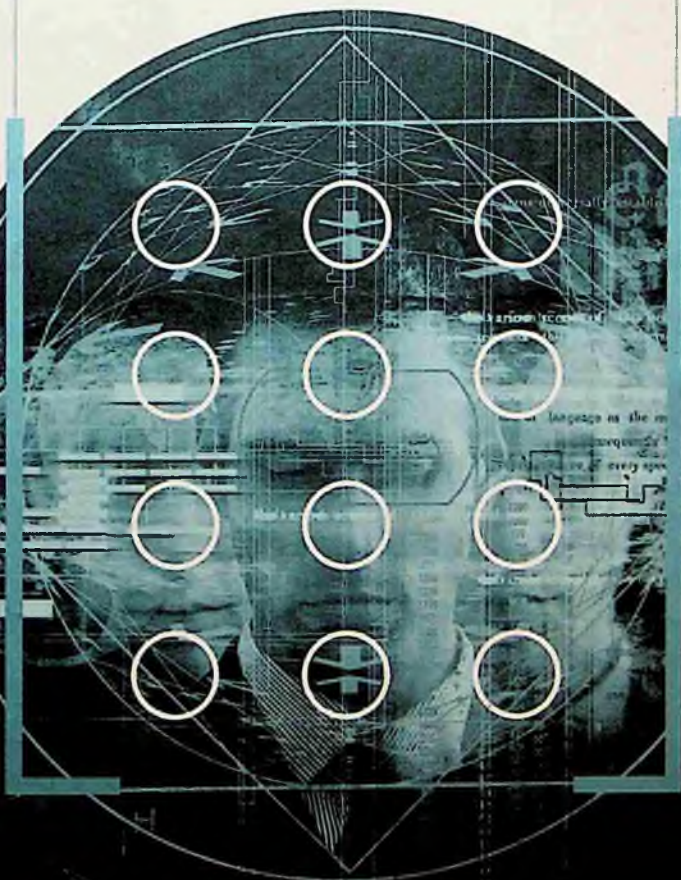


Series on Concrete and Applicable Mathematics – Vol. 3

Ferenc Szidarovszky • Sándor Molnár

Introduction to Matrix Theory

with applications to
business and economics



World Scientific

Introduction to
Matrix Theory
with applications to
business and economics

SERIES ON CONCRETE AND APPLICABLE MATHEMATICS

Series Editor: Professor George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis, TN 38152, USA

Published

- Vol. 1 Long Time Behaviour of Classical and Quantum Systems
edited by S. Graffi & A. Martinez
- Vol. 2 Problems in Probability
by T. M. Mills
- Vol. 3 Introduction to Matrix Theory
With Applications to Business and Economics
by F. Szidarovszky & S. Molnár
- Vol. 4 Stochastic Models with Applications to Genetics, Cancers, Aids and
Other Biomedical Systems
by Tan Wai-Yuan

Series on Concrete and Applicable Mathematics – Vol. 3

Introduction to
Matrix Theory
with applications to
business and economics

Ferenc Szidarovszky
University of Arizona, USA

Sándor Molnár
Szent Istvan University, Hungary

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Szidarovszky, Ferenc.

Introduction to matrix theory ; with applications to business and economics / Ferenc Szidarovszky, Sándor Molnár.

p cm. -- (Series on concrete and applicable mathematics ; vol. 3)

Includes bibliographical references and index.

ISBN 9810245041 (alk. paper) -- ISBN 9810245130 (pbk)

I. Matrices. I. Molnár, Sándor, Dr. II. Title. III. Series.

QA188 .S95 2001

512.9'434--dc21

2001046664

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Copyright © 2002 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.



Preface

Linear algebra and matrix theory are among the most important and most frequently applied branches of mathematics. They are especially important in economic models, where either the model is assumed linear, or the nonlinear model is approximated by a linear model, and the resulting linear model is examined.

This book is mainly a textbook, that covers a one-semester upper division course or a two-semester lower division course on the subject. The book is written for students studying economics and business, however it can also be used in courses offered by the mathematics department, or any kind of engineering departments.

Each chapter consists of three major parts. The first part introduces the new concepts, discusses and proves the main theorems. The new concepts and theoretical results are always illustrated by easy-to-follow numerical examples. The second part of each chapter presents some applications of the material of the chapter. We have selected these applications from special methodology of linear algebra and matrix theory (such as block matrices, matrix exponential, singular value decomposition, pseudoinverses, etc.), linear systems theory (such as discrete and continuous systems), statistics (for example, the least squares method), numerical analysis (such as interpolation polynomials, integral equations), as well as economic modelling (for example, oligopoly, and producer-consumer models). The last section of each chapter offers exercises to improve understanding of the material and to help students to gain experience in problem solving. In each set of exercises we have presented some simple examples which can be solved easily by using the methodology of that chapter, however in each set we also offer some more difficult problems which require deeper understanding and skill in mathematical developments.

The book is organized as follows. Chapter 1 introduces the concept of vectors and matrices, and discusses the elements of matrix algebra. Vector spaces, subspaces, linear independence, basis, and inner-product spaces are examined in Chapter 2. The most important application of linear algebra is

the solution of systems of linear algebraic equations. In Chapter 3 the elimination method is introduced and we demonstrate how to use this method to determine inverses of square matrices. Determinants and their main properties are discussed in Chapter 4. Linear mappings, linear transformations, the vector space of linear mappings, and matrix representations are investigated in Chapter 5. The discussion on diagonal, triangular, and Jordan canonical forms of matrices is based on the theory of eigenvalues and eigenvectors as well as on the main properties of invariant subspaces. The fundamentals of these topics are given in Chapter 6, and special matrices are introduced and examined in Chapter 7. Here we discuss the special properties of diagonal, tridiagonal, triangular, selfadjoint, unitary, and normal matrices, and introduce different kinds of definite matrices and related special matrix classes (such as quasi-definite, quasi-semi-definite, N-, P-, N-P, and P-N matrices). The last chapter of the book introduces and discusses the elements of matrix analysis including vector and matrix norms, and related topics. In a one-semester introductory course we suggest to cover Chapters 1, 2, 3 and large part of Chapters 6. In a one-semester higher level (or in a second) course on the subject we suggest to cover Chapters 4, 5, part of 6, 7, and 8.

It is our pleasure to acknowledge encouragement we have received from our colleagues and friends in both Hungary and the USA. In preparing the manuscript we obtained significant help from our students, and colleagues. Our special thanks should be addressed to Dr. Zoltan Varga for providing us with helpful critique, and to Mark Molnar for all the editorial help we needed in finalizing the manuscript. Finally we mention that it was through various grants from the National Science Foundation and the Joint U.S.–Hungarian Research and Technology Fund that the authors were able to enjoy the continuation of their productive and happy collaboration.

Ferenc Szidarovszky

Tucson, Arizona, USA

Sándor Molnár

Budapest, Hungary

Contents

Chapter 1. Vectors and Matrices	1
1.1 Introduction	1
1.2 Comparison of Matrices	7
1.3 Elementary Matrix Algebra	8
1.4 Inverse of a Matrix	21
1.5 Further Examples and Applications	24
1.6 Exercises.....	52
Chapter 2. Vector Spaces and Inner-Product Spaces.....	57
2.1 Introduction	57
2.2 Subspaces	62
2.3 Linear Independence, Basis.....	70
2.4 Inner-Product Spaces.....	85
2.5 Direct Sums and Orthogonal Complementary Subspaces.....	105
2.6 Applications.....	119
2.7 Exercises.....	127
Chapter 3. Systems of Linear Equations and Inverses of Matrices	133
3.1 Introduction	133
3.2 Existence and Uniqueness of a Solution.....	135
3.3 Systems of Homogeneous Linear Equations	142
3.4 Systems of Inhomogeneous Linear Equations.....	149
3.5 Rank of Matrices	155
3.6 Matrix Equations and Inverses of Matrices.....	157
3.7 The Elimination Method.....	165
3.8 Applications.....	188
3.9 Exercises.....	200
Chapter 4. Determinants	207
4.1 Introduction	207
4.2 Properties of Determinants	215

4.3	Cofactors and Expansion by Cofactors	225
4.4	Determinants and Systems of Linear Equations.....	232
4.5	Further Examples and Applications	238
4.6	Exercises.....	245
Chapter 5. Linear Mappings and Matrices		251
5.1	Introduction	251
5.2	Vector Coordinates.....	251
5.3	Linear Mappings.....	254
5.4	The Vector Space of Linear Mappings.....	271
5.5	Multiplication of Linear Mappings, and Inverses	273
5.6	Matrix Representations of Linear Mappings	284
5.7	Coordinates and Matrix Representation in a New Basis	293
5.8	Applications.....	299
5.9	Exercises.....	306
Chapter 6. Eigenvalues, Invariant Subspaces, Canonical Forms.....		313
6.1	Introduction	313
6.2	Basic Concepts	313
6.3	Matrix Polynomials	325
6.4	The Construction of Invariant Subspaces.....	336
6.5	Diagonal and Triangular Forms.....	341
6.6	The Jordan Canonical Form	351
6.7	Complexification.....	360
6.8	Applications.....	361
6.9.	Exercises.....	378
Chapter 7. Special Matrices.....		383
7.1	Introduction	383
7.2	Diagonal, Tridiagonal, and Triangular Matrices	383
7.3	Idempotent and Nilpotent Matrices	393
7.4	Matrices in Inner Product Spaces	396
7.5	Definite Matrices	410
7.6	Nonnegative Matrices.....	418

7.7 Applications.....	437
7.8 Exercises.....	450
 Chapter 8. Elements of Matrix Analysis.....	 455
8.1 Introduction	455
8.2 Vector Norms	455
8.3 Matrix Norms	461
8.4 Applications.....	474
8.5 Exercises.....	491
 References	 495
 Index	 497

Chapter 1

Vectors and Matrices

1.1 Introduction

In modeling and solving problems in engineering, economics, and in any other field of applied sciences there is often a need to present data in an organized way of a rectangular array. For example, if the prices of five items are listed as $(p_1, p_2, p_3, p_4, p_5)$ then such an array is constructed. This array has the specialty that it consists of only one row, therefore it is often called a *row vector*. Assume next, that a small firm produces three kinds of products. If the production levels are denoted by x_1, x_2 , and x_3 , then another type of special array

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

can be constructed, which consists of only one column. Therefore it is sometimes called a *column vector*. Consider again the same small firm and assume that for the next week the management considers two alternative production plans. Let x_1, x_2, x_3 and y_1, y_2, y_3 denote the alternative production volumes. The data can be conveniently summarized in a rectangular array form:

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix}, \quad (1.1)$$

where the rows correspond to the different products, and the columns correspond to the alternative plans. This array consists of 3 rows and 2 columns, therefore it is usually called a 3×2 *matrix* (pronounced “three by

two" matrix). By constructing such arrays a new mathematical structure is developed.

Definition 1.1. For a given pair (m, n) of positive integers, an $m \times n$ matrix is a rectangular array of real (or complex) numbers given as

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

This matrix has m rows and n columns. Instead of saying that this matrix is $m \times n$ we may say that its *type* is $m \times n$. The numbers a_{11}, a_{12}, \dots are called the *elements* or *entries* of the matrix. Notice, that each matrix element has two subscripts. The first subscript indicates the row in which the element is located, and the second subscript shows the column in which the element is placed. The set of all real (or complex) $m \times n$ matrices is denoted by $R^{m \times n}$ (or $C^{m \times n}$), which is the obvious generalization of the usual notation R (or C) for the set of all real (or complex) numbers.

Example 1.1. The type of matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

is $2 \cdot 3$, since it consists of two rows and three columns, furthermore

$$a_{11} = 1, \quad a_{12} = 2, \quad a_{13} = 3,$$

$$a_{21} = 4, \quad a_{22} = 5, \quad a_{23} = 6.$$

◆

Matrices are usually denoted by underscore capital letters such as $\underline{A}, \underline{B}, \underline{C}$, and so on. Sometimes we refer to the matrix element a_{ij} as the (i, j) element or (i, j) entry of the matrix. In some applications it is convenient to use the notation $\underline{A} = (a_{ij})$, when a special emphasis is placed on the matrix elements. If one needs to indicate the type of matrix \underline{A} , then the simple notation $\underline{A}_{m \times n}$ or the slightly more complicated $\underline{A} = (a_{ij})_{m \times n}$ or

$$\underline{A} = (a_{ij})_{i,j=1}^{m,n}$$

can be used, which shows that the value of i (the row-index) is between 1 and m , and the value of j (the column-index) is between 1 and n .

In most applications the rows and columns of matrices refer to certain quantities, parameters, or alternatives. If the prices of different products are summarized in a row vector (as it was done previously), then the columns refer to the different products, and if the production volumes are summarized in a column vector, then the rows refer to the products. Similarly, in the case of matrix (1.1), the rows correspond to the three products and the columns refer to the two production plans. In many cases it is useful to interchange the meanings of the rows and columns. Then a new matrix is constructed in which each column is formed from the elements of the corresponding row placed in the same order. The same result is obtained, when the elements of each column are placed in the corresponding row of the new matrix.

This matrix operation can be formally defined as follows.

Definition 1.2. Let \underline{A} be an $m \times n$ matrix, then the *transpose* of \underline{A} is the $n \times m$ matrix, the (i, j) element of which is a_{ji} . The transpose of \underline{A} is denoted by \underline{A}^T , and this matrix operation is called *transposition*.

Example 1.2. The transpose of a row vector is a column vector:

$$(1,2,3)^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

the transpose of a column vector is a row vector:

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}^T = (4,5,6);$$

and the transpose of an $m \times n$ matrix is an $n \times m$ matrix:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Notice that $(\underline{A}^T)^T = \underline{A}$, since the (i, j) element of \underline{A}^T is a_{ji} , and therefore the (i, j) element of $(\underline{A}^T)^T$ is a_{ij} , which is the (i, j) element of the original matrix \underline{A} . In several cases it is convenient to emphasize in the notation if a matrix is a column vector or a row vector. Column vectors are denoted by underscore lower case letters such as $\underline{a}, \underline{b}, \underline{c}$, and so on. Since row vectors are the transposes of column vectors, they can be denoted as $\underline{a}^T, \underline{b}^T, \underline{c}^T$, and so on.

For an arbitrary matrix of the type $m \times n$, m of course, need not be equal to n . In the important special case of $m = n$, the matrix is called a *square matrix*. The common value of m and n is called the *order* of the matrix. The entries $a_{11}, a_{22}, \dots, a_{mm}$ of a square matrix of order m are called the *diagonal elements*, and they form the main *diagonal* or simply the diagonal of the matrix.

Example 1.3. Matrix

$$\begin{pmatrix} 1 & 5 & 6 \\ 7 & 2 & 8 \\ 9 & 10 & 3 \end{pmatrix}$$

is a square matrix of order 3, and the elements 1, 2, and 3 form the diagonal of the matrix.

A matrix composed entirely of zeros is called the *zero* (or *null*) *matrix*, and a vector of zeros is called a *zero* (or *null*) *vector*. A zero matrix is denoted by $\underline{0}$, and a zero column (or row) vector is denoted by $\underline{0}$ (or $\underline{0}^T$).

A square matrix with all off-diagonal elements equal to zero is a *diagonal* matrix. A special diagonal matrix, where all diagonal elements are equal to one, is called the *identity* matrix. The $n \times n$ identity matrix is usually denoted by \underline{I}_n . A square matrix in which all elements below the diagonal are equal to zero is called *upper triangular*, and similarly, a square matrix with all zero elements above the diagonal is called *lower triangular*.

Example 1.4. Consider matrices

$$\underline{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \underline{B} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \underline{C} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix},$$

then \underline{A} is diagonal, \underline{B} is upper triangular, and \underline{C} is lower triangular. Notice that \underline{A} (like any other diagonal matrix) satisfies the definitions of both upper and lower triangular matrices, therefore it is also an upper and a lower triangular.

Notice that the transpose of an upper triangular matrix is lower triangular, and the transpose of a lower triangular matrix is upper triangular. It is easy to see that the transpose of a diagonal matrix is itself.

In economic theory, triangular matrices have a special meaning. Consider an $n \times n$ real square matrix \underline{A} , and assume that element a_{ij} represents the effect of unit i towards unit j , where the rows and columns of the matrix correspond to some economic units (for example, sectors in input-output models). In the case of an upper triangular matrix, $a_{ij}=0$ for $i > j$; and for lower triangular matrices $a_{ij}=0$ for $i < j$. That is, the zero matrix elements indicate that there is no effect from higher (or lower) indexed units to lower (or higher) indexed units showing a very strict hierarchy between the economic units.

Definition 1.3. An $n \times n$ matrix \underline{A} is called *decomposable* if there is a nonempty proper subset J of $\{1, 2, \dots, n\}$ such that

$$a_{ij}=0 \text{ for } i \notin J \text{ and } j \in J.$$

An $n \times n$ real matrix is called *indecomposable*, if it is not decomposable and is not the 1×1 zero matrix.

It is easy to see that a matrix is decomposable if and only if its transpose is decomposable. Any decomposable matrix can be transformed into the special form

$$\begin{pmatrix} \underline{A}_{11} & \underline{A}_{12} \\ \underline{O} & \underline{A}_{22} \end{pmatrix}$$

by interchanging its rows and columns, where \underline{A}_{11} is $k \times k$, \underline{A}_{12} is $k \times (n-k)$, \underline{A}_{22} is an $(n-k) \times (n-k)$ matrix, \underline{O} is the $(n-k) \times k$ zero matrix, and set J becomes $\{1, 2, \dots, k\}$. In terms of the above economic interpretation the zero block indicates that there is no effect from the units not belonging to J towards the units of J .

Definition 1.4. A square matrix \underline{A} is called *symmetric* if $\underline{A}^T = \underline{A}$, and it is called *skew-symmetric* if $\underline{A}^T = -\underline{A}$.

Notice that an $n \times n$ matrix is symmetric if and only if for $k = 1, 2, \dots, n$, its k^{th} columns have the same elements in the same order. As a special case, all diagonal matrices are symmetric. The diagonal of a skew symmetric matrix consists of zeros.

Example 1.5. Matrix

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

is a symmetric 3×3 matrix. ♦

1.2 Comparison of Matrices

Matrices \underline{A} and \underline{B} are equal if they have the same type and the corresponding elements in the two matrices are equal. If $\underline{A} = (a_{ij})$ and $\underline{B} = (b_{ij})$, then $\underline{A} = \underline{B}$ if and only if $a_{ij} = b_{ij}$ for all i and j .

Similarly we say that for real matrices \underline{A} and \underline{B} , $\underline{A} \leq \underline{B}$, if they have the same type and for all i and j , $a_{ij} \leq b_{ij}$. Analogously, $\underline{A} < \underline{B}$ if they have the same type and for all i and j , $a_{ij} < b_{ij}$.

Notice that matrices can be compared in the above sense only if they have the same type.

Example 1.6. Let

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \underline{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \underline{C} = \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix}, \underline{D} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix},$$

then for example, $\underline{A} = \underline{B}, \underline{A} \leq \underline{C}, \underline{A} < \underline{D}$. If one defines

$$\underline{E} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

then it cannot be compared to any of matrices \underline{A} , \underline{B} , \underline{C} , or \underline{D} .

◆

In comparing matrices an important comment has to be made. If a and b are two real numbers, then exactly one of the relations $a = b$, $a < b$, and $a > b$ holds. That is, any two real numbers can be compared in this way. However real matrices may not be compared even if they have the same type. For example, row vectors $\underline{a}^T = (1, 2)$ and $\underline{b}^T = (2, 1)$ cannot be compared, since they are different, in the first element \underline{a}^T is smaller but in the second element \underline{b}^T is smaller. This phenomenon plays an important role in many fields of applied sciences. For example, in single objective optimization we are looking for a best solution, since any two values of the objective functions can be compared; however in the case of optimizing for multiple objectives we are looking for so-called efficient solutions, when none of the target function values can be improved without worsening another.

1.3 Elementary Matrix Algebra

In this section matrix operations will be introduced, and their main properties will be discussed.

Definition 1.5. Let \underline{A} be an $m \times n$ real (or complex) matrix and let a be a real (or complex) number. The product $a\underline{A}$ is defined as the $m \times n$ matrix with (i, j) element $a \times a_{ij}$. That is, each element of the matrix is multiplied by a .

This definition can be briefly written as $a\underline{A} = (a \times a_{ij})_{m \times n}$.

As a simple example assume that the elements of a matrix represent cost data, and each element is given in dollars. If someone wants to change the

dimension of the data to \$1000, then each matrix element has to be multiplied by the same constant 0.001.

Example 1.7. For a numerical example assume that $a = 3$ and

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \text{ then } a\underline{A} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}.$$

Notice, if $a = 0$, then $a\underline{A}$ is the zero matrix for all \underline{A} , since each matrix element is multiplied by zero. The real (or complex) number a is sometimes called a *scalar*, and this matrix operation is called the *multiplication by scalars*.

Definition 1.6. The *sum* of matrices \underline{A} and \underline{B} is defined whenever \underline{A} and \underline{B} have the same type. Each element of $\underline{A} + \underline{B}$ equals the sum of the two corresponding elements of \underline{A} and \underline{B} . In other words, $\underline{A} + \underline{B}$ is the matrix the (i, j) element of which is $a_{ij} + b_{ij}$ for all i, j , where a_{ij} and b_{ij} are the (i, j) elements of \underline{A} and \underline{B} , respectively. The *difference* matrix $\underline{A} - \underline{B}$ is analogously defined to be the matrix with (i, j) elements $a_{ij} - b_{ij}$.

We can summarize this definition as

$$\underline{A} + \underline{B} = (a_{ij} + b_{ij})_{m \times n} \quad \text{and} \quad \underline{A} - \underline{B} = (a_{ij} - b_{ij})_{m \times n}.$$

Example 1.8. Matrices

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \underline{B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

cannot be added or subtracted, since they have different types. However

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 4 & 7 \end{pmatrix},$$

and

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

The above matrix operations satisfy the following properties:

(a) If \underline{A} and \underline{B} have the same type, then

$$\underline{A} + \underline{B} = \underline{B} + \underline{A}. \quad (1.2)$$

That is, matrix addition is *commutative*, which is a simple consequence of the fact that in adding matrices we add the corresponding matrix elements, and the addition of real (or complex) numbers is commutative.

(b) If \underline{A} , \underline{B} , and \underline{C} have the same type, then

$$(\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C}). \quad (1.3)$$

That is, matrix addition is *associative*. This property is also the simple consequence of the associativity of the addition of real (or complex) numbers. If $\underline{A}_1, \underline{A}_2, \dots, \underline{A}_k$ are real (or complex) matrices of the same type, then their sum is defined by the recursion $\underline{S}_1 = \underline{A}_1$ and $\underline{S}_i = \underline{S}_{i-1} + \underline{A}_i$ for $i = 2, 3, \dots, k$. Then $\underline{A}_1 + \underline{A}_2 + \dots + \underline{A}_k = \underline{S}_k$.

(c) Let \underline{A} be any matrix, and \underline{O} be the zero matrix of the same order. Then

$$\underline{A} + \underline{O} = \underline{A}. \quad (1.4)$$

(d) If \underline{A} and \underline{B} have the same type, then with any scalar a ,

$$a(\underline{A} + \underline{B}) = a\underline{A} + a\underline{B}; \quad (1.5)$$

and if a and b are two scalars, and \underline{A} is any matrix, then

$$(a + b)\underline{A} = a\underline{A} + b\underline{A}. \quad (1.6)$$

These two equations are called the *distributivity* properties. Equation (1.5) shows distributivity with respect to the addition of matrices, and (1.6) is the distributive property with respect to the addition of scalars.

(e) If \underline{A} and \underline{B} have the same type, then

$$(\underline{A} + \underline{B})^T = \underline{A}^T + \underline{B}^T. \quad (1.7)$$

Assume that both \underline{A} and \underline{B} are $m \times n$, then both sides of this equation are $n \times m$ matrices, and the (i, j) elements are the same: $a_{ji} + b_{ji}$ for all i and j .

(f) For positive integers k ,

$$\underline{A} + \underline{A} + \dots + \underline{A} = k\underline{A}, \quad (1.8)$$

where on the left-hand side we have k terms, each of them equals the same matrix \underline{A} .

(g) For all matrices \underline{A} and \underline{B} of the same type,

$$\underline{A} - \underline{B} = \underline{A} + (-1) \cdot \underline{B}, \quad (1.9)$$

since for all i and j , $a_{ij} - b_{ij} = a_{ij} + (-1) \cdot b_{ij}$.

Multiplication of matrices will be defined next. For simplifying the discussion, particular cases will be introduced before presenting the general definition.

As the first special case we define the product of row vectors by column vectors. Let $\underline{a}^T = (a_i)$ and $\underline{b} = (b_j)$ be a row and a column vector, respectively. The product $\underline{a}^T \underline{b}$ is defined only when the two vectors have the same number of elements, and in this case

$$\underline{a}^T \underline{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i,$$

where n is the common “length” of the vectors.

