An Introduction to LINEAR ALGEBRA

Ravi P. Agarwal and Cristina Flaut

CPD CERTIFIED
The CPD Certification Service

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

An Introduction to

LINEAR ALGEBRA

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

An Introduction to
LINEAR ALGEBRA

Ravi P. Agarwal and Cristina Flaut

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2017 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper

International Standard Book Number-13: 978-1-138-62670-6 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedicated to our mothers:
Godawari Agarwal, Elena Paiu, and Maria Paiu

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contents

Preface

Linear algebra is a branch of both pure and applied mathematics. It provides the foundation for multi-dimensional representations of mathematical reasoning. It deals with systems of linear equations, matrices, determinants, vectors and vector spaces, transformations, and eigenvalues and eigenvectors. The techniques of linear algebra are extensively used in every science where often it becomes necessary to approximate nonlinear equations by linear equations. Linear algebra also helps to find solutions for linear systems of differential and difference equations. In pure mathematics, linear algebra (particularly, vector spaces) is used in many different areas of algebra such as group theory, module theory, representation theory, ring theory, Galöis theory, and this list continues. This has given linear algebra a unique place in mathematics curricula all over the world, and it is now being taught as a compulsory course at various levels in almost every institution.

Although several fabulous books on linear algebra have been written, the present rigorous and transparent introductory text can be used directly in class for students of applied sciences. In fact, in an effort to bring the subject to a wider audience, we provide a compact, but thorough, introduction to the subject in **An Introduction to Linear Algebra**. This book is intended for senior undergraduate and for beginning graduate one-semester courses.

The subject matter has been organized in the form of theorems and their proofs, and the presentation is rather unconventional. It comprises 25 class-tested lectures that the first author has given to math majors and engineering students at various institutions over a period of almost 40 years. It is our belief that the content in a particular chapter, together with the problems therein, provides fairly adequate coverage of the topic under study.

A brief description of the topics covered in this book follows: In **Chapter 1**, we define axiomatically terms such as field, vector, vector space, subspace, linear combination of vectors, and span of vectors. In **Chapter 2**, we introduce various types of matrices and formalize the basic operations: matrix addition, subtraction, scalar multiplication, and matrix multiplication. We show that the set of all $m \times n$ matrices under the operations matrix addition and scalar multiplication is a vector space. In **Chapter 3**, we begin with the definition of a determinant and then briefly sketch the important properties of determinants. In **Chapter 4**, we provide necessary and sufficient conditions for a square matrix to be invertible. We shall show that the theory of determinants can be applied to find an analytical representation of the inverse of a square matrix. Here we also use elementary theory of difference equations to find inverses of some band matrices.

The main purpose of **Chapters 5 and 6** is to discuss systematically Gauss and Gauss–Jordan elimination methods to solve m linear equations in n unknowns. These equations are conveniently written as $Ax = b$, where A is an $m \times n$ matrix, x is an $n \times 1$ unknown vector, and b is an $m \times 1$ vector. For this, we introduce the terms consistent, inconsistent, solution space, null space, augmented matrix, echelon form of a matrix, pivot, elementary row operations, elementary matrix, row equivalent matrix, row canonical form, and rank of a matrix. These methods also provide effective algorithms to compute determinants and inverses of matrices. We also prove several theoretical results that yield necessary and sufficient conditions for a linear system of equations to have a solution. **Chapter 7** deals with a modified but restricted realization of Gaussian elimination. We factorize a given $m \times n$ matrix A to a product of two matrices L and U, where L is an $m \times m$ lower triangular matrix, and U is an $m \times n$ upper triangular matrix. Here we also discuss various variants and applications of this factorization.

