have respective coordinates (a_1, a_2) and (b_1, b_2), then the distance $d(A, B)$ between A and B is given by

$$d(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

We shall not give the proof of 1.17. It follows readily from the Pythagorean Theorem and 1.16.

The set of all ordered triplets of real numbers is denoted by $\mathbb{R}^3$. Thus, we have

$$\mathbb{R}^3 = \{(a, b, c) \mid a, b, c \in \mathbb{R}\}.$$

We can use $\mathbb{R}^3$ as a set of coordinates for the points in space. This is usually accomplished by selecting three mutually perpendicular coordinate axes intersecting at their origins. Let us label these axes the x-axis, the y-axis, and the z-axis. Each point P in space has unique coordinates (a, b, c) in $\mathbb{R}^3$, where the coordinate of the foot of the perpendicular drawn from P to the x-axis is a, from P to the y-axis is b, and from P to the z-axis is c (Fig. 1.3). Also, each ordered triplet (a, b, c) in $\mathbb{R}^3$ are the coordinates of a unique point in space. If P has coordinates (a, b, c), then we call a the *x-coordinate*, b the *y-coordinate*, and c the *z-coordinate* of P. We shall call space *Cartesian three-space* if coordinates are assigned to the points of space in the manner described above.

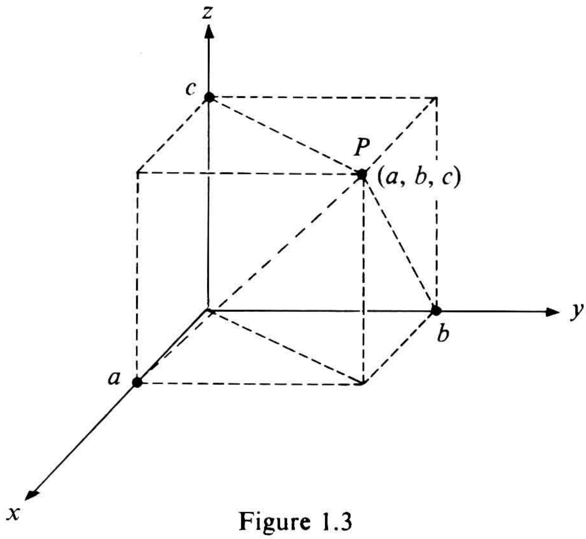

Figure 1.3

If the coordinate axes in a Cartesian three-space have the same scale, then the distance between two points can be found as follows.

1.18. DISTANCE FORMULA IN SPACE. If points A and B in coordinate three-space have respective coordinates (a_1, a_2, a_3) and (b_1, b_2, b_3), then the distance $d(A, B)$ between A and B is given by

$$d(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.$$

The proof of 1.18 follows easily from the Pythagorean Theorem and 1.17.

2. VECTOR SPACES

A vector is an element of a vector space. In turn, a vector space is a set of objects, called vectors, that is closed under operations of addition and scalar multiplication and which satisfies certain algebraic laws. A precise definition of a vector space is given below.

It is worthwhile to study vector spaces for the reason that many of the algebraic systems encountered in applications of mathematics are in essence vector spaces. By studying general vector spaces, without regard to the nature of the elements, we can develop the properties common to all vector spaces.

1.19. DEFINITION OF A VECTOR SPACE. A vector space consists of a set V, an operation of addition in V, and an operation of scalar multiplication of V by a field F. Addition in V has the following properties:

(1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in V$ (*Commutative law*).