Example 1.9. The product

$$(1,2,3) \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

cannot be defined, but

$$(1,2,3) \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = 1 \times 2 + 2 \times 3 + 3 \times 4 = 20.$$

Notice that $\underline{a}^T \underline{b}$ is always a scalar, which can also be considered as a 1×1 matrix. This multiplication can be illustrated and explained by the simple economic example, when a firm produces 3 items, the sale prices of which form the row vector $\underline{p}^T = (p_1, p_2, p_3)$, and the produced quantities are given in a column vector

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Then the revenue (or total sale value) by selling the products is given by the product

$$\underline{p}^T \underline{x} = p_1 x_1 + p_2 x_2 + p_3 x_3.$$

Assume next that \underline{A} is an $m \times n$ matrix, and \underline{x} is a column vector. The product $\underline{A}\underline{x}$ is defined only if the length of \underline{x} equals the length of the rows of \underline{A} (that is, when \underline{x} has n elements), the product is an m -element column vector the i^{th} entry of which is obtained as the product of the i^{th} row of \underline{A} by the column vector \underline{x} . That is, the i^{th} element of $\underline{A}\underline{x}$ equals $\sum_{j=1}^n a_{ij} x_j$.

Example 1.10. The product

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

cannot be defined, but

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \times 1 + 2 \times 2 \\ 3 \times 1 + 4 \times 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}.$$

Assume next that \underline{x}^T is a row vector and \underline{A} is an $m \times n$ matrix. The product $\underline{x}^T \underline{A}$ is defined only if the length of \underline{x}^T equals the length of the columns of \underline{A} (that is, when \underline{x} has m -elements), the product is an n -element row vector, the i^{th} entry of which is obtained as the product of \underline{x}^T by the i^{th} column of \underline{A} . That is, the i^{th} element of $\underline{x}^T \underline{A}$ equals

$$\sum_{j=1}^m x_j a_{ji}.$$

Example 1.11. The product

$$(1, 2, 3) \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

cannot be defined, but

$$(1, 2) \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = (1 \times 1 + 2 \times 3, 1 \times 2 + 2 \times 4) = (7, 10)$$

We are ready now to consider the general case of matrix multiplication.

Definition 1.7. Let \underline{A} and \underline{B} be two matrices. The product \underline{AB} can be defined only if the rows of \underline{A} have the same length as the columns of \underline{B} , and then the (i, j) element of \underline{AB} equals the product of the i^{th} row of \underline{A} by the j^{th} column of \underline{B} . If \underline{A} is $m \times n$ and \underline{B} is $p \times q$, then \underline{AB} is defined only if $n = p$, its type is $m \times q$, and its (i, j) element is obtained as

$$\sum_{k=1}^n a_{ik} b_{kj}.$$

Example 1.12. The product

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

is not defined, but

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ 11 & -11 \end{pmatrix},$$

since

$$(1,2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \times 1 + 2 \times 2 = 5,$$

$$(1,2) \begin{pmatrix} -1 \\ -2 \end{pmatrix} = 1 \times (-1) + 2 \times (-2) = -5,$$

$$(3,4) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 3 \times 1 + 4 \times 2 = 11,$$

and

$$(3,4) \begin{pmatrix} -1 \\ -2 \end{pmatrix} = 3 \times (-1) + 4 \times (-2) = -11.$$

♦

As we have seen before, a row vector can be multiplied by a column vector only if they have the same length, and the product is always a scalar. However a column vector can always be multiplied by a row vector even if they have different lengths, and the product is always a matrix, which is called a *dyad*. If \underline{x} is an m -element column vector and \underline{y}^T is an n -element row vector, then

$$\underline{x}\underline{y}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} (y_1, y_2, \dots, y_n) = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \dots & \dots & \dots & \dots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{pmatrix},$$

since the i^{th} row of \underline{x} is the scalar x_i and the j^{th} column of \underline{y}^T is the number y_j , and their product is $x_i y_j$.

If one selects two arbitrary matrices \underline{A} and \underline{B} , then in their multiplication he/she should face one of the following possibilities:

(i) Neither \underline{AB} nor \underline{BA} exists. Such an example is provided by matrices

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \underline{B} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

(ii) Exactly one of the products \underline{AB} and \underline{BA} exists. For example, select

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \text{ and } \underline{B} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

then

$$\underline{AB} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 3 \\ 2 & 1 & 2 \end{pmatrix},$$

but

$$\underline{BA} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

cannot be defined, since the rows of \underline{B} have 3 elements and the columns of \underline{A} have only 2.

(iii) Both \underline{AB} and \underline{BA} exist, but they have different types. As an example, select

$$\underline{A} = (1,1) \text{ and } \underline{B} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

then

$$\underline{AB} = (1,1) \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 1 \times 2 + 1 \times 2 = 4$$

is a scalar and

$$\underline{BA} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} (1,1) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

is a dyad.

(iv) Both \underline{AB} and \underline{BA} exist, they have the same type but the products are not equal. Select

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \text{ and } \underline{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

then

$$\underline{AB} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}$$

and

$$\underline{BA} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This example also shows a different problem in matrix multiplication. In the second case \underline{BA} is the zero matrix, however neither \underline{A} nor \underline{B} is zero, and furthermore there is no zero element in \underline{A} or \underline{B} . This shows a different phenomenon from the one we used to have and apply in the case of real numbers. The product of real (or complex) numbers is zero only if at least one of the factors equals zero. This idea is used when one solves real equations by factorization. Unfortunately this method cannot be used in solving matrix equations.

(v) Both \underline{AB} and \underline{BA} exist, they have the same type and are equal. Such a special case can be illustrated by matrices

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } \underline{B} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix},$$

when

$$\underline{AB} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}$$

and

$$\underline{BA} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}.$$

Hence, the multiplication of matrices is not a commutative operation in general. However it satisfies the following properties:

(a) If the product $(\underline{AB}) \cdot \underline{C}$ exists, then $\underline{A} \cdot (\underline{BC})$ also exists and

$$(\underline{AB}) \cdot \underline{C} = \underline{A} \cdot (\underline{BC}). \quad (1.10)$$



That is, matrix multiplication is an *associative* operation. This property can be proved as follows. Introduce the notation $\underline{D} = \underline{AB}$ and $\underline{E} = \underline{BC}$, then the (i, j) element of \underline{D} is given as

$$d_{ij} = \sum_l a_{il} b_{lj},$$

and therefore the (i, j) element of the left-hand side of (1.10) is as follows:

$$\sum_k d_{ik} c_{kj} = \sum_k \left(\sum_l a_{il} b_{lk} \right) c_{kj}.$$

On the other hand, the (i, j) element of \underline{E} is

$$e_{ij} = \sum_k b_{ik} c_{kj},$$

and therefore the (i, j) element of the right-hand side of (1.10) equals

$$\sum_l a_{il} e_{lj} = \sum_l a_{il} \left(\sum_k b_{lk} c_{kj} \right).$$

Since a_{il} does not depend on the summation variable k , this expression gives the same value as the (i, j) element of the left-hand side of (1.10).

(b) If one of the matrices $(\underline{A} + \underline{B}) \cdot \underline{C}$ and $\underline{AC} + \underline{BC}$ exists, then the other matrix is also defined and they are equal:

$$(\underline{A} + \underline{B}) \cdot \underline{C} = \underline{AC} + \underline{BC}. \quad (1.11)$$

Similarly, if one of the matrices $\underline{A} \cdot (\underline{B} + \underline{C})$ and $\underline{AB} + \underline{AC}$ exists, then the other matrix is also defined and they are equal:

$$\underline{A} \cdot (\underline{B} + \underline{C}) = \underline{AB} + \underline{AC}. \quad (1.12)$$

These properties show that matrix multiplication is *distributive*, and their proof is similar to the one presented above for associativity.

(c) If \underline{A} is $m \times n$, and \underline{I}_m and \underline{I}_n denote the $m \times m$ and $n \times n$ identity matrices, respectively, then

$$\underline{I}_m \cdot \underline{A} = \underline{A} \quad \text{and} \quad \underline{A} \cdot \underline{I}_n = \underline{A}. \quad (1.13)$$

These properties can also be proved in an easy way, the proofs are left as an exercise. Relations (1.13) show that multiplying by identity matrices leaves matrices unchanged. The same holds for real (or complex) numbers, when we multiply them by 1. Therefore identity matrices can be considered as the matrix-versions of the real number 1.

(d) If \underline{AB} exists, then $\underline{B}^T \underline{A}^T$ also exists, furthermore

$$(\underline{AB})^T = \underline{B}^T \underline{A}^T. \quad (1.14)$$

A simple proof of this equation can be given by comparing the (i, j) elements of the two sides of the equality. The (i, j) element of \underline{AB} equals $\sum_k a_{ik} b_{kj}$, therefore the (i, j) element of $(\underline{AB})^T$ is obtained by interchanging i and j : $\sum_k a_{jk} b_{ki} = \sum_k b_{ki} a_{jk}$.

This is the (i, j) element of the right-hand side, since b_{ki} is the (i, k) element of \underline{B}^T , and a_{jk} is the (k, j) element of \underline{A}^T .

(e) For any matrix \underline{A} and zero matrix \underline{O} ,

$$\underline{AO} = \underline{O} \quad \text{and} \quad \underline{OA} = \underline{O} \quad (1.15)$$

assuming that the left-hand sides are defined. Notice that if \underline{A} is $m \times n$ and \underline{O} is $n \times p$, then \underline{AO} is the $m \times p$ zero matrix, and if \underline{A} is $m \times n$ and \underline{O} is $p \times m$, then \underline{OA} is the $p \times n$ zero matrix.

Let $\underline{A}_1, \underline{A}_2, \dots, \underline{A}_k$ be real (or complex) matrices of the types $m_1 \times n_1, m_2 \times n_2, \dots, m_k \times n_k$, respectively, and assume that $n_1 = m_2$, $n_2 = m_3, \dots, n_{k-1} = m_k$. The product of these matrices is defined by the recursion $\underline{P}_1 = \underline{A}_1$ and $\underline{P}_i = \underline{P}_{i-1} \underline{A}_i$ for $i = 2, 3, \dots, k$ and letting $\underline{A}_1 \cdot \underline{A}_2 \cdot \dots \cdot \underline{A}_k = \underline{P}_k$. In the special case, when \underline{A} is an $n \times n$ square matrix we may select $\underline{A}_1 = \underline{A}_2 = \dots = \underline{A}_k = \underline{A}$, and the product $\underline{A} \cdot \underline{A} \cdot \dots \cdot \underline{A}$ can be

simply denoted by \underline{A}^k . For convenience, we define $\underline{A}^0 = \underline{I}_n$ for all $n \times n$ square matrices \underline{A} .

Example 1.13. Select

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix},$$

then

$$\begin{aligned} \underline{A}^2 &= \underline{A}\underline{A} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix}, \\ \underline{A}^3 &= \underline{A}^2 \underline{A} = \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 9 & 18 \\ 9 & 18 \end{pmatrix}, \end{aligned}$$

and so on.

Assume next that

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$$

is a single-variable polynomial with real (or complex) coefficients, and \underline{A} is an $n \times n$ square matrix. The *matrix-polynomial* $p(\underline{A})$ is defined as

$$p(\underline{A}) = a_0 \underline{I}_n + a_1 \underline{A} + a_2 \underline{A}^2 + \dots + a_m \underline{A}^m, \quad (1.16)$$

where \underline{I}_n is the $n \times n$ identity matrix.

Example 1.14. Let $p(x) = 2 + 2x + x^2$ and \underline{A} as in the previous example, then

$$\begin{aligned}
 p(\underline{A}) &= 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 2\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix} \\
 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 5 & 12 \end{pmatrix}.
 \end{aligned}$$

◆

1.4 Inverse of a Matrix

We start this section with the definition of the inverse of a square matrix.

Definition 1.8. Let \underline{A} be an $n \times n$ square matrix. The *inverse* of \underline{A} is the $n \times n$ matrix \underline{X} which satisfies the relation

$$\underline{AX} = \underline{XA} = \underline{I}_n, \quad (1.17)$$

where \underline{I}_n is the $n \times n$ identity matrix. If such an \underline{X} exists, then \underline{A} is called an *invertible* matrix, and the inverse of \underline{A} is denoted by \underline{A}^{-1} .

These relations show that inverse matrices generalize the concept of the reciprocal of a real (or complex) number, since for all real (or complex) $a \neq 0$,

$$aa^{-1} = a\frac{1}{a} = 1 \quad \text{and} \quad a^{-1}a = \frac{1}{a}a = 1.$$

If $a \neq 0$, then a^{-1} exists. Unfortunately, in the case of matrices the situation is more complicated, as it is illustrated in the following example.

Example 1.15. Consider matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

We will now prove that this matrix has no inverse. Assume that it has, and let

$$\underline{A}^{-1} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then relation

$$\underline{A}\underline{A}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{I}_2$$

implies that

$$\begin{aligned} a_{11} + a_{21} &= 1 \\ a_{12} + a_{22} &= 0 \\ a_{11} + a_{21} &= 0 \\ a_{12} + a_{22} &= 1, \end{aligned}$$

where we equated the (1,1), (1,2), (2,1), and (2,2) elements of the left-hand and right-hand sides, respectively. Notice that the first and third equations contradict each other, since $a_{11} + a_{21}$ must not have two different values at the same time. (Similar contradiction is obtained from the second and fourth equations.) In this case $\underline{A} \neq \underline{O}$, and the matrix even has no zero element. If we change the sign of only one element of \underline{A} to get matrix

$$\underline{B} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

then

$$\underline{B}^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

which can be verified by simple calculation:

$$\underline{B}\underline{B}^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\underline{B}^{-1}\underline{B} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

◆

If \underline{A} is the $n \times n$ zero matrix, then $\underline{A}\underline{X} = \underline{O}$ for all $n \times n$ matrices \underline{X} , therefore \underline{A} has no inverse. If $\underline{A} = \underline{I}_n$, then $\underline{A}^{-1} = \underline{I}_n$, since $\underline{I}_n \underline{I}_n = \underline{I}_n$.

At this moment we do not see an easy way to check if a given matrix has an inverse or not. In later chapters of this book we will introduce simple, practical conditions to check the existence of the inverse of a matrix. However we can easily show that in the case of the existence of an inverse of a given square matrix the inverse must be unique. Assume in contrary to the assertion that both \underline{X} and \underline{Y} are inverses of a matrix \underline{A} . Then

$$\underline{X} = \underline{X}\underline{I}_n = \underline{X}(\underline{A}\underline{Y}) = (\underline{X}\underline{A})\underline{Y} = \underline{I}_n \underline{Y} = \underline{Y};$$

hence \underline{X} and \underline{Y} are necessarily equal to each other. It is also easy to see that if both matrices \underline{A} and \underline{B} are invertible $n \times n$ matrices, then the inverse of their product also exists and

$$(\underline{A}\underline{B})^{-1} = \underline{B}^{-1}\underline{A}^{-1}. \quad (1.18)$$

This relation can be verified by simple calculation:

$$(\underline{AB})(\underline{B}^{-1}\underline{A}^{-1}) = \underline{A}(\underline{BB}^{-1})\underline{A}^{-1} = (\underline{AI})\underline{A}^{-1} = \underline{AA}^{-1} = \underline{I},$$

and

$$(\underline{B}^{-1}\underline{A}^{-1})(\underline{AB}) = \underline{B}^{-1}(\underline{A}^{-1}\underline{A})\underline{B} = (\underline{B}^{-1}\underline{I})\underline{B} = \underline{B}^{-1}\underline{B} = \underline{I}.$$

1.5 Further Examples and Applications

In this section some additional examples and applications of matrix algebra will be outlined.

1. Our first example is the *algebra of block matrices*. Assume that the $m \times n$ real (or complex) matrix \underline{A} is divided into blocks as

$$\underline{A} = \begin{pmatrix} \underline{A}_{11} & \underline{A}_{12} & \cdots & \underline{A}_{1s} \\ \underline{A}_{21} & \underline{A}_{22} & \cdots & \underline{A}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{A}_{r1} & \underline{A}_{r2} & \cdots & \underline{A}_{rs} \end{pmatrix},$$

where \underline{A}_{ij} is an $m_i \times n_j$ matrix. It is assumed that

$$m = \sum_{i=1}^r m_i \quad \text{and} \quad n = \sum_{j=1}^s n_j.$$

Suppose that matrix \underline{B} has the same size as \underline{A} , and it is also divided into blocks as

$$\underline{B} = \begin{pmatrix} \underline{B}_{11} & \underline{B}_{12} & \cdots & \underline{B}_{1s} \\ \underline{B}_{21} & \underline{B}_{22} & \cdots & \underline{B}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{B}_{r1} & \underline{B}_{r2} & \cdots & \underline{B}_{rs} \end{pmatrix},$$

where for all i and j , blocks \underline{A}_{ij} and \underline{B}_{ij} have the same size. Since \underline{A} and \underline{B} are added and subtracted element-wise,

$$\underline{A} + \underline{B} = \begin{pmatrix} \underline{A}_{11} + \underline{B}_{11} & \underline{A}_{12} + \underline{B}_{12} & \cdots & \underline{A}_{1s} + \underline{B}_{1s} \\ \underline{A}_{21} + \underline{B}_{21} & \underline{A}_{22} + \underline{B}_{22} & \cdots & \underline{A}_{2s} + \underline{B}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{A}_{r1} + \underline{B}_{r1} & \underline{A}_{r2} + \underline{B}_{r2} & \cdots & \underline{A}_{rs} + \underline{B}_{rs} \end{pmatrix}, \quad (1.19)$$

and

$$\underline{A} - \underline{B} = \begin{pmatrix} \underline{A}_{11} - \underline{B}_{11} & \underline{A}_{12} - \underline{B}_{12} & \cdots & \underline{A}_{1s} - \underline{B}_{1s} \\ \underline{A}_{21} - \underline{B}_{21} & \underline{A}_{22} - \underline{B}_{22} & \cdots & \underline{A}_{2s} - \underline{B}_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{A}_{r1} - \underline{B}_{r1} & \underline{A}_{r2} - \underline{B}_{r2} & \cdots & \underline{A}_{rs} - \underline{B}_{rs} \end{pmatrix}. \quad (1.20)$$

If the sizes of the corresponding blocks of \underline{A} and \underline{B} are different, we cannot add or subtract the corresponding blocks, since their sum and difference are defined only if they have the same size.

Example 1.16. Let

$$\underline{A} = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & -1 & 1 & 2 \end{pmatrix}$$

be divided into four 2×2 blocks, and assume that

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

is divided also into four blocks, but their sizes are 1×2 , 1×2 , 3×2 , 3×2 , respectively. Notice that both $\underline{A} + \underline{B}$ and $\underline{A} - \underline{B}$ exist, however neither of them can be obtained by adding or subtracting the corresponding blocks. ◆

Matrices can also be multiplied block-wise if the division of both matrices into blocks satisfy certain compatibility conditions. Assume that \underline{A} is $m \times n$ and it is divided into blocks as before. Assume that \underline{B} is an $n \times p$ matrix with block-form

$$\underline{B} = \begin{pmatrix} \underline{B}_{11} & \underline{B}_{12} & \cdots & \underline{B}_{1t} \\ \underline{B}_{21} & \underline{B}_{22} & \cdots & \underline{B}_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{B}_{s1} & \underline{B}_{s2} & \cdots & \underline{B}_{st} \end{pmatrix}$$

where the size of block \underline{B}_{ij} is $n_i \times p_j$ with $p = \sum_{j=1}^t p_j$. For $i = 1, 2, \dots, r$ and $j = 1, 2, \dots, t$ define

$$\underline{C}_{ij} = \sum_{l=1}^s \underline{A}_{il} \underline{B}_{lj}, \quad (1.21)$$

which is the “formal product” of the i^{th} block-row

$$(\underline{A}_{i1}, \underline{A}_{i2}, \dots, \underline{A}_{is})$$

of matrix \underline{A} by the j^{th} block-column

$$\begin{pmatrix} \underline{B}_{1j} \\ \underline{B}_{2j} \\ \cdots \\ \underline{B}_{sj} \end{pmatrix}$$

of matrix \underline{B} . Then it is easy to see that the product $\underline{C} = \underline{A} \cdot \underline{B}$ can be divided into blocks as follows:

$$\underline{C} = \begin{pmatrix} \underline{C}_{11} & \underline{C}_{12} & \cdots & \underline{C}_{1r} \\ \underline{C}_{21} & \underline{C}_{22} & \cdots & \underline{C}_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{C}_{r1} & \underline{C}_{r2} & \cdots & \underline{C}_{rr} \end{pmatrix}.$$

Example 1.17. Select

$$\underline{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \underline{B} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

If both matrices are divided into 2×2 blocks as shown above, we have the blocks

$$\underline{A}_{11} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \underline{A}_{12} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \underline{A}_{21} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \underline{A}_{22} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

and

$$\underline{B}_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \underline{B}_{12} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \underline{B}_{21} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \underline{B}_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then the 2×2 blocks of the product $\underline{C} = \underline{A} \cdot \underline{B}$ can be obtained as follows:

$$\begin{aligned}\underline{C}_{11} &= \underline{A}_{11}\underline{B}_{11} + \underline{A}_{12}\underline{B}_{21} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & 2 \end{pmatrix},\end{aligned}$$

$$\begin{aligned}\underline{C}_{12} &= \underline{A}_{11}\underline{B}_{12} + \underline{A}_{12}\underline{B}_{22} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 3 & 3 \end{pmatrix},\end{aligned}$$

$$\begin{aligned}\underline{C}_{21} &= \underline{A}_{21}\underline{B}_{11} + \underline{A}_{22}\underline{B}_{21} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix},\end{aligned}$$

and

$$\begin{aligned}\underline{C}_{22} &= \underline{A}_{21}\underline{B}_{12} + \underline{A}_{22}\underline{B}_{22} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}.\end{aligned}$$

Therefore,

$$\underline{A} \cdot \underline{B} = \underline{C} = \begin{pmatrix} 3 & 3 & 3 & 2 \\ 1 & 2 & 3 & 3 \\ 3 & 3 & 3 & 2 \\ 3 & 3 & 1 & 0 \end{pmatrix}.$$

The direct multiplication of matrices \underline{A} and \underline{B} has to give the same answer, that can be easily checked.

If the division of \underline{A} and \underline{B} into blocks does not satisfy the above conditions then \underline{A} and \underline{B} cannot be multiplied by using blocks even in cases when $\underline{A} \cdot \underline{B}$ exists. In such cases the original definition of matrix multiplication by using the matrix elements can be used.

As the conclusion of this example we will examine inverses of block matrices. Assume that the $n \times n$ matrix \underline{A} is divided into blocks as

$$\underline{A} = \begin{pmatrix} \underline{P} & \underline{Q} \\ \underline{R} & \underline{S} \end{pmatrix},$$

where \underline{P} is $m \times m$, \underline{Q} is $m \times (n-m)$, \underline{R} is $(n-m) \times m$, and \underline{S} is an $(n-m) \times (n-m)$ matrix. We will determine the inverse of \underline{A} in a similar block-form

$$\underline{A}^{-1} = \begin{pmatrix} \underline{X} & \underline{Y} \\ \underline{U} & \underline{V} \end{pmatrix},$$

where \underline{X} is $m \times m$, \underline{Y} is $m \times (n-m)$, \underline{U} is $(n-m) \times m$, and \underline{V} is $(n-m) \times (n-m)$. The definition of inverse matrices implies that

$$\begin{pmatrix} \underline{P} & \underline{Q} \\ \underline{R} & \underline{S} \end{pmatrix} \begin{pmatrix} \underline{X} & \underline{Y} \\ \underline{U} & \underline{V} \end{pmatrix} = \begin{pmatrix} \underline{I}_m & \underline{O} \\ \underline{O} & \underline{I}_{n-m} \end{pmatrix}.$$

Comparing the corresponding blocks of both sides of this equation gives the relations

$$\begin{aligned}
 \underline{P}\underline{X} + \underline{Q}\underline{U} &= \underline{I}_m \\
 \underline{P}\underline{Y} + \underline{Q}\underline{V} &= \underline{O} \\
 \underline{R}\underline{X} + \underline{S}\underline{U} &= \underline{O} \\
 \underline{R}\underline{Y} + \underline{S}\underline{V} &= \underline{I}_{n-m}.
 \end{aligned}
 \tag{1.22}$$

Assuming that \underline{P} is invertible, the second equation implies that

$$\underline{Y} = -\underline{P}^{-1}\underline{Q}\underline{V}, \tag{1.23}$$

and substituting this relation into the fourth equation gives an equation for block \underline{V} :

$$\left(-\underline{R}\underline{P}^{-1}\underline{Q} + \underline{S}\right)\underline{V} = \underline{I}_{n-m},$$

that is, $\underline{S} - \underline{R}\underline{P}^{-1}\underline{Q}$ must be invertible, and

$$\underline{V} = \left(\underline{S} - \underline{R}\underline{P}^{-1}\underline{Q}\right)^{-1}. \tag{1.24}$$

From the third equation of (1.22) we have

$$\underline{U} = -\underline{S}^{-1}\underline{R}\underline{X} \tag{1.25}$$

assuming that \underline{S} is invertible. Substitute this equation into the first equation of (1.22) to see that $\underline{P} - \underline{Q}\underline{S}^{-1}\underline{R}$ must be invertible, and

$$\underline{X} = \left(\underline{P} - \underline{Q}\underline{S}^{-1}\underline{R}\right)^{-1}. \tag{1.26}$$

Notice that equations (1.26), (1.25), (1.24), (1.23) can be used to recover the unknown blocks \underline{X} , \underline{U} , \underline{V} , and \underline{Y} of the inverse matrix A^{-1} .