In **Chapter 8**, we define the concepts linear dependence and linear independence of vectors. These concepts play an essential role in linear algebra and as a whole in mathematics. Linear dependence and independence distinguish between two vectors being essentially the same or different. In **Chapter 9**, for a given vector space, first we introduce the concept of a basis and then describe its dimension in terms of the number of vectors in the basis. Here we also introduce the concept of direct sum of two subspaces. In **Chapter 10**, we extend the known geometric interpretation of the coordinates of a vector in R^3 to a general vector space. We show how the coordinates of a vector space with respect to one basis can be changed to another basis. Here we also define the terms ordered basis, isomorphism, and transition matrix. In **Chapter 11**, we redefine rank of a matrix and show how this number is directly related to the dimension of the solution space of homogeneous linear systems. Here for a given matrix we also define row space, column space, left and right inverses, and provide necessary and sufficient conditions for their existence. In **Chapter 12**, we introduce the concept of linear mappings between two vector spaces and extend some results of earlier chapters. In **Chapter 13**, we establish a connection between linear mappings and matrices. We also introduce the concept of similar matrices, which plays an important role in later chapters. In **Chapter 14**, we extend the familiar concept inner product of two or three dimensional vectors to general vector spaces. Our definition of inner products leads to the generalization of the notion of perpendicular vectors, called orthogonal vectors. We also discuss the concepts projection of a vector onto another vector, unitary space, orthogonal complement, orthogonal basis, and Fourier expansion. This chapter concludes with the well-known Gram–Schmidt orthogonalization process. In **Chapter 15**, we discuss a special type of linear mapping, known as linear functional. We also address such notions as dual space, dual basis, second dual, natural mapping, adjoint mapping, annihilator, and prove the famous Riesz representation theorem.

Chapter 16 deals with the eigenvalues and eigenvectors of matrices. We summarize those properties of the eigenvalues and eigenvectors of matrices that facilitate their computation. Here we come across the concepts characteristic polynomial, algebraic and geometric multiplicities of eigenvalues, eigenspace, and companion and circulant matrices. We begin **Chapter 17** with the definition of a norm of a vector and then extend it to a matrix. Next, we drive some estimates on the eigenvalues of a given matrix, and prove some useful convergence results. Here we also establish well known Cauchy–Schwarz, Minkowski, and Bessel inequalities, and discuss the terms spectral radius, Rayleigh quotient, and best approximation.

In **Chapter 18**, we show that if algebraic and geometric multiplicities of an $n\times n$ matrix A are the same, then it can be diagonalized, i.e., $A = PDP^{-1}$; here, P is a nonsingular matrix and D is a diagonal matrix. Next, we provide necessary and sufficient conditions for A to be orthogonally diagonalizable, i.e., $A = QDQ^t$, where Q is an orthogonal matrix. Then, we discuss QR factorization of the matrix A. We also furnish complete computationable characterizations of the matrices P, D, Q, and R. In **Chapter 19**, we develop a generalization of the diagonalization procedure discussed in Chapter 18. This factorization is applicable to any real $m \times n$ matrix A, and in the literature has been named singular value decomposition. Here we also discuss reduced singular value decomposition.

In **Chapter 20**, we show how linear algebra (especially eigenvalues and eigenvectors) plays an important role to find the solutions of homogeneous differential and difference systems with constant coefficients. Here we also develop continuous and discrete versions of the famous Putzer's algorithm. In a wide range of applications, we encounter problems in which a given system $Ax = b$ does not have a solution. For such a system we seek a vector(s) $\hat{x}$ so that the error in the Euclidean norm, i.e., $\|A\hat{x} - b\|_2$, is as small as possible (minimized). This solution(s) $\hat{x}$ is called the least squares approximate solution. In **Chapter 21**, we shall show that a least squares approximate solution always exists and can be conveniently computed by solving a related system of n equations in n unknowns (normal equations). In **Chapter 22**, we study quadratic and diagonal quadratic forms in n variables, and provide criteria for them to be positive definite. Here we also discuss maximum and minimum of the quadratic forms subject to some constraints (constrained optimization). In **Chapter 23**, first we define positive definite symmetric matrices in terms of quadratic forms, and then for a symmetric matrix to be positive definite, we provide necessary and sufficient conditions. Next, for a symmetric matrix we revisit LU-factorization, and give conditions for a unique factorization LDL^t, where L is a lower triangular matrix with all diagonal elements 1, and D is a diagonal matrix with all positive elements. We also discuss Cholesky's decomposition $L_cL_c^t$ where $L_c = LD^{1/2}$, and for its computation provide Cholesky's algorithm. This is followed by Sylvester's criterion, which gives easily verifiable necessary and sufficient conditions for a symmetric matrix to be positive definite. We conclude this chapter with a polar decomposition. In **Chapter 24**, we introduce the concept of pseudo/generalized (Moore–Penrose) inverse which is applicable to all $m \times n$ matrices. As an illustration we apply Moore–Penrose inverse to least squares solutions of linear equations. Finally, in **Chapter 25**, we briefly discuss irreducible, nonnegative, diagonally dominant, monotone, and Toeplitz matrices. We state 11 theorems which, from the practical point of view, are of immense value. These types of matrices arise in several diverse fields, and hence have attracted considerable attention in recent years.