Example 1.18. We will now invert matrix

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

In this case we may select

$$\underline{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \underline{Q} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \underline{R} = (1,1), \text{ and } \underline{S} = (1).$$

Equation (1.26) implies that

$$\underline{X} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot (1,1) \right)^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

since $\underline{S}^{-1} = (1)$. From equation (1.25) we have

$$\underline{U} = -1 \cdot (1,1) \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = (1,1).$$

Equation (1.24) is then applied to find \underline{V} :

$$\underline{V} = \left(1 - (1,1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)^{-1} = (1-2)^{-1} = \frac{1}{-1} = -1,$$

where we used the fact that the inverse of the identity matrix is itself. And finally, from equation (1.23) we get

$$\underline{Y} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot (-1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Hence,

$$\underline{A}^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

In many applications a *certain part of a matrix* (a row, column, element, or even a block of the matrix) is needed for further computation. In this example of matrix algebra we will show how to obtain such matrix parts by using only matrix operations. Let \underline{A} be a given $m \times n$ real (or complex) matrix with (i, j) element a_{ij} .

Let $\underline{e}_j^{(n)}$ denote the n -element column vector the j^{th} element of which is one, and all other elements are equal to zero. Then

$$\underline{A} \cdot \underline{e}_j^{(n)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^{\text{th}} \text{ element}$$

\uparrow
 j^{th} column

For $i = 1, 2, \dots, m$, the i^{th} element of the product is obtained by multiplying the i^{th} row of \underline{A} by the column vector $\underline{e}_j^{(n)}$:

$$\begin{array}{c}
 \left(a_{i1}, a_{i2}, \dots, a_{ij}, \dots, a_{in} \right) \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^{\text{th}} \text{ element} = a_{ij}, \\
 \uparrow \\
 j^{\text{th}} \text{ element}
 \end{array}$$

since all other terms equal zero. Hence

$$\underline{A} \cdot \underline{e}_j^{(n)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix},$$

which is the j^{th} column of A .

Let now $\underline{e}_i^{(m)r}$ denote the m -element row vector the i^{th} element of which is equal to one and all other elements are zeros. Then

$$\begin{array}{c}
 \underline{e}_i^{(m)r} \underline{A} = (0, 0, \dots, 1, \dots, 0) \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \leftarrow i^{\text{th}} \text{ row.} \\
 \uparrow \\
 i^{\text{th}} \text{ element}
 \end{array}$$

For $j = 1, 2, \dots, n$, the j th element of the product is obtained by multiplying $\underline{e}_i^{(m)T}$ by the j th column of \underline{A} :

$$(0, 0, \dots, 1, \dots, 0) \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{in} \end{pmatrix} \leftarrow i \text{ th element} = a_{ij},$$

↑
 i th element

since all other terms are equal to zero. Therefore

$$\underline{e}_i^{(m)T} \underline{A} = (a_{i1}, a_{i2}, \dots, a_{ij}, \dots, a_{in}),$$

which is the i th row of \underline{A} .

It is also easy to see that for all i and j ,

$$a_{ij} = \underline{e}_i^{(m)T} \underline{A} \underline{e}_j^{(n)},$$

since

$$\underline{e}_i^{(m)T} (\underline{A} \underline{e}_j^{(n)}) = (0, 0, \dots, 1, \dots, 0) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix} \leftarrow i \text{ th element} = a_{ij},$$

↑
 i th element

since all other terms equal zero.

The above relations can be presented in a much more general framework. Let $1 \leq i_1 < i_2 < \dots < i_r \leq m$ and $1 \leq j_1 < j_2 < \dots < j_s \leq n$ be arbitrary integers. Consider matrix

$$\underline{A}_1 = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_s} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_s} \\ \dots & \dots & \dots & \dots \\ a_{i_r j_1} & a_{i_r j_2} & \dots & a_{i_r j_s} \end{pmatrix}$$

which can be obtained from \underline{A} by deleting all rows except rows i_1, i_2, \dots, i_r and all columns except columns j_1, j_2, \dots, j_s . Define the $r \times m$ matrix \underline{U} and the $n \times s$ matrix \underline{V} such that

$$u_{1i_1} = u_{2i_2} = \dots = u_{ri_r} = 1, \quad \text{all other } u_{ij} = 0;$$

and

$$v_{j_1 1} = v_{j_2 2} = \dots = v_{j_s s} = 1, \quad \text{all other } v_{ij} = 0.$$

Then

$$\underline{A}_1 = \underline{U} \underline{A} \underline{V}. \quad (1.27)$$

In the particular case, when \underline{A} is a square matrix, $r=s$, and $i_1 = j_1, i_2 = j_2, \dots, i_r = j_r$, then matrix \underline{A}_1 called a *principal submatrix* of \underline{A} . Notice that all principal submatrices are square matrices, and if \underline{A} is $n \times n$, then there are $\binom{n}{r} r \times r$ principal submatrices of \underline{A} .

Using a similar idea as before, some further vector characteristics can be derived, that have significant applications in statistics.

Let x_1, \dots, x_n be sample elements. They can be summarized as a column vector

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

By introducing the n -element row vector

$$\underline{1}^T = (1, 1, \dots, 1)$$

it is easy to see that

$$\underline{1}^T \underline{x} = x_1 + x_2 + \dots + x_n,$$

therefore the sample mean can be obtained as

$$\bar{x} = \frac{1}{n} \underline{1}^T \underline{x}.$$

If we notice that $\underline{1}^T \underline{1} = n$, then we also have

$$\bar{x} = \frac{\underline{1}^T \underline{x}}{\underline{1}^T \underline{1}}.$$

The sample variance can be expressed as

$$S^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2.$$

Notice now that

$$\underline{x} - \bar{x} \cdot \underline{1} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \dots \\ x_n - \bar{x} \end{pmatrix},$$

therefore

$$\begin{aligned}
 S_x^2 &= \frac{1}{n-1} (\underline{x} - \bar{x} \cdot \underline{1})^T (\underline{x} - \bar{x} \cdot \underline{1}) = \frac{(\underline{x} - \bar{x} \cdot \underline{1})^T (\underline{x} - \bar{x} \cdot \underline{1})}{n-1} \\
 &= \frac{\left(\underline{x} - \frac{1^T \underline{x}}{1^T \underline{1}} \cdot \underline{1} \right)^T \left(\underline{x} - \frac{1^T \underline{x}}{1^T \underline{1}} \cdot \underline{1} \right)}{1^T \underline{1} - 1}.
 \end{aligned}$$

Consider next two n -element samples, x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_n . The covariance between these samples can be written as follows:

$$\begin{aligned}
 \text{Cov}(\underline{x}, \underline{y}) &= \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y}) = \frac{1}{n} (\underline{x} - \bar{x} \cdot \underline{1})^T (\underline{y} - \bar{y} \cdot \underline{1}) \\
 &= \frac{\left(\underline{x} - \frac{1^T \underline{x}}{1^T \underline{1}} \right) \left(\underline{y} - \frac{1^T \underline{y}}{1^T \underline{1}} \right)}{1^T \underline{1}}.
 \end{aligned}$$

The correlation between the two samples has the general form

$$r = \frac{\text{Cov}(\underline{x}, \underline{y})}{S_x \cdot S_y},$$

it can also be expressed by using vector operations if we substitute the above expressions for the covariance and the two variances.

Time invariant linear *dynamic economic systems with discrete* time scale can be generally formulated as the difference equation

$$\underline{x}(t+1) = \underline{A}\underline{x}(t) + \underline{b}, \quad (1.28)$$

where the n -element vector \underline{x} is the state variable, \underline{A} is a given $n \times n$ real matrix, and \underline{b} is a given n -element real vector. The initial state $\underline{x}_0 = \underline{x}(0)$ is

also assumed to be known. An elementary problem of systems theory is to find $\underline{x}(t)$ for all future times as easily as possible. In this application we will suggest a solution for this problem. Substitute $t=0, 1$ and 2 into equation (1.28) to see that

$$\begin{aligned}\underline{x}(1) &= \underline{A}\underline{x}(0) + \underline{b} = \underline{A}\underline{x}_0 + \underline{b}, \\ \underline{x}(2) &= \underline{A}\underline{x}(1) + \underline{b} = \underline{A}(\underline{A}\underline{x}_0 + \underline{b}) + \underline{b} = \underline{A}^2 \underline{x}_0 + (\underline{A} + \underline{I})\underline{b}, \\ \underline{x}(3) &= \underline{A}\underline{x}(2) + \underline{b} = \underline{A}(\underline{A}^2 \underline{x}_0 + (\underline{A} + \underline{I})\underline{b}) + \underline{b} = \underline{A}^3 \underline{x}_0 + (\underline{A}^2 + \underline{A} + \underline{I})\underline{b}.\end{aligned}$$

These initial solution vectors suggest that in general,

$$\underline{x}(t) = \underline{A}^t \underline{x}_0 + (\underline{A}^{t-1} + \underline{A}^{t-2} + \dots + \underline{A} + \underline{I})\underline{b} = \underline{A}^t \underline{x}_0 + \left(\sum_{i=0}^{t-1} \underline{A}^i \right) \underline{b} \quad (1.29)$$

where we use the fact that $\underline{I} = \underline{A}^0$. This solution formula can be proved by finite induction. For $t=1, 2$, and 3 the formula is valid as it is shown from the above initial values of the state variable. Assume that the formula is valid for an integer $t > 0$. Then from equation (1.28) we conclude that

$$\underline{x}(t+1) = \underline{A}\underline{x}(t) + \underline{b} = \underline{A} \left(\underline{A}^t \underline{x}_0 + \left(\sum_{i=0}^{t-1} \underline{A}^i \right) \underline{b} \right) + \underline{b} = \underline{A}^{t+1} \underline{x}_0 + \left(\sum_{i=0}^t \underline{A}^i \right) \underline{b},$$

that is, the formula remains valid for $t+1$. Hence it holds for all $t \geq 1$.

In applying the solution formula (1.29), we need a fast method to find powers of \underline{A} . Later, in Chapter 6 we will show a general method for the efficient computation of \underline{A}^t for all $t \geq 1$. In many special cases the power matrix \underline{A}^t can be determined by calculating some initial powers $\underline{A}^2, \underline{A}^3, \underline{A}^4$, and observing the general formula, which has to be then proved by finite induction.

Example 1.19. Consider matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Then

$$\underline{A}^2 = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix} = 3 \cdot \underline{A},$$

$$\underline{A}^3 = \underline{A}^2 \cdot \underline{A} = (3 \cdot \underline{A})\underline{A} = 3 \cdot \underline{A}^2 = 3(3 \cdot \underline{A}) = 3^2 \cdot \underline{A},$$

$$\underline{A}^4 = \underline{A}^3 \cdot \underline{A} = (3^2 \cdot \underline{A})\underline{A} = 3^2 \cdot \underline{A}^2 = 3^2(3 \cdot \underline{A}) = 3^3 \cdot \underline{A}.$$

By using finite induction it is easy to show that in general,

$$\underline{A}^t = 3^{t-1} \underline{A}.$$

We will next solve the difference equation

$$\underline{x}(t+1) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

From equation (1.29) we have

$$\underline{x}(t) = 3^{t-1} \underline{A} \underline{x}_0 + \left(\underline{I} + \left(\sum_{i=1}^{t-1} 3^{i-1} \right) \underline{A} \right) \underline{b} = 3^{t-1} \underline{A} \underline{x}_0 + \left(\underline{I} + \frac{3^{t-1} - 1}{3 - 1} \underline{A} \right) \underline{b}.$$

Notice that for \underline{A}^0 we must not use the general formula of \underline{A}^t since it holds usually only for $t \geq 1$. We have to use the fact that $\underline{A}^0 = \underline{I}$. Substituting the actual form of \underline{A} , \underline{x}_0 , and \underline{b} into the above equation gives the solution:

$$\begin{aligned}\underline{x}(t) &= 3^{t-1} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} (3^{t-1} - 1) \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= 3^{t-1} \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{3^{t-1} - 1}{2} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \cdot 3^{t-1} + \frac{1}{2} \\ 5 \cdot 3^{t-1} - 1 \end{pmatrix}.\end{aligned}$$

◆

Time invariant linear *dynamic economic systems with continuous time scale* can be generally formulated as the differential equation

$$\dot{\underline{x}}(t) = \underline{A} \cdot \underline{x}(t) + \underline{b} \quad (1.30)$$

where \underline{x} is the state variable, matrix \underline{A} is $n \times n$, and vector \underline{b} has n elements. It is well known from linear systems theory (see, for example, Szidarovszky and Bahill, 1992) that the solution is given as

$$\underline{x}(t) = e^{\underline{A}t} \underline{x}_0 + \int_0^t e^{\underline{A}(t-\tau)} \underline{b} d\tau. \quad (1.31)$$

In Chapter 6 we will show a general method to compute matrix exponentials, which are defined as the sum of the infinite series

$$e^{\underline{A}t} = \underline{I} + \frac{\underline{A} \cdot t}{1!} + \frac{\underline{A}^2 t^2}{2!} + \frac{\underline{A}^3 t^3}{3!} + \dots \quad (1.32)$$

It is also well known that this series is convergent for all real t and arbitrary $n \times n$ real matrices. In special cases, a general formula may be derived for \underline{A}^k , and this general formula can be substituted into equality (1.32) to get a closed form for $e^{\underline{A}t}$.

Example 1.20. Consider again the 2×2 matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

In Example 1.19 we have seen that for all $k \geq 1$,

$$\underline{A}^k = 3^{k-1} \cdot \underline{A} = 3^{k-1} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Therefore

$$\begin{aligned} e^{\underline{A}t} &= \underline{I} + \sum_{k=1}^{\infty} \frac{3^{k-1} \underline{A}^k t^k}{k!} = \underline{I} + \underline{A} \cdot \sum_{k=1}^{\infty} \frac{3^{k-1} t^k}{k!} \\ &= \underline{I} + \underline{A} \cdot \frac{1}{3} \sum_{k=1}^{\infty} \frac{(3t)^k}{k!} = \underline{I} + \underline{A} \cdot \frac{1}{3} (e^{3t} - 1) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{e^{3t} - 1}{3} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(e^{3t} + 2) & \frac{1}{3}(e^{3t} - 1) \\ \frac{2}{3}(e^{3t} - 1) & \frac{1}{3}(2e^{3t} + 1) \end{pmatrix}. \end{aligned}$$

We will next solve the initial value-problem

$$\dot{\underline{x}}(t) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

that is, we will derive a formula for $\underline{x}(t)$ at all future times $t > 0$ by assuming that the initial state at $t = 0$ is given. From equation (1.31) we have

$$\begin{aligned}
 \underline{x}(t) &= \begin{pmatrix} \frac{1}{3}(e^{3t} + 2) & \frac{1}{3}(e^{3t} - 1) \\ \frac{2}{3}(e^{3t} - 1) & \frac{1}{3}(2e^{3t} + 1) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
 &+ \int_0^t \begin{pmatrix} \frac{1}{3}(e^{3(t-\tau)} + 2) & \frac{1}{3}(e^{3(t-\tau)} - 1) \\ \frac{2}{3}(e^{3(t-\tau)} - 1) & \frac{1}{3}(2e^{3(t-\tau)} + 1) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} d\tau \\
 &= \begin{pmatrix} \frac{1}{3}(2e^{3t} + 1) \\ \frac{1}{3}(4e^{3t} - 1) \end{pmatrix} + \int_0^t \begin{pmatrix} \frac{1}{3}(e^{3(t-\tau)} + 2) \\ \frac{2}{3}(e^{3(t-\tau)} - 1) \end{pmatrix} d\tau.
 \end{aligned}$$

Notice that

$$\begin{aligned}
 \int_0^t \frac{1}{3}(e^{3(t-\tau)} + 2) d\tau &= \frac{1}{3} \left[\frac{e^{3(t-\tau)}}{-3} + 2\tau \right]_{\tau=0}^t \\
 &= \frac{1}{3} \left[\left(-\frac{1}{3} + 2t \right) - \left(-\frac{1}{3} e^{3t} \right) \right] = \frac{1}{9} (e^{3t} + 6t - 1)
 \end{aligned}$$

and

$$\begin{aligned}
 \int_0^t \frac{2}{3}(e^{3(t-\tau)} - 1) d\tau &= \frac{2}{3} \left[\frac{e^{3(t-\tau)}}{-3} - \tau \right]_{\tau=0}^t \\
 &= \frac{2}{3} \left[\left(-\frac{1}{3} - t \right) - \left(-\frac{1}{3} e^{3t} \right) \right] = \frac{2}{9} (e^{3t} - 3t - 1),
 \end{aligned}$$

therefore

$$\underline{x}(t) = \begin{pmatrix} \frac{1}{9}(7e^{3t} + 6t + 2) \\ \frac{1}{9}(14e^{3t} - 6t - 5) \end{pmatrix}.$$



A special linear dynamic system will be introduced next, which is called the *dynamic Cournot oligopoly model*. Assume that n firms produce a homogeneous good and sell it on the same market. Assume that the market demand function $d(p)$ is decreasing and linear, where p is the selling price. Then its inverse is also decreasing and linear:

$$p(d) = a \cdot d + b \quad (a < 0, b > 0).$$

The constant b shows the high price that the consumers would be willing to pay if the product is not available on the market directly. The coefficient a shows the decrease in the price if the quantity of available products increases by a unit. Assume that the production cost of firm k ($1 \leq k \leq n$) is also linear:

$$C_k(x_k) = b_k x_k + c_k \quad (b_k > 0, c_k \geq 0),$$

where x_k denotes the production level of firm k . Here c_k shows the fixed cost, and b_k is the marginal cost (that is, the additional cost arising by the increase of the production level by a unit). The profit of firm k is therefore given as

$$\varphi_k(x_1, \dots, x_n) = x_k \left(a \left(\sum_{l=1}^n x_l \right) + b \right) - (b_k x_k + c_k).$$

Consider next the following *discrete dynamic extension* of this model. Let $x_1(0), \dots, x_n(0)$ denote the production levels of the firms at the initial time period $t = 0$. Assume that at each further time period $t + 1$ ($t \geq 0$), each firm maximizes its profit under the assumption that the competitors do not

change their production levels from the previous time period. That is, firm k ($k = 1, 2, \dots, n$) maximizes its profit

$$x_k \left(a \left(x_k + \sum_{l \neq k} x_l(t) \right) + b \right) - (b_k x_k + c_k).$$

Assuming interior optimum, simple differentiation shows that the optimal solution for x_k is given as

$$x_k(t+1) = -\frac{1}{2} \sum_{l \neq k} x_l(t) + \frac{b_k - b}{2a}. \quad (1.33)$$

This equation is the special case of the difference equation (1.28) with

$$\underline{A} = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{2} & \cdots & -\frac{1}{2} & -\frac{1}{2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \cdots & -\frac{1}{2} & 0 \end{pmatrix} \quad \text{and} \quad \underline{b} = \begin{pmatrix} \frac{b_1 - b}{2a} \\ \frac{b_2 - b}{2a} \\ \cdots \\ \frac{b_n - b}{2a} \end{pmatrix}.$$

The *continuous dynamic extension* of the model can be formulated in the following way. Assume again that $x_1(0), x_2(0), \dots, x_n(0)$ are the initial production levels. At each time period $t \geq 0$, each firm adjusts its production level proportionally to its marginal profit, that is, for all k ,

$$\dot{x}_k(t) = m_k \cdot \left(2ax_k(t) + a \sum_{l \neq k} x_l(t) + b - b_k \right), \quad (1.34)$$

where $m_k > 0$ is a given constant. Notice that the marginal profit of firm k is the derivative of his profit with respect to x_k , and equation (1.34) requires that x_k increases if the profit increases in x_k , and x_k decreases if the profit decreases in x_k . These equations can be summarized as the differential equation (1.30) with

$$\underline{A} = \begin{pmatrix} m_1 & & & & & \\ & m_2 & & & & \\ & & \ddots & & & \\ & & & m_n & & \\ & & & & & \end{pmatrix} \begin{pmatrix} 2a & a & \dots & a & a \\ a & 2a & \dots & a & a \\ \dots & \dots & \dots & \dots & \dots \\ a & a & \dots & a & 2a \end{pmatrix} \quad \text{and}$$

$$\underline{b} = \begin{pmatrix} m_1(b-b_1) \\ m_2(b-b_2) \\ \dots \\ m_n(b-b_n) \end{pmatrix}.$$

Our next application deals with *dynamic producer-consumer models*. Consider a market where a commodity or a service is supplied by n competing firms. Let x_k denote the output of firm k ($k=1,2,\dots,n$), and assume that $C_k(x_k) = B_k x_k^2 + b_k x_k + c_k$ is its cost function ($B_k, b_k, c_k > 0$). At each time period, firm k maximizes its expected profit given its price prediction $p_k^E(t+1)$. The expected profit of this firm is the following

$$x_k p_k^E(t+1) - (B_k x_k^2 + b_k x_k + c_k)$$

and assuming interior optimum, simple differentiation shows that the profit maximizing output is given as

$$x_k(t+1) = \frac{1}{2B_k} (b_k - p_k^E(t+1)).$$

If firm k believes that the price does not change from time period t to $t+1$, then this firm selects $p_k^E(t+1) = p(t)$ and the profit maximizing output is given as

$$x_k(t+1) = \frac{1}{2B_k} (b_k - p(t)). \quad (1.35)$$

Let $d(t) = Dp(t) + d$ denote the market demand function, from which the price $p(t)$ can be obtained as

$$p(t) = \frac{1}{D} (d(t) - d). \quad (1.36)$$

If we assume that at each time period, the supply equals the demand, then $d(t) = \sum_{l=1}^n x_l(t)$, and substituting this equation and (1.36) into (1.35) the following difference equation is obtained:

$$\begin{aligned} x_k(t+1) &= \frac{1}{2B_k} \left(b_k - \frac{1}{D} \left(\sum_{l=1}^n x_l(t) - d \right) \right) \\ &= \sum_{l=1}^n \frac{1}{2B_k D} x_l(t) + \frac{1}{2B_k} \left(b_k + \frac{d}{D} \right), \end{aligned} \quad (1.37)$$

which is the special case of the difference equations (1.28) with

$$\underline{A} = \begin{pmatrix} \frac{1}{2B_1D} & \frac{1}{2B_1D} & \cdots & \frac{1}{2B_1D} \\ \frac{1}{2B_2D} & \frac{1}{2B_2D} & \cdots & \frac{1}{2B_2D} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{2B_nD} & \frac{1}{2B_nD} & \cdots & \frac{1}{2B_nD} \end{pmatrix}, \quad \text{and} \quad \underline{b} = \begin{pmatrix} \frac{1}{2B_1} \left(b_1 + \frac{d}{D} \right) \\ \frac{1}{2B_2} \left(b_2 + \frac{d}{D} \right) \\ \vdots \\ \frac{1}{2B_n} \left(b_n + \frac{d}{D} \right) \end{pmatrix}.$$

The continuous counterpart of this model can be introduced as follows. Assume that at each time period $t \geq 0$, each firm adjusts its production output proportionally to its expected marginal profit:

$$\dot{x}_k(t) = m_k \left(p_k^E(t) - 2B_k x_k(t) - b_k \right),$$

where $p_k^E(t)$ is its price prediction for time period t . Assuming again that $p_k^E(t) = p(t)$, equation (1.36) holds, $d(t) = \sum_{l=1}^n x_l(t)$, and an ordinary differential equation is obtained for x_k :

$$\begin{aligned} \dot{x}_k(t) &= m_k \left(\frac{1}{D} \left(\sum_{l=1}^n x_l(t) - d \right) - 2B_k x_k(t) - b_k \right) \\ &= \left(\frac{m_k}{D} - 2B_k m_k \right) x_k(t) + \frac{m_k}{D} \sum_{l \neq k} x_l(t) + \left(-\frac{dm_k}{D} - m_k b_k \right) \end{aligned}$$

which is a special case of the differential equation (1.30) with

$$\underline{A} = \begin{pmatrix} m_1 \left(\frac{1}{D} - 2B_1 \right) & \frac{m_1}{D} & \frac{m_1}{D} & \cdots & \frac{m_1}{D} \\ \frac{m_2}{D} & m_2 \left(\frac{1}{D} - 2B_2 \right) & \frac{m_2}{D} & \cdots & \frac{m_2}{D} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{m_n}{D} & \frac{m_n}{D} & \frac{m_n}{D} & \cdots & m_n \left(\frac{1}{D} - 2B_n \right) \end{pmatrix}$$

and

$$\underline{b} = \begin{pmatrix} -m_1 \left(\frac{d}{D} + b_1 \right) \\ -m_2 \left(\frac{d}{D} + b_2 \right) \\ \cdots \\ -m_n \left(\frac{d}{D} + b_n \right) \end{pmatrix}. \quad (1.38)$$

Matrix multiplications can be illustrated by the simple example of a bakery. Assume that it makes three kinds of biscuits. The recipes are summarized in the following table:

	Flour	Sugar	Margarine
Type 1	0.5 (kg)	0.2 (kg)	0.3 (kg)
Type 2	0.55 (kg)	0.25 (kg)	0.2 (kg)
Type 3	0.6 (kg)	0.15 (kg)	0.25 (kg)

The quantities of the ingredients are given for one kilogram of each type. Let p_1, p_2, p_3 denote the prices per kg of the raw materials, then the material cost per kg of each biscuit type can be calculated as

$$\begin{aligned}c_1 &= 0.5p_1 + 0.2p_2 + 0.3p_3 \\c_2 &= 0.55p_1 + 0.25p_2 + 0.2p_3 \\c_3 &= 0.6p_1 + 0.15p_2 + 0.25p_3.\end{aligned}$$

By introducing the notation

$$\underline{A} = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.55 & 0.25 & 0.2 \\ 0.6 & 0.15 & 0.25 \end{pmatrix}, \quad \underline{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \quad \underline{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

these relations can be rewritten as

$$\underline{c} = \underline{A}\underline{p}.$$

Assume next that the company sells three kinds of tins of assorted biscuits as given below:

	Type 1	Type 2	Type 3
Assortment 1	0.5 (kg)	0.7 (kg)	0.8 (kg)
Assortment 2	1.2 (kg)	0.3 (kg)	0.5 (kg)
Assortment 3	1 (kg)	0.5 (kg)	0.5 (kg)

The numbers show the amount of each type of biscuits in each type of assortment. Notice that each assortment contains 2 kg, and the costs of the tins are

$$\begin{aligned}k_1 &= 0.5c_1 + 0.7c_2 + 0.8c_3 \\k_2 &= 1.2c_1 + 0.3c_2 + 0.5c_3 \\k_3 &= c_1 + 0.5c_2 + 0.5c_3.\end{aligned}$$

Let

$$\underline{B} = \begin{pmatrix} 0.5 & 0.7 & 0.8 \\ 1.2 & 0.3 & 0.5 \\ 1 & 0.5 & 0.5 \end{pmatrix}, \text{ and } \underline{k} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix},$$

then we have

$$\underline{k} = \underline{B}\underline{c}.$$

Since $\underline{c} = \underline{A}\underline{p}$, this relation can be rewritten as

$$\underline{k} = \underline{B}\underline{A}\underline{p} = (\underline{B}\underline{A})\underline{p}$$

giving a direct relation between the prices of the ingredients and the costs of the assortments.

Matrices are used in describing directed graphs, which model, for example, material flows or many network problems. A directed graph is defined by a finite set of elements, P_1, P_2, \dots, P_n , together with a finite collection of ordered pairs, (P_i, P_j) , of distinct elements where no ordered pair is repeated. The elements P_1, P_2, \dots, P_n are called the vertices, and the ordered pairs are called the directed edges.

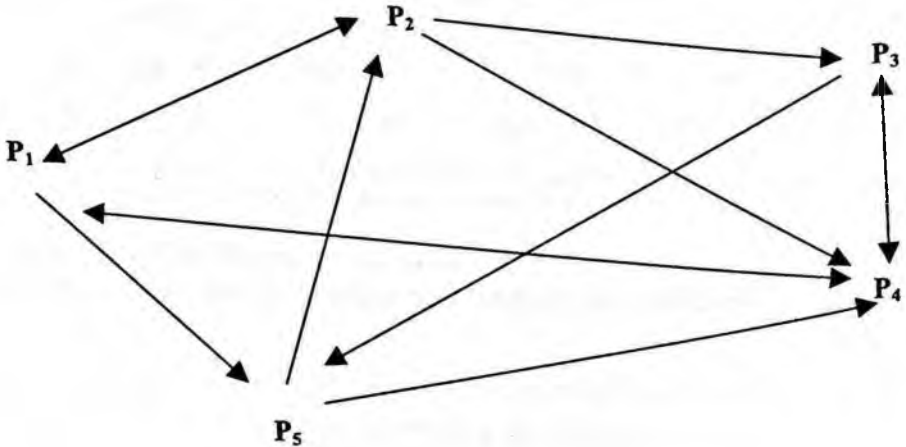


Figure 1.1 A directed graph with vertices

Figure 1.1 shows a directed graph with vertices P_1, \dots, P_5 , and edges (P_1, P_2) , (P_1, P_4) , (P_1, P_5) , (P_2, P_1) , (P_2, P_3) , (P_2, P_4) , (P_3, P_4) , (P_3, P_5) , (P_4, P_3) , (P_4, P_1) , (P_5, P_2) , and (P_5, P_4) . With a directed graph with n vertices, we may associate an $n \times n$ real matrix \underline{A} with elements $a_{ij} = 1$ if (P_i, P_j) is a directed edge, otherwise $a_{ij} = 0$. This matrix is called the vertex matrix of the directed graph. For example, in the previous example we select $n = 5$, and

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

For any square matrix having 0 and 1 elements with all diagonal elements are zero, there exists a directed graph such that the given matrix is its vertex matrix. The elements of the vertex matrix show the direct connections from each vertex to the other vertices.

Let now $r \geq 1$ be a positive integer, and let $a_{ij}^{(r)}$ denote the (i, j) -element of the matrix \underline{A}^r . Then it is easy to show that $a_{ij}^{(r)}$ equals the number of r -step connections from P_i to P_j . In the case of the above matrix,

$$\underline{A}^2 = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 & 2 & 0 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 2 & 2 \\ 2 & 0 & 2 & 1 & 0 \end{pmatrix}.$$

The (1,1)-element with value 2 shows that there are two 2-step connections from vertex 1 to itself. From the graph we see that $P_1 \rightarrow P_2 \rightarrow P_1$ and $P_1 \rightarrow P_4 \rightarrow P_1$ are the two such connections.

The number of at most r -step connections from P_i to P_j is given as $a_{ij} + a_{ij}^{(2)} + \dots + a_{ij}^{(r)}$. For example, the number of at most two-step connections are summarized in matrix

$$\underline{A} + \underline{A}^2 = \begin{pmatrix} 2 & 2 & 2 & 3 & 1 \\ 2 & 1 & 2 & 3 & 2 \\ 1 & 1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 & 0 \end{pmatrix}.$$

Consider as an example, the (1,4) element of this matrix. The direct connection from P_1 to P_4 gives the only 1-step connection, and the two 2-step connections are $P_1 \rightarrow P_2 \rightarrow P_4$ and $P_1 \rightarrow P_5 \rightarrow P_4$.

1.6. Exercises

1. Specify a 3×4 matrix which has only positive entries.
2. Specify 5 different matrices which have entries equal to 1.
3. Compute

$$2 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix} - 3 \begin{pmatrix} 1 & 2 & 2 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

4. Prove that, if $\underline{A}, \underline{B}, \underline{C}$ are matrices of the same size, then

$$(\underline{A} + \underline{B}) - (\underline{A} + \underline{C}) = \underline{B} - \underline{C}.$$

5. Solve equation

$$\underline{X} + \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ 4 & 4 & 4 \end{pmatrix}.$$

What is the size of \underline{X} ?

6. Solve equation

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} + \underline{X} = \begin{pmatrix} 3 & 3 & 3 \\ 4 & 4 & 4 \end{pmatrix}.$$

What is the size of \underline{X} ? Explain why the results of this and the previous problems coincide.

7. Verify relations (1.11), (1.12), and (1.13).
 8. Find values of a, b and c such that

$$\begin{pmatrix} 1+a & b \\ b & c \end{pmatrix} + \begin{pmatrix} b & a \\ c & a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

9. Let

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \underline{B} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \underline{C} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}.$$

Determine the results of the following operations:

a) $3\underline{A} + \underline{B} - 2\underline{C}$

b) $4\underline{A} - \underline{B} + \underline{C}$

c) $-\underline{A} - \underline{B} + 6\underline{C}$.

10. Solve equation

$$2(\underline{X} + \underline{A}) + \frac{1}{2}(\underline{X} + \underline{B} - \underline{C}) = \underline{X}$$

for \underline{X} , where \underline{A} , \underline{B} , and \underline{C} are the same as given in the previous problem.

11. Let

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \underline{B} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}.$$

Compute the results of the following operations

a) \underline{AB}

b) \underline{BA}

c) $(\underline{AB})\underline{A}$

d) $((\underline{AB})\underline{A})\underline{B}$

e) $((\underline{AB})\underline{B})\underline{A}$

f) $\underline{B}(\underline{AB})\underline{A}$.

12. Show that $(\underline{A} + \underline{B})^2 = \underline{A}^2 + 2\underline{AB} + \underline{B}^2$ if and only if $\underline{AB} = \underline{BA}$.

13. Show that $(\underline{A} + \underline{B})(\underline{A} - \underline{B}) = \underline{A}^2 - \underline{B}^2$ if and only if $\underline{AB} = \underline{BA}$.

14. Verify that matrix $\underline{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ satisfies equation $\underline{A}^2 = \underline{O}$.

15. Characterize all 2×2 real matrices such that $\underline{A}^2 = \underline{I}$.

16. Show that for all $k \geq 1$ and $n \times n$ matrices \underline{A} ,

$$(\underline{I}_n - \underline{A})(\underline{I}_n + \underline{A} + \underline{A}^2 + \dots + \underline{A}^k) = \underline{I}_n - \underline{A}^{k+1}.$$

17. Find a matrix satisfying equations

a) $\underline{X}^2 - 3\underline{X} + 2\underline{I} = \underline{O}$

b) $\underline{X}^3 - \underline{I} = \underline{O}$

c) $\underline{X}^3 - \underline{X}^2 - \underline{X} = \underline{O}$.

18. Assume that $ad - bc \neq 0$. Prove that

19.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

19. Find $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}^{-1}$ by using the result of the previous problem.

20. Let $\underline{A} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$.

Find matrices \underline{B} , \underline{C} such that $\underline{B} > \underline{A}$ and $\underline{C} < \underline{A}$.

21. Let $\underline{A} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$ and $\underline{B} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$. Find a matrix \underline{C} such that $\underline{C} > \underline{A} - \underline{B}$.

22. Is $(\underline{A} + \underline{B})^{-1} = \underline{A}^{-1} + \underline{B}^{-1}$?

23. Let $\underline{p}^T = (p_1, p_2, \dots, p_n)$ be a price vector, and let $\underline{1}$ be the n -element vector all elements of which are equal to 1. Show that the average price can be expressed as $\frac{1}{n} \underline{p}^T \underline{1}$.

24. Find \underline{A}' for matrix $\underline{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

25. Find $e^{\underline{A}t}$ for matrix $\underline{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Chapter 2

Vector Spaces and Inner-Product Spaces

2.1 Introduction

In this chapter some structural properties of the set of real (or complex) vectors will be introduced and analysed. Our analysis will be based on a new concept known as vector spaces which can be defined as follows:

Definition 2.1. Let V be a set and assume that the operation of addition is defined for all pairs $\underline{x}, \underline{y} \in V$, and multiplication by real (or complex) scalars is defined for all scalars a and $x \in V$ and the following properties are satisfied:

(i) $\underline{x} + \underline{y} = \underline{y} + \underline{x}$ for all $\underline{x}, \underline{y} \in V$ (commutativity);

(ii) $(\underline{x} + \underline{y}) + \underline{z} = \underline{x} + (\underline{y} + \underline{z})$ for all $\underline{x}, \underline{y}, \underline{z} \in V$ (associativity);

(iii) there exists a zero element $\underline{0} \in V$ such that for all $\underline{x} \in V$,

$$\underline{x} + \underline{0} = \underline{x};$$

(iv) for all $\underline{x} \in V$ there is an element denoted by $(-\underline{x})$ such that

$$\underline{x} + (-\underline{x}) = \underline{0};$$

(v) $a(\underline{x} + \underline{y}) = a\underline{x} + a\underline{y}$ for all scalars a , and $\underline{x}, \underline{y} \in V$ (distributivity);

(vi) $(a+b)\underline{x} = a\underline{x} + b\underline{x}$ for all scalars a and b , and $\underline{x} \in V$ (distributivity);

(vii) $(ab)\underline{x} = a(b\underline{x})$ for all scalars a and b , and $\underline{x} \in V$ (associativity);

(viii) $1\underline{x} = \underline{x}$ for all $\underline{x} \in V$.

Then V is called a *vector space*, and its elements are usually called *vectors*. If scalars are defined as real numbers, then the vector space is called *real*, and if scalars are complex numbers then the vector space is called *complex*.

Notice that in properties (iii) and (iv) we did not assume the uniqueness of the zero element $\underline{0}$ and the negative $(-\underline{x})$ of any \underline{x} . Their uniqueness follows from the other properties, which can be demonstrated in the following way. Assume that $\underline{0}$ and $\underline{0}'$ are both zero elements in a vector space. Then properties (iii) and (i) imply that

$$\underline{0} + \underline{0}' = \underline{0}$$

(when \underline{x} is selected as $\underline{0}$), and

$$\underline{0} + \underline{0}' = \underline{0}' + \underline{0} = \underline{0}'$$

(when \underline{x} is selected as $\underline{0}'$). Therefore $\underline{0} = \underline{0}'$. Assume next that both $(-\underline{x})$ and $(-\underline{x})'$ are negatives of an \underline{x} . Consider the element $(-\underline{x}) + \underline{x} + (-\underline{x})'$. Since

$$\begin{aligned} (-\underline{x}) + \underline{x} + (-\underline{x})' &= ((-\underline{x}) + \underline{x}) + (-\underline{x})' = (\underline{x} + (-\underline{x})) + (-\underline{x})' \\ &= \underline{0} + (-\underline{x})' = (-\underline{x})' + \underline{0} = (-\underline{x})', \end{aligned}$$

and on the other hand

$$(-\underline{x}) + \underline{x} + (-\underline{x})' = (-\underline{x}) + (\underline{x} + (-\underline{x})') = (-\underline{x}) + \underline{0} = (-\underline{x}),$$

the negatives are necessarily equal to each other.

We will next show that the zero element of V can be obtained by multiplying any $\underline{x} \in V$ by zero. Introduce the notation $0 \cdot \underline{x} = \underline{z}$. Then $\underline{x} = \underline{x} + \underline{z}$, since

$$\underline{x} = 1 \cdot \underline{x} = (1 + 0) \cdot \underline{x} = 1 \cdot \underline{x} + 0 \cdot \underline{x} = \underline{x} + \underline{z}.$$

Add the element $(-\underline{x})$ to both sides, to see that

$$\begin{aligned}\underline{0} &= \underline{x} + (-\underline{x}) = \underline{x} + \underline{z} + (-\underline{x}) = \underline{x} + (-\underline{x}) + \underline{z} \\ &= \underline{0} + \underline{z} = \underline{z} + \underline{0} = \underline{z}.\end{aligned}$$

This property implies that for any scalar a , $a \cdot \underline{0} = \underline{0}$, since with an $\underline{x} \in V$,

$$a \cdot \underline{0} = a \cdot (0 \cdot \underline{x}) = (a \cdot 0) \underline{x} = 0 \underline{x} = \underline{0}.$$

Notice that the selection $(-\underline{x}) = (-1) \cdot \underline{x}$ satisfies property (iv), since

$$\underline{x} + (-\underline{x}) = 1 \cdot \underline{x} + (-1) \cdot \underline{x} = (1 + (-1)) \cdot \underline{x} = 0 \cdot \underline{x} = \underline{0}.$$

The uniqueness of $(-\underline{x})$ implies that necessarily $(-\underline{x}) = (-1) \cdot \underline{x}$.

Before analysing further properties of vector spaces, some important examples are presented.

Example 2.1. Let V be the set of all n -element real column vectors, and define vector addition and multiplication by real scalars in the usual way (as given in Chapter 1). Then properties (i), (ii), (v), (vi), (vii), and (viii) are obviously satisfied. If $\underline{0}$ denotes the n -element zero column vector, and $(-\underline{x})$ is defined as $(-1) \cdot \underline{x}$, then (iii) and (iv) are also satisfied. Hence, the set of all n -element real vectors is a vector space, which is usually denoted by R^n . One may show in a similar way that the set of all complex n -element column vectors also forms a vector space with the usual addition and multiplication by real or complex scalars. The set of all n -element complex column vectors is usually denoted by C^n .

Example 2.2. Let now V denote the set of all real (or complex) $m \times n$ matrices, where m and n are fixed. Define the addition and

multiplication by real (or complex) scalars in the way as it was shown in Chapter 1. If \underline{O} is the $m \times n$ zero matrix, the element $(-\underline{A})$ is defined by multiplying \underline{A} by -1, then V is a vector space. The set of all real (or complex) $m \times n$ matrices is usually denoted by $R^{m \times n}$ (or $C^{m \times n}$), as it was introduced in Chapter 1.

◆

Example 2.3. Define next V as the set of all single variable real polynomials of degree at most n , where n is fixed. Let p and q be two polynomials:

$$p(t) = a_0 + a_1 t + \dots + a_n t^n$$

and

$$q(t) = b_0 + b_1 t + \dots + b_n t^n,$$

then the sum of p and q is defined in the following way:

$$p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \dots + (a_n + b_n)t^n.$$

If the degree of either p or q is less than n , then the "missing" coefficients up to a_n or b_n are assumed to be zeros. If a is a scalar, then multiplication by a is defined as

$$a \cdot p(t) = (a \cdot a_0) + (a \cdot a_1)t + \dots + (a \cdot a_n)t^n.$$

The zero polynomial have all zero coefficients, and $(-p)$ is given by changing the signs of all the coefficients of p :

$$(-p)(t) = (-a_0) + (-a_1)t + \dots + (-a_n)t^n.$$

Then V is a vector space.

Example 2.4. Let V be the set of all real functions defined on a closed interval $[\alpha, \beta]$. The sum $f + g$ and scalar multiple $a \cdot f$ of real functions are defined in the usual way:

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

and

$$(a \cdot f)(x) = a \cdot f(x)$$

for all $x \in [\alpha, \beta]$. It is easy to see that V is a vector space, when $(-f)(x) = -f(x)$, and the zero function has zero values for all $x \in [\alpha, \beta]$. This vector space is usually denoted by $C[\alpha, \beta]$.

Example 2.5. Let next V be the set of all real sequences (a_n) . The sum of sequences is defined as

$$(a_n) + (b_n) = (a_n + b_n),$$

that is, all corresponding terms are added. Scalar multiples of sequences are defined by multiplying each term by the scalar:

$$a \cdot (a_n) = (a \cdot a_n).$$

If one defines the sequence $-(a_n)$ as $(-a_n)$ and the zero sequence as $a_n = 0$ for all $n \geq 0$, then V is a vector space. For example,

$$\begin{aligned}(1,1,1,\dots) + (1,2,3,\dots) &= (2,3,4,\dots), \\ 2 \cdot (1,1,1,\dots) &= (2,2,2,\dots), \quad \text{and} \quad -(1,1,1,\dots) = (-1,-1,-1,\dots).\end{aligned}$$

◆

2.2 Subspaces

Before introducing the concept of subspaces in general, a simple example is presented. Consider the vector space of all n -dimensional real column vectors with the same operations as in Example 2.1. Let \underline{z} be a nonzero vector in $V = R^n$, and define V_1 as the set of all real-multiples of \underline{z} :

$$V_1 = \{c \cdot \underline{z} \mid c \in R\}.$$

Here R denotes the set of real numbers. We can easily show that V_1 satisfies all properties of a vector space. Assume that \underline{x} and $\underline{y} \in V_1$, then $\underline{x} = a_1 \cdot \underline{z}$ and $\underline{y} = a_2 \cdot \underline{z}$ with some scalars a_1 and a_2 . Then

$$\underline{x} + \underline{y} = a_1 \cdot \underline{z} + a_2 \cdot \underline{z} = (a_1 + a_2) \cdot \underline{z},$$

and for all scalars a ,

$$a \cdot \underline{x} = a \cdot (a_1 \cdot \underline{z}) = (a \cdot a_1) \cdot \underline{z},$$

that is, both $\underline{x} + \underline{y}$ and $a \cdot \underline{x}$ are in V_1 . Similarly, $\underline{0} = 0 \cdot \underline{z} \in V_1$, and for all $\underline{x} \in V_1$, $\underline{x} = a \cdot \underline{z}$ with some scalar a , therefore

$$(-\underline{x}) = (-1) \cdot \underline{x} = (-1) \cdot (a \cdot \underline{z}) = (-a) \cdot \underline{z} \in V_1.$$

Hence the zero element and $(-\underline{x})$ for all $\underline{x} \in V_1$ are in V_1 . All properties (commutativity, associativity, and distributivity) of the operations also hold in V_1 , since they are true in the larger set V .

◆

Definition 2.2. A subset V_1 of a vector space V is called a *subspace*, if V_1 itself is a vector space.

Notice that the entire vector space V can also be considered as a subspace of itself. Similarly, the set containing only the zero element of V also satisfies the defining properties of a vector space, therefore it is also a subspace in V . These two special subspaces are called the *trivial subspaces* of V .

In verifying the conditions of vector spaces for any subset V_1 we do not need to check the operation properties (i), (ii), (v), (vi), (vii), and (viii), since they are true for all elements of V , and since $V_1 \subseteq V$, they remain true for all elements of V_1 as well. Therefore we only have to check that for all \underline{x} and $\underline{y} \in V_1$, the sums $\underline{x} + \underline{y}$ are in V_1 , and for all scalars a and $\underline{x} \in V_1$, the products $a \cdot \underline{x}$ are in V_1 . Notice that for all \underline{x} , $0 \cdot \underline{x} = \underline{0}$ and $(-1) \cdot \underline{x} = (-\underline{x})$, therefore the zero element and the negative of all $\underline{x} \in V_1$ are necessarily in V_1 . Since properties (iii) and (iv) hold in the entire V , therefore they are valid in the subset V_1 . This observation can be summarized as follows.

Theorem 2.1. If V is a vector space and $V_1 \subseteq V$, then V_1 is a subspace of V if and only if

- (a) $\underline{x} + \underline{y} \in V_1$ for all $\underline{x}, \underline{y} \in V_1$;
 (b) $a \cdot \underline{x} \in V_1$ for all scalars a and $\underline{x} \in V_1$.

The application of the theorem can be illustrated by the following important example.

Example 2.6. Let V be a vector space, and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ be given elements from V . Define

$$V_1 = \{a_1 \cdot \underline{x}_1 + a_2 \cdot \underline{x}_2 + \dots + a_k \cdot \underline{x}_k \mid a_1, a_2, \dots, a_k \in R\}.$$

Notice that the introductory example of this section is the special case of this example, when one selects $k = 1$. We will now verify that V_1 is a subspace.

If $\underline{x}, \underline{y} \in V_1$, then with some scalars a_1, \dots, a_k and b_1, \dots, b_k ,

$$\underline{x} = a_1 \cdot \underline{x}_1 + a_2 \cdot \underline{x}_2 + \dots + a_k \cdot \underline{x}_k$$

and

$$\underline{y} = b_1 \cdot \underline{x}_1 + b_2 \cdot \underline{x}_2 + \dots + b_k \cdot \underline{x}_k,$$

and therefore

$$\underline{x} + \underline{y} = (a_1 + b_1) \cdot \underline{x}_1 + (a_2 + b_2) \cdot \underline{x}_2 + \dots + (a_k + b_k) \cdot \underline{x}_k.$$

Hence, $\underline{x} + \underline{y} \in V_1$. Similarly, for all scalars a ,

$$\begin{aligned} a \cdot \underline{x} &= a \cdot (a_1 \cdot \underline{x}_1 + a_2 \cdot \underline{x}_2 + \dots + a_k \cdot \underline{x}_k) \\ &= (a \cdot a_1) \cdot \underline{x}_1 + (a \cdot a_2) \cdot \underline{x}_2 + \dots + (a \cdot a_k) \cdot \underline{x}_k \in V_1. \end{aligned}$$

Therefore both conditions of the theorem are satisfied showing that V_1 is a vector space. ◆

For any $\underline{x}_1, \dots, \underline{x}_k \in V$ and scalars a_1, \dots, a_k , the element $a_1 \cdot \underline{x}_1 + \dots + a_k \cdot \underline{x}_k \in V$ is called a *linear combination* of the given vectors. Example 2.6 can be reformulated by saying that for any given finite set of the elements of a vector space, the set of all linear combinations of the given elements always form a subspace.

In verifying that a subset V_1 of a given vector space is a subspace, the following theorem is applied most frequently.

Theorem 2.2. If V is a vector space and $V_1 \subseteq V$, then V_1 is a subspace if and only if for all $\underline{x}, \underline{y} \in V_1$ and scalars a and b ,

$$a \cdot \underline{x} + b \cdot \underline{y} \in V_1.$$

Proof. Assume first, that V_1 is a subspace. Then for all a and \underline{x} , $a \cdot \underline{x} \in V_1$ and for all b and \underline{y} , $b \cdot \underline{y} \in V_1$. Since the sum of the elements of V_1 must be in V_1 ,

$$a \cdot \underline{x} + b \cdot \underline{y} \in V_1.$$

Assume next that V_1 satisfies the condition of the theorem. Select an arbitrary a and \underline{x} , and define $b = 0$. Then

$$a \cdot \underline{x} + b \cdot \underline{y} = a \cdot \underline{x} \in V_1.$$

Choose next $a = b = 1$, then

$$a \cdot \underline{x} + b \cdot \underline{y} = \underline{x} + \underline{y} \in V_1,$$

therefore both conditions of Theorem 2.1 are satisfied, and hence, V_1 is a subspace.

♣

The advantage of this result compared to Theorem 2.1 is the fact that only one condition has to be verified to show that a subset of a vector space is a subspace.

Some important properties of the subspaces of a given vector space are presented next.