In this book, there are 148 examples that explain each concept and demonstrate the importance of every result. Two types of 254 problems are also included, those that illustrate the general theory and others designed to fill out text material. The problems form an integral part of the book, and every reader is urged to attempt most, if not all of them. For the convenience of the reader, we have provided answers or hints to all the problems.

In writing a book of this nature, no originality can be claimed, only a humble attempt has been made to present the subject as simply, clearly, and accurately as possible. The illustrative examples are usually very simple, keeping in mind an average student.

It is earnestly hoped that **An Introduction to Linear Algebra** will serve an inquisitive reader as a starting point in this rich, vast, and ever-expanding field of knowledge.

We would like to express our appreciation to our students and Ms. Aastha Sharma at CRC (New Delhi) for her support and cooperation.

Ravi P. Agarwal
Cristina Flaut

Chapter 1

Linear Vector Spaces

A vector space (or linear space) consists of four things $\{F, V, +, \text{s.m.}\}$, where F is a field of scalars, V is the set of vectors, and $+$ and s.m. are binary operations on the set V called vector addition and scalar multiplication, respectively. In this chapter we shall define each term axiomatically and provide several examples.

Fields. A *field* is a set of scalars, denoted by F, in which two binary operations, addition $(+)$ and multiplication $(\cdot)$, are defined so that the following axioms hold:

A1. *Closure property of addition:* If $a, b \in F$, then $a + b \in F$.

A2. *Commutative property of addition:* If $a, b \in F$, then $a + b = b + a$.

A3. *Associative property of addition:* If $a, b, c \in F$, then $(a+b)+c = a+(b+c)$.

A4. *Additive identity:* There exists a zero element, denoted by 0, in F such that for all $a \in F$, $a + 0 = 0 + a = a$.

A5. *Additive inverse:* For each $a \in F$, there is a unique element $(-a) \in F$ such that $a + (-a) = (-a) + a = 0$.

A6. *Closure property of multiplication:* If $a, b \in F$, then $a \cdot b \in F$.

A7. *Commutative property of multiplication:* If $a, b \in F$, then $a \cdot b = b \cdot a$.

A8. *Associative property of multiplication:* If $a, b, c \in F$, then $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

A9. *Multiplicative identity:* There exists a unit element, denoted by 1, in F such that for all $a \in F$, $a \cdot 1 = 1 \cdot a = a$.

A10. *Multiplicative inverse:* For each $a \in F$, $a \neq 0$, there is an unique element $a^{-1} \in F$ such that $a \cdot a^{-1} = a^{-1}a = 1$.

A11. *Left distributivity:* If $a, b, c \in F$, then $a \cdot (b + c) = a \cdot b + a \cdot c$.

A12. *Right distributivity:* If $a, b, c \in F$, then $(a + b) \cdot c = a \cdot c + b \cdot c$.

Example 1.1. The set of rational numbers Q, the set of real numbers R, and the set of complex numbers C, with the usual definitions of addition and multiplication, are fields. The set of natural numbers $N = \{1, 2, \cdots\}$, and the set of all integers $Z = \{\cdots, -2, -1, 0, 1, 2 \cdots\}$ are not fields.

Let F and F_1 be fields and $F_1 \subseteq F$, then F_1 is called a *subfield* of F. Thus, Q is a subfield of R, and R is a subfield of C.

Vector spaces. A *vector space* V over a field F denoted as (V, F) is a nonempty set of elements called *vectors* together with two binary operations, addition of vectors and multiplication of vectors by scalars, so that the following axioms hold:

B1. *Closure property of addition:* If $u, v \in V$, then $u + v \in V$.

B2. *Commutative property of addition:* If $u, v \in V$, then $u + v = v + u$.

B3. *Associativity property of addition:* If $u, v, w \in V$, then $(u + v) + w = u + (v + w)$.

B4. *Additive identity:* There exists a zero vector, denoted by 0, in V such that for all $u \in V$, $u + 0 = 0 + u = u$.

B5. *Additive inverse:* For each $u \in V$, there exists a vector v in V such that $u + v = v + u = 0$. Such a vector v is usually written as $-u$.

B6. *Closure property of multiplication:* If $u \in V$ and $a \in F$, then the product $a \cdot u = au \in V$.

B7. If $u, v \in V$ and $a \in F$, then $a(u + v) = au + av$.

B8. If $u \in V$ and $a, b \in F$, then $(a + b)u = au + bu$.

B9. If $u \in V$ and $a, b \in F$, then $ab(u) = a(bu)$.

B10. *Multiplication of a vector by a unit scalar:* If $u \in V$ and $1 \in F$, then $1u = u$.

In what follows, the subtraction of the vector v from u will be written as $u - v$, and by this we mean $u + (-v)$, or $u + (-1)v$. The spaces (V, R) and (V, C) will be called *real* and *complex vector spaces*, respectively.

Example 1.2 (The n-tuple space). Let F be a given field. We consider the set V of all ordered n-tuples

$$u = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \qquad (\text{or}, \quad (a_1, \cdots, a_n))$$

of scalars (known as *components*) $a_i \in F$. If

$$v = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

is in V, the addition of u and v is defined by

$$u + v = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix},$$

and the product of a scalar $c \in F$ and vector $u \in V$ is defined by

$$cu = \begin{pmatrix} ca_1 \\ \vdots \\ ca_n \end{pmatrix}.$$

It is to be remembered that $u = v$, if and only if their corresponding components are equal, i.e., $a_i = b_i,\ i = 1, \cdots, n$. With this definition of addition and scalar multiplication it is easy to verify all the axioms B1–B10, and hence this (V, F) is a vector space. In particular, if

$$w = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

is in V, then the i-th component of $(u+v)+w$ is $(a_i+b_i)+c_i$, which in view of A3 is the same as $a_i+(b_i+c_i)$, and this is the same as the i-th component of $u+(v+w)$, i.e., B3 holds. If $F = R$, then V is denoted as R^n, which for $n = 2$ and 3 reduces respectively to the two and three dimensional usual vector spaces. Similarly, if $F = C$, then V is written as C^n.

Example 1.3 (The space of polynomials). Let F be a given field. We consider the set $\mathcal{P}_n,\ n \geq 1$ of all polynomials of degree *at most* $n-1$, i.e.,

$$\mathcal{P}_n = \left\{ a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} = \sum_{i=0}^{n-1} a_i x^i : a_i \in F, \quad x \in R \right\}.$$

If $u = \sum_{i=0}^{n-1} a_i x^i,\ v = \sum_{i=0}^{n-1} b_i x^i \in \mathcal{P}_n$, then the addition of vectors u and v is defined by

$$u + v = \sum_{i=0}^{n-1} a_i x^i + \sum_{i=0}^{n-1} b_i x^i = \sum_{i=0}^{n-1}(a_i + b_i)x^i,$$

and the product of a scalar $c \in F$ and vector $u \in \mathcal{P}_n$ is defined by

$$cu = c\sum_{i=0}^{n-1} a_i x^i = \sum_{i=0}^{n-1}(ca_i)x^i.$$

This $(\mathcal{P}_n, F)$ is a vector space. We remark that the set of all polynomials of degree exactly $n-1$ is not a vector space. In fact, if we choose $b_{n-1} = -a_{n-1}$, then $u + v$ is a polynomial of degree $n-2$.

Example 1.4 (The space of functions). Let F be a given field, and $X \subseteq F$. We consider the set V of all functions from the set X to F. The sum of two vectors $f, g \in V$ is defined by $(f+g)$, i.e., $(f+g)(x) = f(x)+g(x),\ x \in X$, and the product of a scalar $c \in F$ and vector $f \in V$ is defined by cf, i.e., $(cf)(x) = cf(x)$. This (V, F) is a vector space. In particular, $(C[X], F)$, where $C[X]$ is the set of all continuous functions from X to F, with the same vector addition, and scalar multiplication is a vector space.