Theorem 2.3. Let V be a vector space, and $\{V_\alpha\}$ ($\alpha \in A$ with some index set A) be a finite or infinite set of (not necessarily all) subspaces of V . Then the intersection of the given subspaces is also a subspace.

Proof. Let V_1 denote the intersection of the subspaces V_α . We will show that V_1 satisfies the condition of Theorem 2.2. Let \underline{x} and $\underline{y} \in V_1$ and let a and b be two scalars. The definition of V_1 implies that for all α , \underline{x} and \underline{y} are in V_α and since V_α is a subspace, $a \cdot \underline{x} + b \cdot \underline{y} \in V_\alpha$ for all α . Therefore $a \cdot \underline{x} + b \cdot \underline{y}$ has to belong to the intersection of sets V_α . That is,

$$a \cdot \underline{x} + b \cdot \underline{y} \in V_1,$$

which completes the proof.

♣

Definition 2.3. Let S be an arbitrary (finite or infinite) subset of a vector space V . The subspace *generated* by the set S is the intersection of all subspaces of V containing S .

Since V itself contains S and it is a trivial subspace in itself, there is at least one subspace containing S . From Theorem 2.3 we know that the intersection of all such subspaces is also a subspace. We will use the

notation $V(S)$ for the subspace generated by the set S . From the definition we may conclude that if V_1 is any subspace of V containing S , then $V(S)$ is also contained in V_1 , therefore $V(S)$ is the *smallest* subspace containing S . The construction of $V(S)$ is given in the following result.

Theorem 2.4. Let V be a vector space and let S be an arbitrary non-empty subset of V . Then $V(S)$ is the set of all linear combinations of finitely many elements of S :

$$V(S) = \left\{ \sum_{i=1}^N a_i \underline{u}_i \mid N \in \mathbb{N}^+, \underline{u}_i \in S, i = 1, 2, \dots, N \text{ and } a_i \text{ is a scalar for all } i \right\} \quad (2.1)$$

Proof. First we will show that the set of all linear combinations of finitely many elements of S is a subspace in V . For simple reference, let $L(S)$ denote this set. It is easy to verify that $L(S)$ satisfies the condition of Theorem 2.2. Let \underline{x} and \underline{y} be two elements of $L(S)$, and let a and b be two scalars. Then we may assume that

$$\underline{x} = a_1 \underline{x}_1 + \dots + a_k \underline{x}_k + b_1 \underline{y}_1 + \dots + b_l \underline{y}_l$$

and

$$\underline{y} = c_1 \underline{x}_1 + \dots + c_k \underline{x}_k + d_1 \underline{z}_1 + \dots + d_m \underline{z}_m,$$

where $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l, \underline{z}_1, \dots, \underline{z}_m \in S$ and the common elements in the linear combinations are denoted by $\underline{x}_1, \dots, \underline{x}_k$. Simple calculation shows that

$$\begin{aligned}
 a \cdot \underline{x} + b \cdot \underline{y} &= (a \cdot a_1) \cdot \underline{x}_1 + \dots + (a \cdot a_k) \cdot \underline{x}_k + (a \cdot b_1) \cdot \underline{y}_1 + \dots + (a \cdot b_l) \cdot \underline{y}_l \\
 &\quad + (b \cdot c_1) \cdot \underline{x}_1 + \dots + (b \cdot c_k) \cdot \underline{x}_k + (b \cdot d_1) \cdot \underline{z}_1 + \dots + (b \cdot d_m) \cdot \underline{z}_m \\
 &= (a \cdot a_1 + b \cdot c_1) \cdot \underline{x}_1 + \dots + (a \cdot a_k + b \cdot c_k) \underline{x}_k + (a \cdot b_1) \cdot \underline{y}_1 + \dots \\
 &\quad + (a \cdot b_l) \cdot \underline{y}_l + (b \cdot d_1) \cdot \underline{z}_1 + \dots + (b \cdot d_m) \cdot \underline{z}_m,
 \end{aligned}$$

which is the linear combination of the elements $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l, \underline{z}_1, \dots, \underline{z}_m$, therefore it is in $L(S)$. Hence $L(S)$ is a vector space. Notice that for all $\underline{x} \in S$, $\underline{x} = 1 \cdot \underline{x} \in L(S)$, consequently $S \subseteq L(S)$.

Assume next that V_1 is a subspace of V such that $S \subseteq V_1$. Let $\underline{x}_1, \dots, \underline{x}_k \in S$ be arbitrary elements and let a_1, \dots, a_k be arbitrary scalars. Since V_1 is a subspace containing $\underline{x}_1, \dots, \underline{x}_k$, the element $\underline{y}_1 = a_1 \underline{x}_1 + a_2 \underline{x}_2$ must belong to V_1 . Then $\underline{y}_2 = 1 \underline{y}_1 + a_3 \underline{x}_3 = a_1 \underline{x}_1 + a_2 \underline{x}_2 + a_3 \underline{x}_3$ also belongs to V_1 . Similarly, $\underline{y}_3 = 1 \underline{y}_2 + a_4 \underline{x}_4 = a_1 \underline{x}_1 + \dots + a_4 \underline{x}_4 \in V_1$. Repeating the same reasoning $k-1$ times we conclude that $\underline{y}_{k-1} = 1 \underline{y}_{k-2} + a_k \underline{x}_k = a_1 \underline{x}_1 + \dots + a_{k-1} \underline{x}_{k-1} + a_k \underline{x}_k \in V_1$. In other words, all linear combinations of finitely many elements of S necessarily belong to V_1 . That is, $L(S) \subseteq V_1$. Hence $L(S)$ is the smallest subspace containing S , therefore it is the subspace generated by S .

✦

The assertion of the theorem is illustrated by the following examples.

Example 2.7. Consider the set R^3 of real 3-element column vectors, and let $\underline{x} \neq \underline{0}$ be an element of R^3 . Then the subspace generated by vector \underline{x} is the set of all real multiples of \underline{x} :

$$V(\{\underline{x}\}) = \{a\underline{x} \mid a \text{ is a real scalar}\}$$

forming a straight line which passes through the origin, since $0\underline{x} = \underline{0}$.

Example 2.8. Consider again the vector space R^3 , and let \underline{x} and \underline{y} be two nonzero vectors such that \underline{y} is not a constant multiple of \underline{x} . Then the set

$$V(\{\underline{x}, \underline{y}\}) = \{a\underline{x} + b\underline{y} \mid a, b \text{ are real scalars}\}$$

is a plane containing the origin. Figure 2.1 illustrates this subspace, which is the plane containing the origin and the endpoints of vectors \underline{x} and \underline{y} .

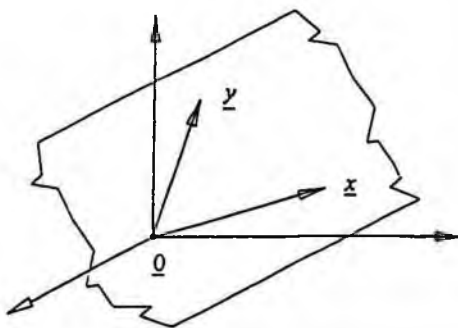


Figure 2.1. Subspace generated by two vectors.

Example 2.9. Let V be the set of all real variable, real valued functions defined on the interval $(-\infty, \infty)$, and let real numbers be considered as scalars. For a given integer $n \geq 1$, select the functions

$$p_0(x) = 1, p_1(x) = x, p_2(x) = x^2, \dots, p_n(x) = x^n.$$

Then the subspace generated by these functions is the set of all linear combinations of these functions, which is the set of all real polynomials of degree of most n :

$$V(\{p_0, p_1, \dots, p_n\}) = \{a_0 + a_1x + \dots + a_nx^n \mid a_0, a_1, \dots, a_n \text{ are real scalars}\}$$

♦

2.3 Linear Independence, Basis

In the previous section we have seen that any element of a subspace generated by a subset of a vector space can be obtained as a linear combination of finitely many elements from that subset. The description of the generated subspace becomes more simple if the necessary number of the elements in the linear combinations is smaller. In this section we will find the minimum number of such elements and the corresponding minimal realization of subspaces. Our analysis will be based on the concept of linear independence, which will be defined next.

Definition 2.4. Let V be a vector space. The elements $\underline{x}_1, \dots, \underline{x}_n \in V$ are called *linearly independent*, if

$$a_1\underline{x}_1 + a_2\underline{x}_2 + \dots + a_n\underline{x}_n = \underline{0}$$

implies that $a_1 = a_2 = \dots = a_n = 0$. If the elements are not linearly independent, then we say that they are *linearly dependent*.

That is, a linear combination of linearly independent vectors is zero only if all coefficients are equal to zero.

Example 2.10. Consider the vector space R^n , and select the vectors:

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Notice that for $k = 1, 2, \dots, n$, the k^{th} entry of \underline{e}_k equals 1, and all other elements are equal to zero. We will next prove that these vectors are linearly independent. A linear combination of these vectors is zero, when with some constants a_1, \dots, a_n ,

$$a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \dots + a_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

The first element of the left-hand side vector is

$$a_1 \cdot 1 + a_2 \cdot 0 + \dots + a_n \cdot 0,$$

which is zero only if $a_1 = 0$. The second element of the left-hand side vector equals

$$a_1 \cdot 0 + a_2 \cdot 1 + \dots + a_n \cdot 0,$$

which is zero only if $a_2 = 0$. In general, the k^{th} ($k = 1, 2, \dots, n$) element of the left-hand side vector is

$$a_1 \cdot 0 + \dots + a_{k-1} \cdot 0 + a_k \cdot 1 + a_{k+1} \cdot 0 + \dots + a_n \cdot 0,$$

which is zero only if $a_k = 0$. Thus all coefficients are necessarily zero. ♦

Example 2.11. Next let V be the set of all real functions defined for all $x \in (-\infty, \infty)$. Select the functions

$$p_0(x) = 1, p_1(x) = x, p_2(x) = x^2, \dots, p_n(x) = x^n$$

from V . We will now prove that these functions are linearly independent. If their linear combination is the zero element in V , then for all real x ,

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0,$$

which is possible only if $a_0 = a_1 = \dots = a_n = 0$.

Contrary to this, assume that at least one coefficient is nonzero. Assume that $a_k \neq 0$ and $a_{k+1} = \dots = a_n = 0$. If $k = 0$, then the polynomial is a nonzero constant. Otherwise, for all real $x \neq 0$,

$$a_0 + a_1x + a_2x^2 + \dots + a_kx^k = x^k \left(a_k + \frac{a_{k-1}}{x} + \dots + \frac{a_1}{x^{k-1}} + \frac{a_0}{x^k} \right).$$

If $x \rightarrow \infty$, then $x^k \rightarrow \infty$, and all terms $\frac{a_{k-1}}{x}, \dots, \frac{a_1}{x^{k-1}}, \frac{a_0}{x^k}$ converge to zero. Therefore this polynomial converges to ∞ if $a_k > 0$ and it converges to $-\infty$ if $a_k < 0$. Hence the polynomial cannot have zero value for all real x .

◆

Some elementary properties of linearly independent elements are listed below:

1. If $\underline{0}$ is among the selected elements, then they are linearly dependent. Assume that $\underline{0}, \underline{x}_2, \dots, \underline{x}_k$ are selected. Then

$$1 \cdot \underline{0} + 0 \cdot \underline{x}_2 + \dots + 0 \cdot \underline{x}_k = \underline{0}$$

showing that a linear combination equals zero with a nonzero coefficient. Hence these elements are linearly dependent.

2. If among $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$, at least one element is repeated, then they are linearly dependent.

Assume now that $\underline{x}_2 = \underline{x}_1$, otherwise we have to renumber the elements. Then

$$1 \cdot \underline{x}_1 + (-1) \cdot \underline{x}_2 + 0 \cdot \underline{x}_3 + \dots + 0 \cdot \underline{x}_k = \underline{0},$$

and there are two nonzero coefficients.

3. Any subset of the set of linearly independent elements contains only linearly independent vectors.

Assume that $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$, are linearly independent, and let $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_l$ ($l < k$) be a subset of the original set of elements. In contrary to the assertion assume that the subset has linearly dependent vectors. Then with some scalars a_1, \dots, a_l ,

$$a_1 \cdot \underline{x}_1 + \dots + a_l \cdot \underline{x}_l = \underline{0}$$

and there is at least one nonzero coefficient a_i ($1 \leq i \leq l$). Consider now the linear combination

$$a_1 \cdot \underline{x}_1 + \dots + a_l \cdot \underline{x}_l + 0 \cdot \underline{x}_{l+1} + \dots + 0 \cdot \underline{x}_k$$

of the original elements, which is the zero vector, but the coefficient a_i is nonzero. Therefore the original elements are also linearly dependent, which contradicts the assumption. Hence, the subset contains only linearly independent elements.

4. Let $\underline{x} \neq \underline{0}$ be an element of a vector space. Then it forms a linearly independent (one element) set.

In contrary to the assertion assume that $a \cdot \underline{x} = \underline{0}$ with some $a \neq 0$. Then $\frac{1}{a}$ exists and

$$\underline{x} = 1 \cdot \underline{x} = \left(\frac{1}{a} \cdot a \right) \cdot \underline{x} = \frac{1}{a} \cdot (a \cdot \underline{x}) = \frac{1}{a} \cdot \underline{0} = \underline{0},$$

which is an obvious contradiction to the assumption that \underline{x} is nonzero.

Theorem 2.5. The elements $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ of a vector space V are linearly dependent if and only if at least one of them can be expressed as the linear combination of the others.

Proof. Assume first that $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ are linearly dependent. Then with some scalars a_1, a_2, \dots, a_n ,

$$a_1 \underline{x}_1 + a_2 \underline{x}_2 + \dots + a_n \underline{x}_n = \underline{0},$$

where at least one coefficient is nonzero, say $a_i \neq 0$. Then

$$a_i \underline{x}_i = -a_1 \underline{x}_1 - \dots - a_{i-1} \underline{x}_{i-1} - a_{i+1} \underline{x}_{i+1} - \dots - a_n \underline{x}_n,$$

therefore

$$\underline{x}_i = \left(-\frac{a_1}{a_i} \right) \underline{x}_1 + \dots + \left(-\frac{a_{i-1}}{a_i} \right) \underline{x}_{i-1} + \left(-\frac{a_{i+1}}{a_i} \right) \underline{x}_{i+1} + \dots + \left(-\frac{a_n}{a_i} \right) \underline{x}_n$$

where the right-hand side is a linear combination of the elements $\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_n$.

Assume next that at least one element, say \underline{x}_i , equals the linear combination of the others. Then

$$\underline{x}_i = a_1 \cdot \underline{x}_1 + \dots + a_{i-1} \cdot \underline{x}_{i-1} + a_{i+1} \cdot \underline{x}_{i+1} + \dots + a_n \cdot \underline{x}_n,$$

which can be rewritten as

$$a_1 \underline{x}_1 + \dots + a_{i-1} \underline{x}_{i-1} + (-1) \cdot \underline{x}_i + a_{i+1} \cdot \underline{x}_{i+1} + \dots + a_n \cdot \underline{x}_n = \underline{0}.$$

The left-hand side is a linear combination of all the elements, and the coefficient of \underline{x}_i is nonzero, therefore the elements are linearly dependent.

♣

In the case when all selected elements are nonzero, this theorem can be reformulated as follows.

Theorem 2.6. Assume that $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ are nonzero elements of a vector space V . Then they are linearly dependent if and only if there exists an i ($i = 2, 3, \dots, n$) such that \underline{x}_i is the linear combination of $\underline{x}_1, \dots, \underline{x}_{i-1}$.

Proof. Assume first that the vectors are linearly dependent. This assumption implies that

$$a_1 \cdot \underline{x}_1 + a_2 \cdot \underline{x}_2 + \dots + a_n \cdot \underline{x}_n = \underline{0}$$

where at least one of the coefficients is nonzero. Let i be the largest integer ($1 \leq i \leq n$) such that $a_i \neq 0$. Then

$$a_1 \cdot \underline{x}_1 + \dots + a_{i-1} \underline{x}_{i-1} + a_i \cdot \underline{x}_i = \underline{0},$$

since $a_{i+1} = \dots = a_n = 0$. This equation implies that

$$a_i \cdot \underline{x}_i = -a_1 \underline{x}_1 - \dots - a_{i-1} \underline{x}_{i-1},$$

that is,

$$\underline{x}_i = \left(-\frac{a_1}{a_i} \right) \cdot \underline{x}_1 + \dots + \left(-\frac{a_{i-1}}{a_i} \right) \underline{x}_{i-1}.$$

Assume next that \underline{x}_i can be obtained as a linear combination of $\underline{x}_1, \dots, \underline{x}_{i-1}$. Then it is the linear combination of $\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_n$, where $\underline{x}_{i+1}, \dots, \underline{x}_n$ are multiplied by zero. Therefore Theorem 2.5 implies that $\underline{x}_1, \dots, \underline{x}_n$ are linearly dependent.

♣

Example 2.12. The vectors

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \text{ and } \underline{x}_3 = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}$$

are linearly dependent, since $\underline{x}_3 = 2 \cdot \underline{x}_1 + 1 \cdot \underline{x}_2$, which can be shown by simple calculation:

$$2 \cdot \underline{x}_1 + 1 \cdot \underline{x}_2 = 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} = \underline{x}_3.$$

♦

Theorem 2.7. The elements $\underline{x}_1, \dots, \underline{x}_n$ of a vector space V are linearly independent if and only if every element of the subspace generated by these vectors can be uniquely expressed as the linear combination of these vectors.

Proof. Assume first that an element of the generated subspace can be expressed as two different linear combinations:

$$\underline{x} = a_1 \underline{x}_1 + \dots + a_n \underline{x}_n = b_1 \underline{x}_1 + \dots + b_n \underline{x}_n.$$

This equation implies that

$$(a_1 - b_1) \cdot \underline{x}_1 + \dots + (a_n - b_n) \cdot \underline{x}_n = \underline{0},$$

where for at least one i , $a_i - b_i \neq 0$. This implies that $\underline{x}_1, \dots, \underline{x}_n$ are linearly dependent.

Assume next that each element of $V(\{\underline{x}_1, \dots, \underline{x}_n\})$ can be uniquely expressed as the linear combination of $\underline{x}_1, \dots, \underline{x}_n$. Since

$$0 \cdot \underline{x}_1 + 0 \cdot \underline{x}_2 + \dots + 0 \cdot \underline{x}_n = \underline{0},$$

there is no other linear combination of these elements that equals $\underline{0}$. Therefore $\underline{x}_1, \dots, \underline{x}_n$ are linearly independent.

♣

Before defining the basis and dimension of vector spaces and subspaces, one more concept should be introduced.

Definition 2.5. Let V be a vector space and let S be a subset of V . We say that S is a *generating system* of V , if $L(S) = V$, that is, if V coincides with the set of all linear combinations of finitely many elements of S .

Every vector space V has generating systems, since V itself is a generating system. (Notice that each element $\underline{x} \in V$ is a linear combination of itself: $1 \cdot \underline{x}$.) Vector spaces (or subspaces) with finite generating systems are called *finitely generated*.

Definition 2.6. A vector space $V \neq \{0\}$ is said to be *n-dimensional*, if it has at least one *n*-element generating system and any set with less than *n* elements is not a generating system of V . The dimension of V is denoted by $\dim(V)$. If a vector space does not have finite generating system, then it is said to be *infinite dimensional*.

This concept is illustrated next.

Example 2.13. Vector space $V = R^n$ is finite dimensional, since the vectors

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \dots, \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

form a generating system. Let

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} \in V$$

be an arbitrary vector, then it is easy to see by comparing the entries of the vectors on both sides of the equality, that

$$\underline{a} = a_1 \cdot \underline{e}_1 + a_2 \cdot \underline{e}_2 + \dots + a_n \cdot \underline{e}_n.$$

Hence, arbitrary $\underline{a} \in V$ can be expressed as the linear combination of the vectors $\underline{e}_1, \dots, \underline{e}_n$.

◆

Example 2.14. Let V be the set of all single-variable real polynomials. Then V is infinite-dimensional, since there is no finite generating system. Assume to the contrary that a finite set of polynomials p_1, p_2, \dots, p_N is a generating system. Let n_1, n_2, \dots, n_N denote the degrees of these polynomials, respectively, and define $n = \max\{n_1, n_2, \dots, n_N\}$, which is the largest among the degrees. Since every linear combination of these polynomials is a polynomial of degree at most n , no polynomial of higher degree can be obtained as a linear combination. Therefore polynomials p_1, p_2, \dots, p_N do not form a generating system.

◆

Definition 2.7. Let V be a finite dimensional vector space. A finite set of elements $\underline{x}_1, \dots, \underline{x}_n$ is called a *basis* of V , if

- (i) they are linearly independent;
- (ii) they form a generating system of V .

Example 2.15. Combining the results of Examples 2.10 and 2.13 we conclude that vectors $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$ form a basis in R^n .

◆

Example 2.16. Let V be the set of single variable, real polynomials of degree at most n . Then

$$p_0(t) = 1, p_1(t) = t, p_2(t) = t^2, \dots, p_n(t) = t^n$$

is a basis of V .

◆

Example 2.17. Let V be an arbitrary vector space, and let $\underline{0}$ denote the zero element of V . Assume that S consists of only the zero element: $S = \{\underline{0}\}$. As we have seen earlier, this is a trivial subspace of V . Notice that S has no basis, since equality $1 \cdot \underline{0} = \underline{0}$ shows that the zero vector itself forms a linearly dependent (one-element) set. That is, no independent generating system exists. We will say that $\dim(S) = 0$. The dimension of any other subspace of V is given in Definition 2.6.

◆

Some important properties of generating system are given in the following theorems.

Theorem 2.8. Assume that $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_N\}$ is a generating system of a vector space V , and for some i , \underline{x}_i can be expressed as the linear combination of elements $\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_N$. Then $\{\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_N\}$ is also a generating system of V .

Proof. We will show that any arbitrary element $\underline{x} \in V$ can be expressed as a linear combination of $\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_N$, that is, \underline{x}_i can be deleted from the generating system. For any \underline{x} ,

$$\underline{x} = a_1 \underline{x}_1 + a_2 \underline{x}_2 + \dots + a_{i-1} \underline{x}_{i-1} + a_i \underline{x}_i + a_{i+1} \underline{x}_{i+1} + \dots + a_N \underline{x}_N. \quad (2.2)$$

Since \underline{x}_i is a linear combination of the other elements, with some scalars $c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_N$,

$$\underline{x}_i = c_1 \underline{x}_1 + \dots + c_{i-1} \underline{x}_{i-1} + c_{i+1} \underline{x}_{i+1} + \dots + c_N \underline{x}_N.$$

Substituting this equality into (2.2) gives the relation

$$\begin{aligned} \underline{x} &= a_1 \underline{x}_1 + \dots + a_{i-1} \underline{x}_{i-1} + a_i (c_1 \underline{x}_1 + \dots + c_{i-1} \underline{x}_{i-1} + c_{i+1} \underline{x}_{i+1} + \dots + c_N \underline{x}_N) \\ &\quad + \underline{a}_{i+1} \underline{x}_{i+1} + \dots + \underline{a}_N \underline{x}_N \\ &= (a_1 + a_i c_1) \underline{x}_1 + \dots + (a_{i-1} + a_i c_{i-1}) \underline{x}_{i-1} + (a_{i+1} + a_i c_{i+1}) \underline{x}_{i+1} + \dots \\ &\quad + (a_N + a_i c_N) \underline{x}_N, \end{aligned}$$

which is a linear combination of only $\underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_N$. Hence, the proof is completed. ♣

Corollary. Let S be a finite generating system of a vector space $V \neq \{0\}$. Then it contains a basis of V .

Proof. If the elements of S are linearly independent, then they form a basis. Otherwise, at least one element of S can be deleted such that the remaining set is still a generating system. If it has only linearly independent vectors, then a basis is obtained. Otherwise another element can be deleted, and so on. The process has to continue until only linearly independent vectors are obtained. ♣

Theorem 2.9. Let V be a finite dimensional vector space, and assume that $\{\underline{x}_1, \dots, \underline{x}_n\}$ is a generating system of V , and let $\{\underline{y}_1, \dots, \underline{y}_m\}$ be a linearly independent set in V . Then $m \leq n$.

Proof. Assumption that $\{\underline{x}_1, \dots, \underline{x}_n\}$ is a basis implies that \underline{y}_1 is a linear combination of these elements. Therefore $\{\underline{y}_1, \underline{x}_1, \dots, \underline{x}_n\}$ is a linearly

dependent set. From Theorem 2.6 we know that there exists an \underline{x}_i ($1 \leq i \leq n$) which is a linear combination of $\underline{y}_1, \underline{x}_1, \dots, \underline{x}_{i-1}$, and so $\{\underline{y}_1, \underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_n\}$ is also a generating system of V , and \underline{y}_2 is a linear combination of these elements. Therefore set $\{\underline{y}_1, \underline{y}_2, \underline{x}_1, \dots, \underline{x}_{i-1}, \underline{x}_{i+1}, \dots, \underline{x}_n\}$ consists of linearly dependent elements, therefore there is a j ($j = 1, \dots, i-1, i+1, \dots, n$) such that \underline{x}_j is a linear combination of the preceding elements. Notice that the linear independence of $\{\underline{y}_1, \dots, \underline{y}_m\}$ implies that this element cannot be \underline{y}_2 . Therefore \underline{x}_j can be deleted from the generating system. Continue this process until all elements $\underline{y}_1, \dots, \underline{y}_m$ are entered into the generating system by deleting one \underline{x}_k at each step. Hence the resulting generating system will consist of n elements and will contain all of $\underline{y}_1, \dots, \underline{y}_m$ implying that $m \leq n$.