Example 1.5 (The space of sequences). Let F be a given field. Consider the set S of all sequences $a = \{a_n\}_{n=1}^{\infty}$, where $a_n \in F$. If a and b are in S and $c \in F$, we define $a + b = \{a_n\} + \{b_n\} = \{a_n + b_n\}$ and $ca = c\{a_n\} = \{ca_n\}$. Clearly, (S, F) is a vector space.

Example 1.6. Let $F = R$ and V be the set of all solutions of the homogeneous ordinary linear differential equation with real constant coefficients

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n y \;=\; 0, \quad a_0 \neq 0, \quad x \in R.$$

This (V, F) is a vector space with the same vector addition and scalar multiplication as in Example 1.4. Note that if the above differential equation is nonhomogeneous then (V, F) is not a vector space.

Theorem 1.1. Let V be a vector space over the field F, and let $u, v \in V$. Then,

1. $u + v = u$ implies $v = 0 \in V$.
2. $0u = 0 \in V$.
3. $-u$ is unique.
4. $-u = (-1)u$.

Proof. 1. On adding $-u$ on both sides of $u + v = u$, we have

$$-u + u + v \;=\; -u + u \;\Rightarrow\; (-u + u) + v = 0 \;\Rightarrow\; 0 + v = 0 \;\Rightarrow\; v = 0.$$

2. Clearly, $0u = (0 + 0)u = 0u + 0u$, and hence $0u = 0 \in V$.

3. Assume that v and w are such that $u + v = 0$ and $u + w = 0$. Then, we have

$$v \;=\; v + 0 \;=\; v + (u + w) \;=\; (v + u) + w \;=\; (u + v) + w \;=\; 0 + w \;=\; w,$$

i.e., $-u$ of any vector $u \in V$ is unique.

4. Since

$$0 \;=\; 0u \;=\; [1 + (-1)]u \;=\; 1u + (-1)u \;=\; u + (-1)u,$$

it follows that $(-1)u$ is a negative for u. The uniqueness of this negative vector now follows from Part 3. ∎

Subspaces. Let (V, F) and (W, F) be vector spaces and $W \subseteq V$, then (W, F) is called a *subspace* of (V, F). It is clear that the smallest subspace

(W, F) of (V, F) consists of only the zero vector, and the largest subspace (W, F) is (V, F) itself.

Example 1.7. Let $F = R$,

$$W = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} : a_1, a_2 \in R \right\} \quad \text{and} \quad V = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} : a_1, a_2, a_3 \in R \right\}.$$

Clearly, (W, R) is a subspace of (V, R). However, if we let

$$W = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} : a_1 > 0, a_2 > 0, a_3 > 0 \right\},$$

then (W, R) is not a subspace of (V, R).

Example 1.8. Let F be a given field. Consider the vector spaces $(\mathcal{P}_4, F)$ and $(\mathcal{P}_3, F)$. Clearly, $(\mathcal{P}_3, F)$ is a subspace of $(\mathcal{P}_4, F)$. However, the set of all polynomials of degree exactly two over the field F is not a subspace of $(\mathcal{P}_4, F)$.

Example 1.9. Consider the vector spaces (V, F) and $(C[X], F)$ considered in Example 1.4. Clearly, $(C[X], F)$ is a subspace of (V, F).

To check if the nonempty subset W of V over the field F is a subspace requires the verification of all the axioms B1–B10. However, the following result simplifies this verification considerably.

Theorem 1.2. If (V, F) is a vector space and W is a nonempty subset of V, then (W, F) is a subspace of (V, F) if and only if for each pair of vectors $u, v \in W$ and each scalar $a \in F$ the vector $au + v \in W$.

Proof. If (W, F) is a subspace of (V, F), and $u, v \in W,\ a \in F$, then obviously $au + v \in W$. Conversely, since $W \neq \emptyset$, there is a vector $u \in W$, and hence $(-1)u + u = 0 \in W$. Further, for any vector $u \in W$ and any scalar $a \in F$, the vector $au = au + 0 \in W$. This in particular implies that $(-1)u = -u \in W$. Finally, we notice that if $u, v \in W$, then $1u + v \in W$. The other axioms can be shown similarly. Thus (W, F) is a subspace of (V, F). ∎

Thus (W, F) is a subspace of (V, F) if and only if for each pair of vectors $u, v \in W, u + v \in W$ and for each scalar $a \in F,\ au \in W$.