♣

Corollary 1. In a finite dimensional vector space any two bases consist of the same number of elements, which is the dimension of the vector space. This coincides with our earlier discussion when we defined the dimension of the subspace $\{0\}$ as zero, since this subspace has no basis.

Corollary 2. Let $\underline{x}_1, \dots, \underline{x}_k$ be linearly independent elements in an n -dimensional vector space V , where $k < n$. Then there are certain elements $\underline{y}_{k+1}, \dots, \underline{y}_n \in V$ such that $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}, \dots, \underline{y}_n\}$ is a basis of V . That is, in finite dimensional vector spaces each set of linearly independent elements can be completed to form a basis.

Proof. If $\{\underline{x}_1, \dots, \underline{x}_k\}$ is a generating system, then it is a basis. Otherwise there is an element \underline{y}_{k+1} , which cannot be expressed as the linear combination of $\underline{x}_1, \dots, \underline{x}_k$. Therefore $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}$ are linearly independent. If these vectors generate the entire vector space, then they form

a basis. Otherwise, there is a \underline{y}_{k+2} , which cannot be expressed as a linear combination of these vectors, so $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}, \underline{y}_{k+2}$ are linearly independent. Since the vector space is n -dimensional, after $n-k$ steps we obtain a set $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}, \dots, \underline{y}_n\}$ of linearly independent vectors that generates V , which is therefore a basis.

+

Corollary 3. Let V be an n -dimensional vector space and assume that $\{\underline{x}_1, \dots, \underline{x}_n\}$ is a linearly independent set. Then it is a basis of V .

Example 2.18. The vector space R^n is n -dimensional, since from Examples 2.10 and 2.13 we know that vectors $\underline{e}_1, \dots, \underline{e}_n$ form a basis. This basis is usually called the *natural basis* of R^n .

◆

Example 2.19. For given positive integers m and n , let V be the set of all real or complex $m \times n$ matrices. For $i = 1, \dots, m$ and $j = 1, \dots, n$, introduce matrices \underline{E}_{ij} , the (i, j) -entry of which equals 1 and all other elements equal 0. Similar to the case of R^n one may easily show that these matrices form a basis of V , therefore this vector space is $(m \cdot n)$ -dimensional.

◆

Example 2.20. Assume next that V is the set of all real polynomials of degree at most n . Then this vector space is $(n+1)$ -dimensional, since from Examples 2.9 and 2.11 we conclude that $p_0(x) = 1, p_1(x) = x, \dots, p_n(x) = x^n$ form a basis in V .

◆

2.4 Inner-Product Spaces

Let V be a real vector space.

Definition 2.8. An *inner product* on V is a bivariable function, which assigns to each ordered pair $\underline{x}, \underline{y}$ of elements of V a real number, denoted by $(\underline{x}, \underline{y})$, such that the following conditions are satisfied:

- (i) $(\underline{x}, \underline{x}) \geq 0$ for all $\underline{x} \in V$, and $(\underline{x}, \underline{x}) = 0$ if and only if $\underline{x} = \underline{0}$;
- (ii) $(\underline{x}, \underline{y}) = (\underline{y}, \underline{x})$ for all $\underline{x}, \underline{y} \in V$;
- (iii) $(\underline{x}_1 + \underline{x}_2, \underline{y}) = (\underline{x}_1, \underline{y}) + (\underline{x}_2, \underline{y})$ for all $\underline{x}_1, \underline{x}_2, \underline{y} \in V$;
- (iv) $(c\underline{x}, \underline{y}) = c(\underline{x}, \underline{y})$ for all $\underline{x}, \underline{y} \in V$ and $c \in R$.

Example 2.21. Let $V = R^n$ and for all $\underline{x}, \underline{y} \in V$, define

$$(\underline{x}, \underline{y}) = \underline{x}^T \underline{y}$$

(which is the product of \underline{x} as a row vector by the column vector \underline{y}). If x_i and y_i denote the elements of \underline{x} and \underline{y} , respectively, then

$$(\underline{x}, \underline{y}) = x_1 y_1 + \dots + x_n y_n.$$

We will now show that this inner product satisfies all the above properties:

(i) $(\underline{x}, \underline{x}) = x_1^2 + x_2^2 + \dots + x_n^2$ is always nonnegative, and since all terms are nonnegative, it equals zero if and only if all terms equal zero, which occurs if and only if $x_i = 0$ for all i .

$$(ii) \quad (\underline{x}, \underline{y}) = x_1 y_1 + \dots + x_n y_n = y_1 x_1 + \dots + y_n x_n = (\underline{y}, \underline{x});$$

$$(iii) \quad (\underline{x}_1 + \underline{x}_2, \underline{y}) = (x_{11} + x_{21})y_1 + \dots + (x_{1n} + x_{2n})y_n$$

$$= (x_{11}y_1 + \dots + x_{1n}y_n) + (x_{21}y_1 + \dots + x_{2n}y_n) = (\underline{x}_1, \underline{y}) + (\underline{x}_2, \underline{y}),$$

where x_{1i} and x_{2i} denote the entries of \underline{x}_1 and \underline{x}_2 , respectively;

$$(iv) \quad (c\underline{x}, \underline{y}) = (cx_1)y_1 + \dots + (cx_n)y_n$$

$$= c(x_1 y_1) + \dots + c(x_n y_n) = c(x_1 y_1 + \dots + x_n y_n) = c(\underline{x}, \underline{y}).$$

◆

The inner product of two vectors introduced in this example plays an important role in economics. One important application is the following.

Let \underline{x} and \underline{y} be two vectors with nonnegative entries. Then obviously

$$(\underline{x}, \underline{y}) = x_1 y_1 + \dots + x_n y_n \geq 0,$$

and

$$(\underline{x}, \underline{y}) = 0$$

if and only if for all $k = 1, 2, \dots, n$, either $x_k = 0$ or $y_k = 0$. This property is known as the complementarity condition.

Example 2.22. Let now V be the set of all continuous real functions on an interval $[a, b]$. As in Example 2.4 this vector space is usually denoted by $C[a, b]$. Let f and g be two functions from this set. Define the inner product as

$$(f, g) = \int_a^b f(x)g(x)dx.$$

Notice that the integral exists, since both functions are continuous. We will now prove that this inner product satisfies the conditions of Definition 2.8.

$$(i) \quad (f, f) = \int_a^b f^2(x)dx \geq 0, \text{ since } f^2(x) \geq 0 \text{ for all } x \in [a, b].$$

Assume next that $(f, f) = 0$, that is,

$$\int_a^b f^2(x)dx = 0.$$

Assume that for some $x_0 \in [a, b]$, $f(x_0) \neq 0$. Then $f^2(x_0) > 0$, and the continuity of function f implies that f^2 is also continuous, and therefore there is a neighborhood of x_0 such that $f^2(x) > 0$ for every x in this neighborhood. Since for all $x \in [a, b]$, $f^2(x) \geq 0$, the integral of f^2 must be positive contradicting the assumption. Therefore $f(x) = 0$ for all $x \in [a, b]$.

$$(ii) \quad (f, g) = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = (g, f);$$

$$\begin{aligned}
 \text{(iii)} \quad (f_1 + f_2, g) &= \int_a^b (f_1(x) + f_2(x))g(x)dx \\
 &= \int_a^b (f_1(x)g(x) + f_2(x)g(x))dx \\
 &= \int_a^b f_1(x)g(x)dx + \int_a^b f_2(x)g(x)dx = (f_1, g) + (f_2, g);
 \end{aligned}$$

$$\text{(iv)} \quad (cf, g) = \int_a^b cf(x)g(x)dx = c \int_a^b f(x)g(x)dx = c(f, g).$$

Definition 2.8 implies the additional properties of the inner product:

$$\text{(v)} \quad (\underline{x}, \underline{y}_1 + \underline{y}_2) = (\underline{x}, \underline{y}_1) + (\underline{x}, \underline{y}_2),$$

since

$$(\underline{x}, \underline{y}_1 + \underline{y}_2) = (\underline{y}_1 + \underline{y}_2, \underline{x}) = (\underline{y}_1, \underline{x}) + (\underline{y}_2, \underline{x}) = (\underline{x}, \underline{y}_1) + (\underline{x}, \underline{y}_2);$$

$$\text{(vi)} \quad (\underline{x}, c\underline{y}) = c(\underline{x}, \underline{y}) \text{ for all } c \in R,$$

which can be proven as follows:

$$(\underline{x}, c\underline{y}) = (c\underline{y}, \underline{x}) = c(\underline{y}, \underline{x}) = c(\underline{x}, \underline{y}).$$

$$\text{(vii)} \quad (c_1\underline{x}, c_2\underline{y}) = c_1c_2(\underline{x}, \underline{y}) \text{ for all } c_1, c_2 \in R,$$

which is the consequence of the previous property:

$$(c_1\underline{x}, c_2\underline{y}) = c_1(\underline{x}, c_2\underline{y}) = c_1c_2(\underline{x}, \underline{y}).$$

(viii) $(\underline{x}, \underline{0}) = 0$ for all \underline{x} ,

since

$$(\underline{x}, \underline{0}) = (\underline{x}, \underline{0} + \underline{0}) = (\underline{x}, \underline{0}) + (\underline{x}, \underline{0}) = 2(\underline{x}, \underline{0}),$$

from which the assertion follows immediately.

Definition 2.9. A vector space V is called an *inner-product space*, if an inner-product is defined on V .

The previous examples show that both R^n and $C[a, b]$ are inner-product spaces.

Definition 2.10. Let $\underline{x}, \underline{y}$ be two elements of an inner-product space V . We say that \underline{x} and \underline{y} are orthogonal, if $(\underline{x}, \underline{y}) = 0$.

Example 2.23. In R^n , the natural basis vectors $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$ are pair-wise orthogonal, since for all k and l ,

$$(\underline{e}_k, \underline{e}_l) = \begin{cases} 0 & \text{if } k \neq l \\ 1 & \text{if } k = l. \end{cases}$$

◆

Example 2.24. In the interval $[0, 2\pi]$ the functions $f_0(x) = 1$, $f_k(x) = \sin(kx)$ ($k \geq 1$), $g_k(x) = \cos(kx)$ ($k \geq 1$) are pair-wise orthogonal. This property has an important role in the theory of Fourier-series, and can be proved by simple integration:

$$\int_0^{2\pi} f_0(x) f_k(x) dx = \int_0^{2\pi} \sin(kx) dx = \left[\frac{-\cos(kx)}{k} \right]_0^{2\pi} = 0 \quad (k \geq 1),$$

$$\int_0^{2\pi} f_0(x) g_k(x) dx = \int_0^{2\pi} \cos(kx) dx = \left[\frac{\sin(kx)}{k} \right]_0^{2\pi} = 0 \quad (k \geq 1),$$

$$\begin{aligned} \int_0^{2\pi} f_k(x) f_l(x) dx &= \int_0^{2\pi} \sin(kx) \sin(lx) dx = \frac{1}{2} \int_0^{2\pi} [\cos((k-l)x) - \cos((k+l)x)] dx \\ &= \frac{1}{2} \int_0^{2\pi} \cos((k-l)x) dx - \frac{1}{2} \int_0^{2\pi} \cos((k+l)x) dx = 0 \quad \text{for } k \neq l, \end{aligned}$$

since both terms equal zero;

$$\begin{aligned} \int_0^{2\pi} g_k(x) g_l(x) dx &= \int_0^{2\pi} \cos(kx) \cos(lx) dx = \frac{1}{2} \int_0^{2\pi} [\cos((k-l)x) + \cos((k+l)x)] dx \\ &= \frac{1}{2} \int_0^{2\pi} \cos((k-l)x) dx + \frac{1}{2} \int_0^{2\pi} \cos((k+l)x) dx = 0 \quad \text{for } k \neq l, \end{aligned}$$

since both terms equal zero again, and finally, for all k and l ,

$$\int_0^{2\pi} f_k(x) g_l(x) dx = \int_0^{2\pi} \sin(kx) \cos(lx) dx = \frac{1}{2} \int_0^{2\pi} [\sin((k+l)x) + \sin((k-l)x)] dx$$

$$= \frac{1}{2} \int_0^{2\pi} \sin((k+l)x) dx + \frac{1}{2} \int_0^{2\pi} \sin((k-l)x) dx = 0,$$

since both terms are equal to zero even if $k = l$. This set of orthogonal functions is usually called the *trigonometric system*.

Definition 2.11. A set of elements in an inner-product space is called an *orthogonal system*, if it consists of pair-wise orthogonal elements.

For example, the natural basis in R^n , and system $\{1; \sin kx, \cos kx, k \geq 1\}$ in $C[0, 2\pi]$ are orthogonal systems.

Definition 2.12. Let \underline{x} be an element of an inner-product space. Then the *length* of \underline{x} is defined as $\sqrt{(\underline{x}, \underline{x})}$ and is denoted by $|\underline{x}|$.

For example, if $\underline{x} \in R^n$, then

$$|\underline{x}| = \sqrt{x_1^2 + \dots + x_n^2},$$

assuming that the inner product is defined the same way as in Example 2.21; and if $f \in C[a, b]$ and the inner product is defined as in Example 2.22, then

$$|f| = \sqrt{\int_a^b f^2(x) dx}.$$

Definition 2.13. A set of elements in an inner-product space is called an *orthonormal system*, if it consists of pair-wise orthogonal elements and each element from the set has unit length.

For example, the natural basis in R^n with the inner product given above is an orthonormal system, since it is an orthogonal system and for all k , $|\underline{e}_k| = 1$.

The next theorem is known as the *Cauchy-Schwarz inequality*, and it is considered as one of the most fundamental properties of inner-product spaces.

Theorem 2.10. Let $\underline{x}, \underline{y}$ be arbitrary elements of an inner-product space V . Then

$$|(\underline{x}, \underline{y})| \leq |\underline{x}| \cdot |\underline{y}|. \quad (2.3)$$

Proof. Notice first that if $\underline{x} = \underline{0}$ or/and $\underline{y} = \underline{0}$, then both sides of the inequality are equal to zero, consequently the assertion holds. Assume next that $\underline{x} \neq \underline{0}$ and $\underline{y} \neq \underline{0}$. Let λ be a real number, and consider the square of the length of the element $\underline{x} + \lambda \underline{y}$:

$$\begin{aligned} |\underline{x} + \lambda \underline{y}|^2 &= (\underline{x} + \lambda \underline{y}, \underline{x} + \lambda \underline{y}) = (\underline{x}, \underline{x} + \lambda \underline{y}) + (\lambda \underline{y}, \underline{x} + \lambda \underline{y}) \\ &= (\underline{x}, \underline{x}) + (\underline{x}, \lambda \underline{y}) + (\lambda \underline{y}, \underline{x}) + (\lambda \underline{y}, \lambda \underline{y}) \\ &= (\underline{x}, \underline{x}) + \lambda(\underline{x}, \underline{y}) + \lambda(\underline{y}, \underline{x}) + \lambda^2(\underline{y}, \underline{y}) = \lambda^2(\underline{y}, \underline{y}) + 2\lambda(\underline{x}, \underline{y}) + (\underline{x}, \underline{x}). \end{aligned}$$

Observe that this is a quadratic function of λ , the graph of which is a parabola that opens up and its value is always nonnegative (being the square of the length of an element). Therefore this quadratic function has at most one real root, otherwise between the roots it would have negative values. Therefore the discriminant is zero or negative:

$$4(\underline{x}, \underline{y})^2 - 4 \cdot (\underline{x}, \underline{x}) \cdot (\underline{y}, \underline{y}) \leq 0,$$

that is

$$(\underline{x}, \underline{y})^2 \leq (\underline{x}, \underline{x}) \cdot (\underline{y}, \underline{y}).$$

Recall that $(\underline{x}, \underline{x}) = |\underline{x}|^2$ and $(\underline{y}, \underline{y}) = |\underline{y}|^2$ to have the assertion.

Corollary. For all $\underline{x}, \underline{y}$ from an inner-product space, ♣

$$|\underline{x} + \underline{y}| \leq |\underline{x}| + |\underline{y}|,$$

which relation is called the *triangle inequality*. This name originates from the geometric representation of the sum of two-dimensional real vectors as shown in Figure 2.2, and this inequality can be interpreted as the length of a side of a triangle is never longer than the sum of the lengths of the two other sides.

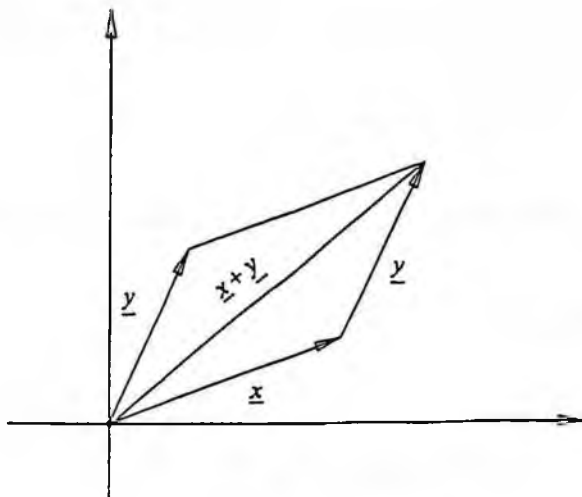


Figure 2.2. Illustration of the triangle inequality.

Proof. Simple calculation and the Cauchy-Schwarz inequality imply that

$$\begin{aligned} |\underline{x} + \underline{y}|^2 &= (\underline{x} + \underline{y}, \underline{x} + \underline{y}) = (\underline{x}, \underline{x} + \underline{y}) + (\underline{y}, \underline{x} + \underline{y}) \\ &= (\underline{x}, \underline{x}) + (\underline{x}, \underline{y}) + (\underline{y}, \underline{x}) + (\underline{y}, \underline{y}) = (\underline{x}, \underline{x}) + 2(\underline{x}, \underline{y}) + (\underline{y}, \underline{y}) \\ &= |\underline{x}|^2 + 2(\underline{x}, \underline{y}) + |\underline{y}|^2 \leq |\underline{x}|^2 + 2|\underline{x}, \underline{y}| + |\underline{y}|^2 \\ &\leq |\underline{x}|^2 + 2|\underline{x}| \cdot |\underline{y}| + |\underline{y}|^2 = (|\underline{x}| + |\underline{y}|)^2. \end{aligned}$$

♣

Definition 2.14. Let \underline{x} and \underline{y} be two nonzero elements of an inner-product space. The angle φ between these elements is defined as the unique $\varphi \in [0, \pi]$ such that

$$\cos \varphi = \frac{(\underline{x}, \underline{y})}{|\underline{x}| \cdot |\underline{y}|}. \quad (2.4)$$

The right hand side of this equation is always between -1 and $+1$, therefore φ exists and is unique. Notice that this definition is equivalent to the definition of inner-products

$$(\underline{x}, \underline{y}) = |\underline{x}| \cdot |\underline{y}| \cdot \cos \varphi$$

of two or three dimensional vectors, which is known from analytic geometry.

Example 2.25. Let

$$\underline{x} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{y} = \begin{pmatrix} -4 \\ 1 \\ 2 \end{pmatrix}.$$

Then

$$(\underline{x}, \underline{y}) = 1 \cdot (-4) + 2 \cdot 1 + 1 \cdot 2 = 0,$$

$$|\underline{x}| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6} \quad \text{and} \quad |\underline{y}| = \sqrt{(-4)^2 + 1^2 + 2^2} = \sqrt{21},$$

therefore $\cos \varphi = 0$, consequently $\varphi = 90^\circ$. The two vectors are orthogonal. ♦

We will next show that an orthogonal system is necessarily linearly independent.

Theorem 2.11. Let $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k\}$ be an orthogonal system of nonzero elements in an inner product space. Then vectors $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ are linearly independent.

Proof. Assume that $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ are linearly dependent. Then with some scalars,

$$a_1 \underline{x}_1 + \dots + a_i \underline{x}_i + \dots + a_k \underline{x}_k = \underline{0},$$

where at least one coefficient is nonzero. Say, $a_i \neq 0$. Multiply both sides by \underline{x}_i to see that

$$a_1 (\underline{x}_1, \underline{x}_i) + \dots + a_i (\underline{x}_i, \underline{x}_i) + \dots + a_k (\underline{x}_k, \underline{x}_i) = (\underline{0}, \underline{x}_i). \tag{2.5}$$

The right hand side is zero, and the orthogonality of the vectors $\underline{x}_1, \dots, \underline{x}_k$ implies that

$$(\underline{x}_1, \underline{x}_i) = \dots = (\underline{x}_{i-1}, \underline{x}_i) = (\underline{x}_{i+1}, \underline{x}_i) = \dots = (\underline{x}_k, \underline{x}_i) = 0,$$

therefore equation (2.5) can be rewritten as

$$a_i(\underline{x}_i, \underline{x}_i) = 0.$$

This is a contradiction, since $a_i \neq 0$ and $(\underline{x}_i, \underline{x}_i) > 0$.

✱

We have seen earlier that vector space R^n has an orthonormal basis consisting of the vectors $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$. In the next theorem we will show that any finite dimensional inner-product space has an orthonormal basis.

Theorem 2.12. Any finite dimensional inner-product space V has an orthonormal basis.

Proof. Assume that $\dim(V) = k$, and let $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k\}$ denote a basis of V . Based on these elements we introduce an inductive procedure for constructing an orthonormal basis. The method is the following. Select

$$\underline{z}_1 = \frac{\underline{y}_1}{|\underline{y}_1|}, \quad (2.6)$$

where $\underline{y}_1 = \underline{x}_1$, and for $i = 1, 2, 3, \dots, k-1$, let

$$\underline{z}_{i+1} = \frac{\underline{y}_{i+1}}{|\underline{y}_{i+1}|} \quad (2.7)$$

with

$$\underline{y}_{i+1} = \underline{x}_{i+1} - \sum_{j=1}^i (\underline{x}_{i+1}, \underline{z}_j) \underline{z}_j. \quad (2.8)$$

In order to prove that $\{\underline{z}_1, \underline{z}_2, \dots, \underline{z}_k\}$ is an orthonormal basis we have to verify that

- (a) $|\underline{z}_i| = 1$ for $i = 1, 2, \dots, k$;
- (b) $(\underline{z}_i, \underline{z}_j) = 0$ for all $i \neq j$;
- (c) set $\{\underline{z}_1, \dots, \underline{z}_k\}$ is a generating system of V .

Property (a) follows from (2.7), since

$$|\underline{z}_i|^2 = (\underline{z}_i, \underline{z}_i) = \left(\frac{\underline{y}_i}{|\underline{y}_i|}, \frac{\underline{y}_i}{|\underline{y}_i|} \right) = \frac{1}{|\underline{y}_i|^2} (\underline{y}_i, \underline{y}_i) = \frac{1}{|\underline{y}_i|^2} |\underline{y}_i|^2 = 1.$$

The orthogonality of the elements $\underline{z}_1, \dots, \underline{z}_k$ is proved by finite induction. Notice first that

$$(\underline{z}_2, \underline{z}_1) = \left(\frac{\underline{y}_2}{|\underline{y}_2|}, \frac{\underline{y}_1}{|\underline{y}_1|} \right) = \frac{1}{|\underline{y}_1| \cdot |\underline{y}_2|} (\underline{y}_2, \underline{y}_1),$$

where

$$\begin{aligned} (\underline{y}_2, \underline{y}_1) &= (\underline{x}_2 - (\underline{x}_2, \underline{z}_1)\underline{z}_1, \underline{y}_1) = (\underline{x}_2, \underline{y}_1) - (\underline{x}_2, \underline{z}_1) \cdot (\underline{z}_1, \underline{y}_1) \\ &= (\underline{x}_2, \underline{y}_1) - \left(\underline{x}_2, \frac{\underline{y}_1}{|\underline{y}_1|} \right) \cdot \left(\frac{\underline{y}_1}{|\underline{y}_1|}, \underline{y}_1 \right) = (\underline{x}_2, \underline{y}_1) - \frac{1}{|\underline{y}_1|} (\underline{x}_2, \underline{y}_1) \cdot \frac{1}{|\underline{y}_1|} (\underline{y}_1, \underline{y}_1) \\ &= (\underline{x}_2, \underline{y}_1) - \frac{1}{|\underline{y}_1|^2} (\underline{x}_2, \underline{y}_1) \cdot |\underline{y}_1|^2 = 0. \end{aligned}$$

Assume next that for some i , \underline{z}_i is orthogonal to $\underline{z}_1, \dots, \underline{z}_{i-1}$. Then we will prove that \underline{z}_{i+1} is orthogonal to all elements $\underline{z}_1, \underline{z}_2, \dots, \underline{z}_i$. Let l ($1 \leq l \leq i$) be arbitrary. Then

$$(\underline{z}_{i+1}, \underline{z}_l) = \left(\frac{\underline{y}_{i+1}}{|\underline{y}_{i+1}|}, \frac{\underline{y}_l}{|\underline{y}_l|} \right) = \frac{1}{|\underline{y}_{i+1}| \cdot |\underline{y}_l|} (\underline{y}_{i+1}, \underline{y}_l),$$

where

$$\begin{aligned} (\underline{y}_{i+1}, \underline{y}_l) &= \left(\underline{x}_{i+1} - \sum_{j=1}^i (\underline{x}_{i+1}, \underline{z}_j) \underline{z}_j, \underline{y}_l \right) = (\underline{x}_{i+1}, \underline{y}_l) \\ &\quad - \sum_{j=1}^i (\underline{x}_{i+1}, \underline{z}_j) \cdot (\underline{z}_j, \underline{y}_l) \\ &= (\underline{x}_{i+1}, \underline{y}_l) - \sum_{j=1}^i (\underline{x}_{i+1}, \underline{z}_j) \cdot (\underline{z}_j, |\underline{y}_l| \cdot \underline{z}_l) = (\underline{x}_{i+1}, \underline{y}_l) \\ &\quad - \sum_{j=1}^i (\underline{x}_{i+1}, \underline{z}_j) \cdot |\underline{y}_l| \cdot (\underline{z}_j, \underline{z}_l). \end{aligned}$$

Notice that $(\underline{z}_j, \underline{z}_l) = 0$ if $j \neq l$ and $(\underline{z}_l, \underline{z}_l) = 1$, therefore

$$\begin{aligned} (\underline{y}_{i+1}, \underline{y}_l) &= (\underline{x}_{i+1}, \underline{y}_l) - (\underline{x}_{i+1}, \underline{z}_l) \cdot |\underline{y}_l| \cdot (\underline{z}_l, \underline{z}_l) = (\underline{x}_{i+1}, \underline{y}_l) - (\underline{x}_{i+1}, |\underline{y}_l| \cdot \underline{z}_l) \\ &= 0, \end{aligned}$$

since $|\underline{y}_l| \cdot \underline{z}_l = \underline{y}_l$. Hence, property (b) is verified.

In order to prove that $\{\underline{z}_1, \dots, \underline{z}_k\}$ is a basis in V it is sufficient to mention that $\dim(V) = k$ and system $\{\underline{z}_1, \dots, \underline{z}_k\}$ consists of k linearly independent elements.

Remark. Procedure (2.6), (2.7), (2.8) is known as the *Gram-Schmidt orthogonalization process*.

Assume now that $\{\underline{z}_1, \underline{z}_2, \dots, \underline{z}_k\}$ is an orthonormal basis in V . The arbitrary vectors $\underline{x}, \underline{y} \in V$ can be uniquely expressed as the linear combinations of the basis elements:

$$\underline{x} = x_1 \underline{z}_1 + x_2 \underline{z}_2 + \dots + x_k \underline{z}_k$$

and

$$\underline{y} = y_1 \underline{z}_1 + y_2 \underline{z}_2 + \dots + y_k \underline{z}_k,$$

where x_i and y_i are real scalars for $i = 1, 2, \dots, k$. Then the inner product of \underline{x} and \underline{y} has the simple form:

$$(\underline{x}, \underline{y}) = \left(\sum_{i=1}^k x_i \underline{z}_i, \sum_{j=1}^k y_j \underline{z}_j \right) = \sum_{i=1}^k \sum_{j=1}^k x_i y_j (\underline{z}_i, \underline{z}_j) = \sum_{i=1}^k x_i y_i,$$

since for $i \neq j$, $(\underline{z}_i, \underline{z}_j) = 0$, and for all i , $(\underline{z}_i, \underline{z}_i) = 1$. Notice that this simple derivation shows that the inner product can be easily obtained as the product of the row vector $\underline{x}^T = (x_1, x_2, \dots, x_k)$ by the column vector

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix}.$$

The Gram-Schmidt orthogonalization procedure is illustrated by the following examples.

Example 2.26. Let $V = R^3$ and consider the vectors

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{x}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \underline{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

It is easy to show that they are linearly independent, therefore they form a basis in V . The application of the Gram-Schmidt process consists of the following elementary steps.

For $i = 1$,

$$\underline{y}_1 = \underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\underline{y}_1| = \sqrt{1^2 + 0^2 + 0^2} = 1, \quad \underline{z}_1 = \frac{\underline{y}_1}{|\underline{y}_1|} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

For $i = 2$,

$$\underline{y}_2 = \underline{x}_2 - (\underline{x}_2, \underline{z}_1)\underline{z}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\underline{y}_2| = \sqrt{0^2 + 1^2 + 0^2} = 1,$$

and so,

$$\underline{z}_2 = \frac{\underline{y}_2}{|\underline{y}_2|} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

For $i = 3$,

$$\underline{y}_3 = \underline{x}_3 - (\underline{x}_3, \underline{z}_1)\underline{z}_1 - (\underline{x}_3, \underline{z}_2)\underline{z}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and since $|\underline{y}_3| = \sqrt{0^2 + 0^2 + 1^2} = 1$,

$$\underline{z}_3 = \frac{\underline{y}_3}{|\underline{y}_3|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Example 2.27. Consider next the set of all real polynomials of degree at most 2 on the interval $[-1, 1]$. We know that

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2$$

is a basis. In this case the Gram-Schmidt orthogonalization process has the following steps, where for the sake of using usual notation, q and r will be used instead of y and z .

For $i = 0$,

$$q_0(x) = p_0(x) = 1, \quad |q_0| = \sqrt{\int_{-1}^1 1 dx} = \sqrt{2},$$

$$r_0(x) = \frac{q_0(x)}{|q_0|} = \frac{1}{\sqrt{2}}.$$

For $i = 1$,

$$q_1(x) = p_1(x) - (p_1, r_0)r_0(x) = x$$

since

$$(p_1, r_0) = \int_{-1}^1 x \cdot \frac{1}{\sqrt{2}} dx = \left[\frac{x^2}{2\sqrt{2}} \right]_{-1}^1 = 0.$$

Notice that

$$|q_1| = \sqrt{\int_{-1}^1 x^2 dx} = \sqrt{\left[\frac{x^3}{3} \right]_{-1}^1} = \sqrt{\frac{2}{3}},$$

therefore

$$r_1(x) = \frac{q_1(x)}{|q_1|} = \frac{x}{\sqrt{\frac{2}{3}}} = \sqrt{\frac{3}{2}} \cdot x.$$

For $i = 2$,

$$q_2(x) = p_2(x) - (p_2, r_0)r_0(x) - (p_2, r_1)r_1(x).$$

Simple calculation shows that

$$(p_2, r_0) = \int_{-1}^1 x^2 \frac{1}{\sqrt{2}} dx = \left[\frac{x^3}{3\sqrt{2}} \right]_{-1}^1 = \frac{2}{3\sqrt{2}} = \frac{\sqrt{2}}{3},$$

$$(p_2, r_1) = \int_{-1}^1 x^2 \sqrt{\frac{3}{2}} x dx = \left[\frac{x^4 \sqrt{3}}{4\sqrt{2}} \right]_{-1}^1 = 0,$$

which imply that

$$q_2(x) = x^2 - \frac{\sqrt{2}}{3} \frac{1}{\sqrt{2}} - 0 \cdot \sqrt{\frac{3}{2}} \cdot x = x^2 - \frac{1}{3}.$$

Since

$$\begin{aligned}
 |q_2| &= \sqrt{\int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx} = \sqrt{\int_{-1}^1 \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right) dx} = \sqrt{\left[\frac{x^5}{5} - \frac{2x^3}{9} + \frac{x}{9}\right]_{-1}^1} \\
 &= \sqrt{\frac{2}{5} - \frac{4}{9} + \frac{2}{9}} = \sqrt{\frac{8}{45}}, \\
 r_2(x) &= \frac{x^2 - \frac{1}{3}}{\sqrt{\frac{8}{45}}} = \sqrt{\frac{45}{8}}x^2 - \sqrt{\frac{5}{8}}.
 \end{aligned}$$

Hence

$$r_0(x) = \frac{1}{\sqrt{2}}, r_1(x) = \sqrt{\frac{3}{2}} \cdot x, r_2(x) = \sqrt{\frac{45}{8}}x^2 - \sqrt{\frac{5}{8}}$$

is an orthonormal basis. ♦

In this section inner products were defined only in real vector spaces, and the properties of only real inner product spaces have been examined. Similar to the real case, inner products in complex vector spaces as well as complex inner product spaces can be introduced. If V is a complex vector space, then it is assumed that the inner product is a bivariable function that assigns to each pair $\underline{x}, \underline{y}$ of elements of V a complex number denoted by $(\underline{x}, \underline{y})$, such that

- (i) $(\underline{x}, \underline{x}) \geq 0$ for all $\underline{x} \in V$, and $(\underline{x}, \underline{x}) = 0$ if and only if $\underline{x} = \underline{0}$;
- (ii) $(\underline{x}, \underline{y}) = \overline{(\underline{y}, \underline{x})}$ $\underline{x}, \underline{y} \in V$, where overbar denotes complex

conjugate;

- (iii) $(\underline{x}_1 + \underline{x}_2, \underline{y}) = (\underline{x}_1, \underline{y}) + (\underline{x}_2, \underline{y})$ for all $\underline{x}_1, \underline{x}_2, \underline{y} \in V$;

(iv) $(c\underline{x}, \underline{y}) = c(\underline{x}, \underline{y})$ for all $\underline{x}, \underline{y} \in V$ and complex number c .

Notice that properties (ii) and (iv) imply that for all $\underline{x}, \underline{y} \in V$ and complex number c ,

$$(\underline{x}, c\underline{y}) = \overline{(c\underline{y}, \underline{x})} = \overline{c \cdot (\underline{y}, \underline{x})} = \overline{c} \cdot \overline{(\underline{y}, \underline{x})} = \overline{c} \cdot (\underline{x}, \underline{y})$$

Orthogonality in complex inner product spaces can be defined in the same way as it has been done for the real case. With obvious modifications all previous results and the results of the next section can be extended to the complex case. For example, assume that $(\underline{z}_1, \underline{z}_2, \dots, \underline{z}_k)$ is an orthonormal basis in a complex inner product space V , then any arbitrary elements $\underline{x}, \underline{y} \in V$ can be uniquely represented as

$$\underline{x} = x_1 \underline{z}_1 + x_2 \underline{z}_2 + \dots + x_k \underline{z}_k$$

and

$$\underline{y} = y_1 \underline{z}_1 + y_2 \underline{z}_2 + \dots + y_k \underline{z}_k,$$

where x_i, y_i are complex scalars for $i = 1, 2, \dots, k$. Then the inner product of \underline{x} and \underline{y} can be given as

$$(\underline{x}, \underline{y}) = \left(\sum_{i=1}^k x_i \underline{z}_i, \sum_{j=1}^k y_j \underline{z}_j \right) = \sum_{i=1}^k \sum_{j=1}^k x_i \overline{y_j} (\underline{z}_i, \underline{z}_j) = \sum_{i=1}^k x_i \overline{y_i},$$

since for $i \neq j$, $(\underline{z}_i, \underline{z}_j) = 0$, and for all i , $(\underline{z}_i, \underline{z}_i) = 1$. Notice that in the above equation $\overline{y_i}$ denotes the complex conjugate of y_i . Hence the inner

product of \underline{x} and \underline{y} can be easily obtained as the product of the row vector $\underline{x}^T = (x_1, x_2, \dots, x_k)$ by the column vector

$$\underline{y} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_k \end{pmatrix}.$$

Finally we note that in the linear algebra literature real and complex inner-product spaces are called *Euclidean* and *Unitary spaces*, respectively.

2.5 Direct Sums and Orthogonal Complementary Subspaces

Let V be a vector-space, and let V_1 and V_2 be two subspaces of V . We do not assume first that there is an inner product in V . From Theorem 2.3 we know that the intersection $V_1 \cap V_2$ is also a subspace of V . However the union $V_1 \cup V_2$ is not necessarily a subspace as the following example illustrates.

Example 2.28. Assume that $V = \mathbb{R}^2$, and let $V_1 = \left\{ a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid a \text{ is real} \right\}$, and $V_2 = \left\{ a \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid a \text{ is real} \right\}$. The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are in V_1 and V_2 , respectively. However their sum, the vector $\underline{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ does not belong to $V_1 \cup V_2$, since $\underline{x} \notin V_1$ and $\underline{x} \notin V_2$.



The subspace generated by the union $V_1 \cup V_2$ is the smallest subspace containing both V_1 and V_2 . By using the notation introduced in Section 2.2, we might use the symbols $V(V_1 \cup V_2)$ or $V(V_1, V_2)$ to denote this subspace. The construction of $V(V_1 \cup V_2)$ is given in the following theorem.

Theorem 2.13. Let V_1 and V_2 be subspaces of a vector space V . Then

$$V(V_1 \cup V_2) = \{\underline{x} + \underline{y} \mid \underline{x} \in V_1 \text{ and } \underline{y} \in V_2\}. \quad (2.9)$$

Proof. Let U denote the right hand side of (2.9). First we show that any element of $V(V_1 \cup V_2)$ can be written as the sum of an element of V_1 and an element of V_2 . Theorem 2.4 implies that any arbitrary element $\underline{z} \in V(V_1 \cup V_2)$ can be written as a linear combination of finitely many elements of $V_1 \cup V_2$, that is,

$$\underline{z} = a_1 \underline{x}_1 + \dots + a_k \underline{x}_k + b_1 \underline{y}_1 + \dots + b_l \underline{y}_l,$$

where $\underline{x}_1, \dots, \underline{x}_k \in V_1$ and $\underline{y}_1, \dots, \underline{y}_l \in V_2$. Here we used the fact that each element of the union $V_1 \cup V_2$ belongs either to V_1 or to V_2 . Since V_1 and V_2 are subspaces,

$$\underline{x} = a_1 \underline{x}_1 + \dots + a_k \underline{x}_k \in V_1 \text{ and } \underline{y} = b_1 \underline{y}_1 + \dots + b_l \underline{y}_l \in V_2,$$

and therefore, $\underline{z} = \underline{x} + \underline{y}$. Hence $V(V_1 \cup V_2) \subseteq U$.

In order to complete the proof of the theorem we have to show that $U \subseteq V(V_1 \cup V_2)$ by verifying that if $\underline{x} \in V_1$ and $\underline{y} \in V_2$ are arbitrary elements, then $\underline{x} + \underline{y} \in V(V_1 \cup V_2)$. Since V_1 and V_2 are subsets of $V(V_1 \cup V_2)$, both \underline{x} and \underline{y} belong to $V(V_1 \cup V_2)$, and the fact that $V(V_1 \cup V_2)$ is a subspace implies that $\underline{x} + \underline{y} \in V(V_1 \cup V_2)$. Therefore $U \subseteq V(V_1 \cup V_2)$, which completes the proof.

+

Remark. The above theorem guarantees that every element of the generated subspace $V(V_1 \cup V_2)$ can be written as the sum of an element of V_1 and an element of V_2 . In most cases this decomposition is not unique, as is shown in the following example.

Example 2.29. Select the vector space $V = R^3$, and the vectors

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ and } \underline{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Let V_1 be the subspace generated by \underline{e}_1 and \underline{e}_2 and let V_2 be generated by \underline{e}_2 and \underline{e}_3 . Then

$$V_1 = \left\{ a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \mid a_1 \text{ and } a_2 \text{ are real} \right\}$$

and

$$V_2 = \left\{ b_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + b_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ b_1 \\ b_2 \end{pmatrix} \mid b_1 \text{ and } b_2 \text{ are real} \right\}.$$

In Example 2.13 we have seen that \underline{e}_1 , \underline{e}_2 and \underline{e}_3 generate the entire vector space R^3 , therefore $V(V_1 \cup V_2) = R^3$, since all of the vectors \underline{e}_1 , \underline{e}_2 and \underline{e}_3 belong to $V_1 \cup V_2$. Consider next a three-element vector with entries x_1, x_2 , and x_3 . Therefore it can be rewritten as the sum of a vector of V_1 and a vector of V_2 . That is,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b_1 \\ b_2 \end{pmatrix}. \quad (2.10)$$

Comparing the corresponding elements of the left-hand side and right-hand side vectors we see that this equation is equivalent to the following system of equations:

$$\begin{aligned} x_1 &= a_1 \\ x_2 &= a_2 + b_1 \\ x_3 &= b_2. \end{aligned}$$

If vector \underline{x} is given, then the values of x_1, x_2 , and x_3 are also given. The values of a_1 and b_2 are unique, however x_2 can be rewritten as the sum of two real numbers in infinitely many different ways, therefore infinitely many decompositions (2.10) exist.

The following result gives a sufficient and necessary condition for the uniqueness of the decomposition of the elements of $V(V_1 \cup V_2)$ as the sum of an element of V_1 and an element of V_2 .

Theorem 2.14. Let V_1 and V_2 be two subspaces of a vector space V . Each element of $V(V_1 \cup V_2)$ can be uniquely written as $\underline{x} + \underline{y}$ with $\underline{x} \in V_1$ and $\underline{y} \in V_2$ if and only if the intersection $V_1 \cap V_2$ consists of only the zero vector.

Proof. Assume first, that $V_1 \cap V_2 = \{0\}$, and assume in addition, that for a \underline{z} , two such decompositions hold:

$$\underline{z} = \underline{x}_1 + \underline{y}_1 = \underline{x}_2 + \underline{y}_2,$$

where $\underline{x}_1, \underline{x}_2 \in V_1$ and $\underline{y}_1, \underline{y}_2 \in V_2$. We can rewrite this equality as

$$\underline{x}_1 - \underline{x}_2 = \underline{y}_2 - \underline{y}_1.$$

The left-hand side is in V_1 and the right-hand side is in V_2 , therefore both sides are in $V_1 \cap V_2$, which consists of only the zero element. So,

$$\underline{x}_1 - \underline{x}_2 = \underline{y}_2 - \underline{y}_1 = \underline{0},$$

that is, $\underline{x}_1 = \underline{x}_2$ and $\underline{y}_1 = \underline{y}_2$ proving the uniqueness of the decomposition.

Assume next, that $V_1 \cap V_2$ contains at least one nonzero element which can be denoted by \underline{w} . If $\underline{z} = \underline{x} + \underline{y}$ with some $\underline{x} \in V_1$ and $\underline{y} \in V_2$, then

$\underline{z} = (\underline{x} + \underline{w}) + (\underline{y} - \underline{w})$, where $\underline{x} \neq \underline{x} + \underline{w} \in V_1$ and $\underline{y} \neq \underline{y} - \underline{w} \in V_2$ showing that the decomposition is not unique.

♣

Definition 2.15. Let V_1 and V_2 be subspaces of a vector space V such that $V_1 \cap V_2 = \{0\}$. Then the subspace $V(V_1 \cup V_2)$ is called the *direct sum* of V_1 and V_2 , and is denoted by $V_1 \oplus V_2$.

From Theorem 2.14 we know that any element of $V_1 \oplus V_2$ can be uniquely represented as $\underline{x} + \underline{y}$ with some $\underline{x} \in V_1$ and $\underline{y} \in V_2$. This fact has the following consequence. Assume that V_1 and V_2 are finitely generated, and let $(\underline{x}_1, \dots, \underline{x}_k)$ and $(\underline{y}_1, \dots, \underline{y}_l)$ be a basis of V_1 and V_2 , respectively. Then $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l\}$ is a basis of $V_1 \oplus V_2$. In order to prove this assertion we have to show that

- (a) these elements are linearly independent;
- (b) they generate the entire subspace $V_1 \oplus V_2$.

In order to verify (a), assume that a linear combination of the elements $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l$ is zero. Then with some scalars a_i and b_j ($1 \leq i \leq k, 1 \leq j \leq l$),

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k + b_1 \underline{y}_1 + \dots + b_l \underline{y}_l = \underline{0},$$

which implies that

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k = (-b_1) \underline{y}_1 + \dots + (-b_l) \underline{y}_l.$$

The left-hand side is in V_1 , the right-hand side is in V_2 , therefore both are in $V_1 \cap V_2$, which contains only the zero element. Therefore,

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k = \underline{0} \text{ and } (-b_1) \underline{y}_1 + \dots + (-b_l) \underline{y}_l = \underline{0}.$$

Since $\underline{x}_1, \dots, \underline{x}_k$ as well as $\underline{y}_1, \dots, \underline{y}_l$ are linearly independent, $a_1 = \dots = a_k = 0$ and $b_1 = \dots = b_l$ showing that the set $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l\}$ consists of only linearly independent elements. In order to prove statement (b), consider an arbitrary element \underline{z} of $V_1 \oplus V_2$. Then with some $\underline{x} \in V_1$ and $\underline{y} \in V_2$, $\underline{z} = \underline{x} + \underline{y}$. Since $\{\underline{x}_1, \dots, \underline{x}_k\}$ is a basis of V_1 , $\underline{x} = a_1 \underline{x}_1 + \dots + a_k \underline{x}_k$ with some scalars a_1, \dots, a_k , and since $\{\underline{y}_1, \dots, \underline{y}_l\}$ is a basis in V_2 , $\underline{y} = b_1 \underline{y}_1 + \dots + b_l \underline{y}_l$. Therefore

$$\underline{z} = \underline{x} + \underline{y} = a_1 \underline{x}_1 + \dots + a_k \underline{x}_k + b_1 \underline{y}_1 + \dots + b_l \underline{y}_l,$$

that is, \underline{z} is a linear combination of vectors $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l$. Hence, they generate the entire direct sum $V_1 \oplus V_2$. This statement has the following important consequence:

$$\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2), \quad (2.11)$$

since the dimension of finitely generated subspaces equals the number of basis elements.

Let's drop next the assumption that $V_1 \cap V_2 = \{0\}$. Then $V_1 \oplus V_2$ is not defined, but the subspace $V(V_1 \cup V_2)$ generated by the union of V_1 and V_2 is defined. We will next show that a straightforward generalization of relation (2.11) holds in this more general case.

Theorem 2.15. Let V_1 and V_2 be finitely generated subspaces. Then

$$\dim(V(V_1 \cup V_2)) + \dim(V_1 \cap V_2) = \dim(V_1) + \dim(V_2).$$

Proof. If $V_1 \cap V_2 = \{0\}$, then relation (2.11) is equivalent to the assertion, since $\dim(\{0\}) = 0$. Assume next that $V_1 \cap V_2 \neq \{0\}$. Let $\{\underline{x}_1, \dots, \underline{x}_k\}$ be a basis of $V_1 \cap V_2$. Corollary 2 of Theorem 2.9 implies the existence of elements $\underline{y}_{k+1}, \dots, \underline{y}_m$ and $\underline{z}_{k+1}, \dots, \underline{z}_n$ such that $\{\underline{x}_1, \dots, \underline{x}_k\}$ is a basis of V_1 and $\{\underline{x}_1, \dots, \underline{x}_k, \underline{z}_{k+1}, \dots, \underline{z}_n\}$ is a basis of V_2 . By using the method which was applied in proving relation (2.11) one may easily verify that $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}, \dots, \underline{y}_m, \underline{z}_{k+1}, \dots, \underline{z}_n\}$ is a basis of $V(V_1 \cup V_2)$. That is, $\dim(V(V_1 \cup V_2)) = m + n - k$. The assertion then becomes clear, since $\dim(V_1 \cap V_2) = k$, $\dim(V_1) = m$, and $\dim(V_2) = n$.

♣

Direct sums of more than two subspaces can be defined in the following way.

Definition 2.16. Let V_1, V_2, \dots, V_k be subspaces of V . Then $V(V_1, V_2, \dots, V_k)$ is said to be the *direct sum* of V_1, \dots, V_k , if each element $\underline{v} \in V(V_1, V_2, \dots, V_k)$ can be uniquely expressed as a sum

$$\underline{v} = \underline{v}_1 + \dots + \underline{v}_k, \tag{2.12}$$

where for $i = 1, 2, \dots, k$, $\underline{v}_i \in V_i$. In this case we use the notation

$$V(V_1, V_2, \dots, V_k) = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

Notice first that for $k = 2$, this definition is equivalent to Definition 2.15. To show that a vector space (or a subspace) V is the direct sum of given subspaces V_1, V_2, \dots, V_k , we have to verify two things. First, we have to prove that each element can be written as the sum (2.12), and second, that this decomposition is unique for all $\underline{v} \in V$. In practical cases, however we do not need to check uniqueness for all elements of \underline{v} as it is given in the following result.

Theorem 2.16. Let V be a vector space (or a subspace) and V_1, V_2, \dots, V_k be subspaces of V . Then $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ if and only if

a) for all $\underline{v} \in V$, \underline{v} can be expressed as

$$\underline{v} = \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_k,$$

where $\underline{v}_i \in V_i$ ($i = 1, 2, \dots, k$);

b) if $\underline{v}_i \in V_i$ ($i = 1, 2, \dots, k$) are elements such that

$$\underline{v}_1 + \dots + \underline{v}_k = \underline{0}, \tag{2.13}$$

then $\underline{v}_1 = \underline{v}_2 = \dots = \underline{v}_k = \underline{0}$.

Proof. If $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$, then (a) is obviously satisfied. Since the selection $\underline{v}_1 = \underline{v}_2 = \dots = \underline{v}_k = \underline{0}$ satisfies equation (2.13), the uniqueness of such decomposition implies that necessarily $\underline{v}_i = \underline{0}$ for all i .