Let $u^1, \cdots, u^n$ be vectors in a given vector space (V, F), and $c_1, \cdots, c_n \in F$ be scalars. The vector $u = c_1 u^1 + \cdots + c_n u^n$ is known as *linear combination* of $u^i,\ i = 1, \cdots, n$. By mathematical induction it follows that $u \in (V, F)$.

Theorem 1.3. Let $u^i \in (V, F),\ i = 1, \cdots, n(\geq 1)$, and

$$W = \{c_1 u^1 + \cdots + c_n u^n \ : \ c_i \in F,\ i = 1, \cdots, n\}$$

then (W, F) is a subspace of (V, F), and W contains each of the vectors $u^i,\ i = 1, \cdots, n$.

Proof. Clearly, each u^i is a linear combination of the form

$$u^i \;=\; \sum_{j=1}^{n} \delta_{ij} u^j,$$

where δ_{ij} is the Kronecker delta defined by

$$\delta_{ij} \;=\; \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

Thus, each $u^i \in W$. Now, if $v = \sum_{i=1}^n c_i u^i,\ w = \sum_{i=1}^n d_i u^i$ and $a \in F$, then we have

$$av + w = a\sum_{i=1}^{n} c_i u^i + \sum_{i=1}^{n} d_i u^i = \sum_{i=1}^{n} (ac_i + d_i) u^i = \sum_{i=1}^{n} \alpha_i u^i, \quad \alpha_i \in F$$

which shows that $av + w \in W$. The result now follows from Theorem 1.2. ∎

The subspace (W, F) in Theorem 1.3 is called the subspace *spanned* or *generated* by the vectors $u^i,\ i = 1, \cdots, n$, and written as $\mathrm{Span}\{u^1, \cdots, u^n\}$. If $(W, F) = (V, F)$, then the set $\{u^1, \cdots, u^n\}$ is called a *spanning set* for the vector space (V, F). Clearly, in this case each vector $u \in V$ can be expressed as a linear combination of vectors $u^i,\ i = 1, \cdots, n$.

Example 1.10. Since

$$2\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} - 3\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + 5\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 12 \\ 10 \\ 7 \end{pmatrix}$$

it follows that

$$\begin{pmatrix} 12 \\ 10 \\ 7 \end{pmatrix} \;\in\; \mathrm{Span}\left\{\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}\right\}.$$

However,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \;\notin\; \mathrm{Span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right\}.$$

Example 1.11. For the vector space (V, F) considered in Example 1.2

the set $\{e^1, \cdots, e^n\}$, where

$$e^i \;=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \;\in\; V \quad (1 \;\text{ at the }\; i\text{-th place})$$

is a spanning set. Similarly, for the vector space $(\mathcal{P}_n, F)$ considered in Example 1.3, the set $\{1, x, \cdots, x^{n-1}\}$ is a spanning set.

Problems

1.1. Show that the set of all real numbers of the form $a + \sqrt{2}b$, where a and b are rational numbers, is a field.

1.2. Show that

(i) if $u^1, \cdots, u^n$ span V and $u \in V$, then $u, u^1, \cdots, u^n$ also span V

(ii) if $u^1, \cdots, u^n$ span V and u^k is a linear combination of u^i, $i = 1, \cdots, n$, $i \neq k$, then u^i, $i = 1, \cdots, n$, $i \neq k$ also span V

(iii) if $u^1, \cdots, u^n$ span V and $u^k = 0$, then u^i, $i = 1, \cdots, n$, $i \neq k$ also span V.

1.3. Show that the intersection of any number of subspaces of a vector space V is a subspace of V.

1.4. Let U and W be subspaces of a vector space V. The space

$$U + W \;=\; \{v : v = u + w \;\;\text{where}\;\; u \in U, \;\; w \in W\}$$

is called the *sum* of U and W. Show that

(i) $U + W$ is also a subspace of V

(ii) U and W are contained in $U + W$

(iii) $U + U = U$

(iv) $U \cup W$ is a subspace of V?.

1.5. Consider the following polynomials of degree three:

$$L_1(x) = \frac{(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)}, \;\; L_2(x) = \frac{(x-x_1)(x-x_3)(x-x_4)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)}$$

$$L_3(x) = \frac{(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)}, \;\; L_4(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_1)(x_4-x_2)(x_4-x_3)},$$

where $x_1 < x_2 < x_3 < x_4$. Show that

(i) if $P_3(x) \in \mathcal{P}_4$ is an arbitrary polynomial of degree three, then $P_3(x) = L_1(x)P_3(x_1) + L_2(x)P_3(x_2) + L_3(x)P_3(x_3) + L_4(x)P_3(x_4)$

(ii) the set $\{L_1(x), L_2(x), L_3(x), L_4(x)\}$ is a spanning set for $(\mathcal{P}_4, R)$.

1.6. Prove that the sets $\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$ and $\{1, (1-x), (1-x)^2, (1-x)^3\}$ are spanning sets for $(\mathcal{P}_4, R)$.

1.7. Let S be a subset of R^n consisting of all vectors with components $a_i,\ i = 1, \cdots, n$ such that $a_1 + \cdots + a_n = 0$. Show that S is a subspace of R^n.

1.8. On R^3 we define the following operations

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 0 \\ x_3 + y_3 \end{pmatrix} \text{ and } a \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ ax_3 \end{pmatrix}, \ a \in R.$$

With these operations, is R^3 a vector space over the field R?

1.9. Consider the following subsets of the vector space R^3:

(i) $V_1 = \{x \in R^3 : 3x_3 = x_1 - 5x_2\}$ (ii) $V_2 = \{x \in R^3 : x_1^2 = x_2 + 6x_3\}$

(iii) $V_3 = \{x \in R^3 : x_2 = 0\}$ (iv) $V_4 = \{x \in R^3 : x_2 = a,\ a \in R - \{0\}\}$.

Find if the above sets V_1, V_2, V_3, and V_4 are vector subspaces of R^3.

1.10. Let (V, X) be the vector space of functions considered in Example 1.4 with $X = F = R$, and $W \subset V$. Show that W is a subspace of V if

(i) W contains all bounded functions

(ii) W contains all even functions $(f(-x) = f(x))$

(iii) W contains all odd functions $(f(-x) = -f(x))$.

Answers or Hints

1.1. Verify A1–A12.

1.2. (i) Since $u^1, \cdots, u^n$ span V and $u \in V$ there exist scalars $c_1, \cdots, c_n$ such that $u = \sum_{i=1}^n c_i u^i$. Let $W = \{v : v = \sum_{i=1}^n \alpha_i u^i + \alpha_{n+1} u\}$. We need to show that $(V, F) = (W, F)$. Clearly, $V \subseteq W$. Now let $v \in W$, then $v = \sum_{i=1}^n \alpha_i u^i + \alpha_{n+1} \sum_{i=1}^n c_i u^i = \sum_{i=i}^n (\alpha_i + \alpha_{n+1} c_i) u^i$. Hence, $W \subseteq V$.

(ii) Similar as (i).

(iii) Similar as (i).

1.3. Let U, W be subspaces of V. It suffices to show that $U \cap W$ is also a subspace of V. Since $0 \in U$ and $0 \in W$ it is clear that $0 \in U \cap W$. Now let $u, w \in U \cap W$, then $u, w \in U$ and $u, w \in W$. Further for all scalars $a, b \in F$, $au + bw \in U$ and $au + bw \in W$. Thus $au + bw \in U \cap W$.

1.4. (i) Let $v^1, v^2 \in U + W$, where $v^1 = u^1 + w^1,\ v^2 = u^2 + w^2$. Then, $v^1 + v^2 = u^1 + w^1 + u^2 + w^2 = (u^1 + u^2) + (w^1 + w^2)$. Now since U and W are subspaces, $u^1 + u^2 \in U$ and $w^1 + w^2 \in W$. This implies that $v^1 + v^2 \in U + W$. Similarly we can show that $cv^1 \in U + W,\ c \in F$.

(ii) If $u \in U$, then since $0 \in W,\ u = u + 0 \in U + W$.

(iii) Since U is a subspace of V it is closed under vector addition, and hence $U + U \subseteq U$. We also have $U \subseteq U + U$ from (i).

(iv) $U \cup W$ need not be a subspace of V. For example, consider $V = R^3$,

$$U = \left\{ \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} : a_1 \in R \right\},\ W = \left\{ \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} : a_3 \in R \right\}.$$

Then

$$U \cup W = \left\{ \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix} : a_1 \in R,\ a_3 \in R \right\}.$$

Clearly,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in U \cup W,\ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in U \cup W, \quad \text{but} \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \notin U \cup W.$$

1.5. (i) The function $f(x) = L_1(x)P_3(x_1) + L_2(x)P_3(x_2) + L_3(x)P_3(x_3) + L_4(x)P_4(x_4)$ is a polynomial of degree at most three, and $f(x_i) = L_i(x_i) \times P_3(x_i) = P_3(x_i),\ i = 1, 2, 3, 4$. Thus $f(x) = P_3(x)$ follows from the uniqueness of interpolating polynomials.

(ii) Follows from (i).

1.6. It suffices to note that $a + bx + cx^2 + dx^3 = (a - b) + (b - c)(1 + x) + (c - d)(1 + x + x^2) + d(1 + x + x^2 + x^3)$.

1.7. Use Theorem 1.2.

1.8. No.

1.9. V_1 and V_3 are vector subspaces, whereas V_2 and V_4 are not vector subspaces of R^3.

1.10. Use Theorem 1.2.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Chapter 2

Matrices

Matrices occur in many branches of applied mathematics and social sciences, such as algebraic and differential equations, mechanics, theory of electrical circuits, nuclear physics, aerodynamics, and astronomy. It is, therefore, necessary for every young scientist and engineer to learn the elements of matrix algebra.

A system of $m \times n$ elements from a field F arranged in a rectangular formation along m rows and n columns and bounded by the brackets ( ) is called an $m \times n$ *matrix*. Usually, a matrix is written by a single capital letter. Thus,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

is an $m \times n$ matrix. In short, we often write $A = (a_{ij})$, where it is understood that the suffix $i = 1, \cdots, m$ and $j = 1, \cdots, n$, and ij indicates the i-th row and the j-th column. The numbers $(A)_{ij} = a_{ij}$ are called the *elements* of the matrix A. For example, the following matrices A and B are of order 2×3 and 3×2,

$$A = \begin{pmatrix} 3 & 5 & 7 \\ 1 & 4 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 1+i & 1-i \\ 2+3i & 2-5i \\ 7 & 5+3i \end{pmatrix}, \quad i = \sqrt{-1}.$$

A matrix having a single row, i.e., $m = 1$, is called a *row matrix* or a *row vector*, e.g., $(2 \;\; 3 \;\; 5 \;\; 7)$.

A matrix having a single column, i.e., $n = 1$, is called a *column matrix* or a *column vector*, e.g.,

$$\begin{pmatrix} 5 \\ 7 \\ 3 \end{pmatrix}.$$

Thus the columns of the matrix A can be viewed as vertical m-tuples (see

Example 1.2), and the rows as horizontal n-tuples. Hence, if we let

$$a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad j = 1, 2, \cdots, n$$

then the above matrix A can be written as

$$A = (a^1, a^2, \cdots, a^n).$$

A matrix having n rows and n columns is called a *square matrix* of order n, e.g.,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} \tag{2.1}$$

is a square matrix of order 3.

For a square matrix A of order n, the elements $a_{ii}, \ i = 1, \cdots, n$, lying on the *leading* or *principal diagonal* are called the *diagonal elements* of A, whereas the remaining elements are called the *off-diagonal* elements. Thus for the matrix A in (2.1) the diagonal elements are $1, 3, 5$.

A square matrix all of whose elements except those in the principal diagonal are zero, i.e., $a_{ij} = 0, \ |i - j| \geq 1$ is called a *diagonal matrix*, e.g.,

$$A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A diagonal matrix of order n that has unity for all its diagonal elements, i.e., $a_{ii} = 1$, is called a *unit* or *identity matrix* of order n and is denoted by I_n or simply by I. For example, identity matrix of order 3 is

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and of nth order $I_n = (e^1, e^2, \cdots, e^n)$.

If all the elements of a matrix are zero, i.e., $a_{ij} = 0$, it is called a *null* or *zero matrix* and is denoted by 0, e.g.,

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

A square matrix $A = (a_{ij})$ is called *symmetric* when $a_{ij} = a_{ji}$. If $a_{ij} = -a_{ji}$, so that all the principal diagonal elements are zero, then the matrix is