Assume next that conditions (a) and (b) are satisfied. From (a) we know that all elements \underline{v} of V can be written as the sum (2.12). In order to show that $V = V_1 \oplus V_2 \oplus \dots \oplus V_k$ it is sufficient to prove that decomposition (2.12) is unique. Assume that for some $\underline{v} \in V$, we have two such decompositions:

$$\underline{v} = \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_k = \underline{v}'_1 + \underline{v}'_2 + \dots + \underline{v}'_k,$$

where $\underline{v}_i, \underline{v}'_i \in V_i$ for all i . Then

$$(\underline{v}_1 - \underline{v}'_1) + (\underline{v}_2 - \underline{v}'_2) + \dots + (\underline{v}_k - \underline{v}'_k) = \underline{0},$$

and since $\underline{v}_i, \underline{v}'_i \in V_i$, condition (b) implies that for all i , $\underline{v}_i - \underline{v}'_i = \underline{0}$, that is, $\underline{v}_i = \underline{v}'_i$ proving the required uniqueness.

♣

We conclude this section with a particular case of direct sums of subspaces of inner-product spaces, which will play important roles in later chapters of this book.

Definition 2.17. Let V_1 and V_2 be two subspaces of an inner-product space V . We say that V_2 is the *orthogonal complementary subspace* of V_1 in V , if

(i) $V_1 \cap V_2 = \{\underline{0}\}$;

(ii) $V_1 \oplus V_2 = V$;

(iii) for all $\underline{x} \in V_1$ and $\underline{y} \in V_2$, $(\underline{x}, \underline{y}) = 0$. That is, all elements of V_1 are orthogonal to all elements of V_2 .

If V_2 is the orthogonal complementary subspace of V_1 then V_1 is also the orthogonal complementary subspace of V_2 in the same inner-product space.

Example 2.30. Select $V = R^2$, and let $V_1 = V(\underline{e}_1)$ and $V_2 = V(\underline{e}_2)$. That is, V_1 and V_2 are the subspaces generated by vectors \underline{e}_1 and \underline{e}_2 , respectively. We will show that V_1 and V_2 are orthogonal complementary subspaces in V . Notice first, that

$$V_1 = \left\{ a \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} \mid a \text{ is real} \right\}$$

and

$$V_2 = \left\{ b \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \mid b \text{ is real} \right\}.$$

Property (i) is obvious, since for any vector $\underline{x} = (x_i)$ such that $\underline{x} \in V_1 \cap V_2$, $x_2 = 0$ (since $\underline{x} \in V_1$) and $x_1 = 0$ (since $\underline{x} \in V_2$), that is $\underline{x} = \underline{0}$. In order to prove condition (ii) assume that $\underline{x} = (x_i)$ is an arbitrary vector in R^2 . Then $\underline{x} = x_1 \underline{e}_1 + x_2 \underline{e}_2$, where the first term is in V_1 and the second term is in V_2 , and this is the unique such decomposition. Property (iii) is obvious, since with arbitrary real numbers a and b , the inner product of vectors $\begin{pmatrix} a \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ b \end{pmatrix}$ is zero, since

$$\left(\begin{pmatrix} a \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ b \end{pmatrix} \right) = (a, 0) \begin{pmatrix} 0 \\ b \end{pmatrix} = a \cdot 0 + 0 \cdot b = 0.$$



Consider next a subspace $V_1 \neq \{0\}$ in a finitely generated inner-product space V , and let V_2 be the orthogonal complementary subspace of V_1 in V . The following characterization results are useful in constructing V_2 .

Theorem 2.17. The orthogonal complementary subspace consists of all elements of V which are orthogonal to all elements of V_1 , that is,

$$V_2 = \{ \underline{y} \mid \underline{y} \in V, \text{ such that } (\underline{y}, \underline{x}) = 0 \text{ for all } \underline{x} \in V_1 \}.$$

Proof. We have to verify that all conditions of Definition 2.17 are satisfied.

(i) If $\underline{x} \in V_1 \cap V_2$, then $\underline{x} \in V_1$ and $\underline{x} \in V_2$, therefore \underline{x} is orthogonal to itself. That is, $(\underline{x}, \underline{x}) = 0$, which implies that $\underline{x} = \underline{0}$;

(ii) Let $\{\underline{x}_1, \dots, \underline{x}_k\}$ be a basis in V_1 . Corollary 2 of Theorem 2.9 implies that there are elements $\underline{y}_{k+1}, \dots, \underline{y}_n \in V$ such that $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_{k+1}, \dots, \underline{y}_n\}$ is a basis of V . Apply the Gram-Schmidt orthogonalization process for these vectors, then an orthonormal basis $\{\underline{z}_1, \dots, \underline{z}_k, \underline{z}_{k+1}, \dots, \underline{z}_n\}$ of V is obtained. From the inductive nature of the procedure it follows that $\{\underline{z}_1, \dots, \underline{z}_k\}$ is a basis of V_1 . Denote the subspace generated by the other basis elements $\underline{z}_{k+1}, \dots, \underline{z}_n$ by W . We will verify that $W = V_2$. First we show that $W \subseteq V_2$. Let $\underline{z} \in W$ be arbitrary, then

$$\underline{z} = a_{k+1} \underline{z}_{k+1} + \dots + a_n \underline{z}_n,$$

and if $\underline{x} \in V_1$, then

$$\underline{x} = b_1 \underline{z}_1 + \dots + b_k \underline{z}_k.$$

Obviously, \underline{z} is orthogonal to \underline{x} , since

$$(\underline{z}, \underline{x}) = \left(\sum_{i=k+1}^n a_i \underline{z}_i, \sum_{j=1}^k b_j \underline{z}_j \right) = \sum_{i=k+1}^n \sum_{j=1}^k a_i b_j (\underline{z}_i, \underline{z}_j) = 0.$$

That is, all elements of \mathcal{W} are orthogonal to all elements of V_1 implying that $\mathcal{W} \subseteq V_2$. Next we show that $V_2 \subseteq \mathcal{W}$. Let $\underline{y} \in V_2$ be an arbitrary element. Since $V_2 \subseteq V$, necessarily $\underline{y} \in V$, therefore

$$\underline{y} = a_1 \underline{z}_1 + \dots + a_k \underline{z}_k + a_{k+1} \underline{z}_{k+1} + \dots + a_n \underline{z}_n$$

with some scalars a_1, \dots, a_n . In order to verify that $\underline{y} \in \mathcal{W}$ we have to prove that $a_1 = \dots = a_k = 0$. For $i = 1, 2, \dots, k$, multiply both sides of the above equation by \underline{z}_i to see that

$$(\underline{z}_i, \underline{y}) = a_1 (\underline{z}_i, \underline{z}_1) + \dots + a_i (\underline{z}_i, \underline{z}_i) + \dots + a_n (\underline{z}_i, \underline{z}_n).$$

Notice that $\underline{z}_i \in V_1$ and $\underline{y} \in V_2$, therefore the left-hand side is zero. If $i \neq j$, $(\underline{z}_i, \underline{z}_j) = 0$, which implies that the right-hand side equals $a_i (\underline{z}_i, \underline{z}_i)$. Since $\underline{z}_i \neq \underline{0}$ (otherwise vectors $\underline{z}_1, \dots, \underline{z}_k$ would be linearly dependent), $(\underline{z}_i, \underline{z}_i) > 0$, therefore $a_i = 0$ for $i = 1, 2, \dots, k$ implying that $\underline{y} \in \mathcal{W}$. Condition (iii) follows from the definition of V_2 , which completes the proof. ◆

Example 2.31. Consider the inner-product space R^2 , and let

$$V_1 = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} \mid x \text{ is real} \right\}.$$

We first show that V_1 is a subspace by applying Theorem 2.2. Let a and b be two scalars, and \underline{x} and \underline{y} two elements from V_1 . Then with some real values x and y ,

$$\underline{x} = \begin{pmatrix} x \\ x \end{pmatrix} \text{ and } \underline{y} = \begin{pmatrix} y \\ y \end{pmatrix},$$

furthermore

$$a\underline{x} + b\underline{y} = \begin{pmatrix} ax \\ ax \end{pmatrix} + \begin{pmatrix} by \\ by \end{pmatrix} = \begin{pmatrix} ax + by \\ ax + by \end{pmatrix} \in V_1.$$

The orthogonal complementary subspace of V_1 in R^2 consists of all vectors which are orthogonal to all elements of V_1 . Now let $\underline{z} = (z_i)$ denote such a vector. Then for all $\underline{x} \in V_1$,

$$(\underline{z}, \underline{x}) = (z_1, z_2) \begin{pmatrix} x \\ x \end{pmatrix} = z_1 x + z_2 x = x(z_1 + z_2) = 0,$$

which holds for all real x if and only if $z_2 = -z_1$.

Hence

$$V_2 = \left\{ \begin{pmatrix} z \\ -z \end{pmatrix} \mid z \text{ is real} \right\}.$$

◆

2.6 Applications

In this final section two applications will be briefly outlined, and then simple examples will be presented.

1. First a simple algorithm is presented to select a maximum number of linearly independent vectors from a set of finitely many vectors. Consider vectors $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m \in R^n$. By using Theorem 2.6, the following algorithm can be suggested:

Step 1. Select $S = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m\}$, and $k = 2$.

Step 2. Check if \underline{x}_k is the linear combination of vectors $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_{k-1}$. If not, then let $k := k + 1$ and go to Step 3. Otherwise let $m := m - 1$, for $i = k, k + 1, \dots, m - 1$, set $x_i := x_{i+1}$, and go to Step 3.

Step 3. If $k \leq m$, then go to Step 2, otherwise stop. The remaining vectors in S form the requested maximum number of linearly independent vectors.

Example 2.32. Select $m = 4$ and

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

We first select $k = 2$. Since \underline{x}_2 is not a constant multiple of \underline{x}_1 , it is not a linear combination of \underline{x}_1 . Therefore we let $k = 3$, and go to Step 3. Since $k < m$, we have to go back to Step 2, where we have to check if \underline{x}_3 is a linear combination of \underline{x}_1 and \underline{x}_2 . It is easy to see that $\underline{x}_3 = \underline{x}_1 + \underline{x}_2$,

therefore we let $m = 3$, and $\underline{x}_3 = \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}$. Therefore system S is modified as

$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} \right\}$. In Step 3 we see that $k = m$, so we have to go back to

Step 2, where we see that $\underline{x}_3 = \underline{x}_1 + 2\underline{x}_2$. Therefore we let $m = 2$, and the final set becomes

$$S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

2. Assume next that system $S = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k\}$ is orthonormal, and a vector \underline{x} does not belong to the subspace generated by the elements of S . That is, \underline{x} cannot be expressed as the linear combination of vectors $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$. We will now find the linear combination of these vectors that has the minimal distance from \underline{x} . This is a particular case of the least squares problem known from statistics. Mathematically this problem can be formulated as an unconstrained optimization problem:

$$\text{Minimize } |c_1 \underline{x}_1 + c_2 \underline{x}_2 + \dots + c_k \underline{x}_k - \underline{x}|!$$

In order to find the optimal values of c_1, c_2, \dots, c_k , consider the square of the objective function:

$$\left| \sum_{i=1}^k c_i \underline{x}_i - \underline{x} \right|^2 = \left(\sum_{i=1}^k c_i \underline{x}_i - \underline{x}, \sum_{i=1}^k c_i \underline{x}_i - \underline{x} \right) = \sum_{i=1}^k c_i^2 - 2 \sum_{i=1}^k c_i (\underline{x}_i, \underline{x}) + (\underline{x}, \underline{x}),$$

where we used the fact that the elements of S form an orthonormal system. It is easy to see that this objective function can be rewritten as

$$\sum_{i=1}^k (c_i - (\underline{x}_i, \underline{x}))^2 + (\underline{x}, \underline{x}) - \sum_{i=1}^k (\underline{x}_i, \underline{x})^2.$$

It is clear that this function is minimal if we select

$$c_i = (\underline{x}_i, \underline{x}) \quad (2.14)$$

for all $i = 1, 2, \dots, k$. If the elements of S are not orthonormal, then Theorem 2.12 implies that the subspace generated by the elements of S has an orthonormal basis $\underline{z}_1, \underline{z}_2, \dots, \underline{z}_l$. Then we have to use the above procedure with the new system $S' = \{\underline{z}_1, \underline{z}_2, \dots, \underline{z}_l\}$.

Example 2.33. Assume that

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right\} \quad \text{and} \quad \underline{x} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

First we show that \underline{x} cannot be expressed as the linear combination of the two vectors of S . The equation

$$a \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

can be rewritten as the system

$$a + b = 2$$

$$2b = 2$$

$$a + 2b = 2.$$

From the second equation we see that $b = 1$, and from the first equation we obtain that $a = 1$. However with these values the third equation is not satisfied. Notice that the inner product of the two given vectors of S equals

$$(1,0,1) \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = 1 + 0 + 2 = 3 \neq 0,$$

therefore they are not orthonormal. In order to use the above algorithm, an orthonormal basis of the subspace generated by these two vectors has to be first constructed. The Gram-Schmidt orthogonalization process will be used. In our case, $k = 2$,

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

We will follow the method introduced in the proof of Theorem 2.12. From equation (2.6) we have

$$\underline{y}_1 = \underline{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad |\underline{y}_1| = \sqrt{1^2 + 0^2 + 1^2} = \sqrt{2},$$

$$\underline{z}_1 = \frac{\underline{y}_1}{|\underline{y}_1|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{0}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Equation (2.8) implies that

$$\underline{y}_2 = \underline{x}_2 - (\underline{x}_2, \underline{z}_1)\underline{z}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{3}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{0}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 2 \\ \frac{1}{2} \end{pmatrix},$$

$$|\underline{y}_2| = \sqrt{\left(-\frac{1}{2}\right)^2 + 2^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{18}{4}} = \frac{3}{\sqrt{2}},$$

$$\underline{z}_2 = \frac{\underline{y}_2}{|\underline{y}_2|} = \begin{pmatrix} -\frac{\sqrt{2}}{6} \\ \frac{2\sqrt{2}}{3} \\ \frac{\sqrt{2}}{6} \end{pmatrix}.$$

And finally, from equation (2.14) we conclude that

$$c_1 = (\underline{z}_1, \underline{x}) = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2} \right) \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2\sqrt{2},$$

$$c_2 = (\underline{z}_2, \underline{x}) = \left(-\frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}, \frac{\sqrt{2}}{6} \right) \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \frac{4}{3}\sqrt{2},$$

and so, vector

$$c_1 \underline{z}_1 + c_2 \underline{z}_2 = 2\sqrt{2} \begin{pmatrix} \frac{\sqrt{2}}{2} \\ 2 \\ 0 \\ \frac{\sqrt{2}}{2} \\ 2 \end{pmatrix} + \frac{4}{3}\sqrt{2} \begin{pmatrix} -\frac{\sqrt{2}}{6} \\ \frac{2\sqrt{2}}{3} \\ \frac{\sqrt{2}}{6} \end{pmatrix} = \begin{pmatrix} \frac{14}{9} \\ 9 \\ \frac{16}{9} \\ 22 \\ 9 \end{pmatrix}$$

belonging to S has the smallest distance from \underline{x} .

The least squares method will be discussed later in Section 3.8, where a different solution method will be introduced.

3. Assume now that $\underline{x}_1, \dots, \underline{x}_n$ form an orthonormal basis in R^n , and construct the matrix

$$\underline{A} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$$

with columns $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$. The transpose of this matrix can be written as

$$\underline{A}^T = \begin{pmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \dots \\ \underline{x}_n^T \end{pmatrix},$$

and since $\underline{x}_i^T \underline{x}_i = 1$ and $\underline{x}_i^T \underline{x}_j = 0$ ($i \neq j$), $\underline{A}^T \underline{A} = \underline{I}$ showing that in this special case,

$$\underline{A}^{-1} = \underline{A}^T.$$

In Chapter 7 we will examine special matrices including the ones satisfying this relation.

4. If $\underline{x}_1, \dots, \underline{x}_n$ form an orthogonal basis, then $\underline{x}_i^T \underline{x}_j = 0$ for all $i \neq j$. Denote the products $\underline{x}_i^T \underline{x}_i$ by d_i . Then similar to the previous application we have

$$\underline{A}^T \underline{A} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

5. Many problems of operational research can be formulated by using inner-products of finite dimensional vectors.

The most simple example is the objective function of linear programming problem, which are usually written as $c_1 x_1 + c_2 x_2 + \dots + c_n x_n$, where c_1, c_2, \dots, c_n are given coefficients. Notice that this objective function can be rewritten in the form of $\underline{c}^T \underline{x} = (\underline{c}, \underline{x})$,

where the inner-product introduced earlier in Example 2.21 is used. Here c_i is the i^{th} entry of the column vector \underline{c} , and x_i is the i^{th} entry of \underline{x} .

Nonlinear complementarity problems can be formulated in the following way. Let function $\underline{f} : R^n \mapsto R^n$ be defined for all nonnegative vectors \underline{x} . Then the corresponding nonlinear complementarity problem has the form:

$$\begin{aligned}\underline{x}^T \underline{f}(\underline{x}) &= 0 \\ \underline{x} &\geq \underline{0} \\ \underline{f}(\underline{x}) &\geq \underline{0}.\end{aligned}$$

Using again the inner-product of Example 2.21, this problem can be reformulated as

$$\begin{aligned}(\underline{x}, \underline{f}(\underline{x})) &= 0 \\ \underline{x} &\geq \underline{0} \\ \underline{f}(\underline{x}) &\geq \underline{0}.\end{aligned}$$

Notice that for all i , either x_i or $f_i(\underline{x})$ has to be zero, where $f_i(\underline{x})$ denote the i^{th} entry of vector $\underline{f}(\underline{x})$. In the special case, when $\underline{f}(\underline{x}) = \underline{Mx} + \underline{b}$ with some $n \times n$ matrix \underline{M} and $\underline{b} \in R^n$, the problem simplifies to the following:

$$\begin{aligned}\underline{x}^T (\underline{Mx} + \underline{b}) &= 0 \\ \underline{x} &\geq \underline{0} \\ \underline{Mx} + \underline{b} &\geq \underline{0},\end{aligned}$$

which is known as a linear complementarity problem.

Variational inequalities have the usual form

$$(\underline{x} - \underline{x}^*)^T \underline{f}(\underline{x}^*) \leq 0, (\text{all } \underline{x} \in \mathbf{R}^n)$$

where $\underline{f} : \mathbf{R}^n \mapsto \mathbf{R}^n$ is a given function, and $\underline{x}^* \in \mathbf{R}^n$ is a given vector from the domain of \underline{f} . Notice again that the left hand side can be expressed as the inner-product of Example 2.21, the problem can be rewritten as

$$(\underline{x} - \underline{x}^*, \underline{f}(\underline{x}^*)) \leq 0 (\text{all } \underline{x} \in \mathbf{R}^n)$$

2.7 Exercises

1. Show that the real numbers form a real vector space with the usual addition and multiplication. Is the set of rational numbers a subspace of this vector space?

2. Let V and W be two real vector spaces. Define $V \times W$ as the set of all ordered pairs (v, w) such that $v \in V$ and $w \in W$. Define addition and multiplication by scalars element-wise, that is, $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ and $a(v, w) = (av, aw)$. Is $V \times W$ a real vector space?

3. Determine which of the following subsets of \mathbf{R}^n are subspaces. The set of

- all vectors (x_i) such that $x_1 = 0$;
- all vectors (x_i) such that $x_1 \neq 0$;
- all vectors (x_i) such that $x_1 = 1$;

- d) all vectors (x_i) such that $x_1 + x_2 + \dots + x_n = 0$;
 e) all vectors (x_i) such that $x_1 + x_2 + \dots + x_n \neq 0$;
 f) all vectors (x_i) such that $x_1 + x_2 + \dots + x_n = 1$;
 g) all vectors (x_i) such that $a_1x_1 + a_2x_2 + \dots + a_nx_n = 0$
 with fixed real numbers a_1, \dots, a_n ;
 h) all vectors (x_i) such that $a_1x_1 + a_2x_2 + \dots + a_nx_n \neq 0$
 with fixed real numbers a_1, \dots, a_n ;
 i) all vectors (x_i) such that $a_1x_1 + a_2x_2 + \dots + a_nx_n = 1$
 with fixed real numbers a_1, \dots, a_n .

4. Let V be the set of all continuous functions on $[0, 1]$. Determine which of the following subsets are subspaces. Set of

- a) all polynomials;
 b) all functions such that $f(0) = 0$;
 c) all functions such that $f(0) \neq 0$;
 d) all functions such that $f(0) = 1$;
 e) all functions such that $\int_0^1 f(x) dx = 0$;
 f) all functions such that $\int_0^1 f(x) dx \neq 0$;
 g) all functions such that $\int_0^1 f(x) dx = 1$;
 h) all functions such that $f(0) + f(1) = 0$.

5. Do vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ span the vector space R^2 ?

6. Show that vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ cannot be expressed as the linear combination of $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$.
7. Verify that vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ form a basis in R^2 .

8. A vector space is spanned by four vectors. What can be said about the dimension of this vector space? How your answer has to be modified if the four vectors are linearly independent?

9. Assume that $\underline{x} \in R^n$ is a linear combination of vectors $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$, and for all $i = 1, 2, \dots, k$, \underline{x}_i is the linear combination of vectors $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_l$. Prove that \underline{x} is a linear combination of $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_l$.

10. Assume that $V = V(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k)$, where for $i = 1, 2, \dots, k$, \underline{x}_i is the linear combination of the linearly independent vectors $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_l$. What can you say about $\dim(V)$?

11. Assume that vectors $\underline{a}_1, \dots, \underline{a}_k$ are linearly independent. Examine the linear independence of vectors

a) $\underline{a}_1, \underline{a}_1 + \underline{a}_2, \underline{a}_3, \dots, \underline{a}_k$;

b) $\underline{a}_1, \underline{a}_1, \underline{a}_2, \underline{a}_3, \dots, \underline{a}_k$;

c) $-\underline{a}_1, \underline{a}_2, \underline{a}_3, \dots, \underline{a}_k$;

d) $\underline{a}_1, 2\underline{a}_2, 3\underline{a}_3, \dots, k\underline{a}_k$.

12. Are the following polynomials linearly independent?

a) $1, x, x^2, \dots, x^k$;

b) $1, 1+x, 1+x+x^2, 1+2x+x^2, 1+x+x^2+x^3$;

c) $1, 1+x, 1+x+x^2, 1+x+x^2+x^3, \dots, 1+x+x^2+\dots+x^k$.

13. Determine whether the row vector $(1, 2, 3, 4)$ belongs to the subspace spanned by vectors $(1, 1, 1, 1)$, $(0, 1, 1, 1)$, and $(0, 0, 1, 1)$.

14. Are the following vectors linearly independent?

a) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$;

b) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix}$;

c) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$.

15. Prove that if V_1 is a subspace of V_2 , and V_2 is a subspace of V_3 , then V_1 is a subspace of V_3 .

16. Let V be a real vector space, and $A \subseteq B \subseteq V$. Prove that $V(A) \subseteq V(B)$.

17. Let $V = \{\underline{x} \mid \underline{x} = (x_i) \in R^n, x_1 + x_2 + \dots + x_n = 0\}$. Find a basis in V , and determine the dimension of V .

18. Repeat Example 2.25 with vectors

$$\underline{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{y} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

19. Prove that if a vector \underline{u} has zero length, then $\underline{u} = \underline{0}$.

20. Prove that for any two vectors of the same size,

$$|\underline{u} - \underline{v}|^2 + |\underline{u} + \underline{v}|^2 = 2(|\underline{u}|^2 + |\underline{v}|^2).$$

21. Prove that if $|\underline{u} + \underline{v}| = |\underline{u}| + |\underline{v}|$ then \underline{u} is a scalar multiple of \underline{v} .

22. Repeat Example 2.26 with vectors

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{x}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

23. Repeat Example 2.27 with functions

$$p_0(x) = 1, p_1(x) = 1 + x, p_2(x) = 1 + x + x^2, p_3(x) = 1 + x + x^2 + x^3.$$

24. Select a maximum number of linearly independent vectors from the set below

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \right\}.$$

25. Repeat Example 2.33 with vectors

$$\underline{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \underline{x}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Chapter 3

Systems of Linear Equations and Inverses of Matrices

3.1 Introduction

Before presenting a general definition for systems of linear equations, consider the following problem. Let $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$ and \underline{b} be given m -element real (or complex) vectors. Assume that we wish to check if \underline{b} is in the subspace generated by the given vectors $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$. From Theorem 2.4 we know that $\underline{b} \in V(\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n)$ if and only if there exist scalars x_1, x_2, \dots, x_n such that

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \dots + x_n \underline{a}_n = \underline{b}. \quad (3.1)$$

To obtain an equivalent formulation for this equation, introduce the notation

$$\underline{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \underline{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \underline{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Here a_{ij} denotes the i^{th} element of vector \underline{a}_j . That is, the second subscript refers to the vector, and the first subscript indicates the position of the element. By using this notation, equation (3.1) can be rewritten as

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

and by comparing the like elements of the two sides we obtain the following set of equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots \quad \quad \quad \quad \quad \quad \quad &\quad \quad \quad \quad \quad \quad \quad \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m. \end{aligned} \tag{3.2}$$

Definition 3.1. If the coefficients a_{ij} and right-hand side scalars b_i are given real (or complex) numbers for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$, equations (3.2) are called a *system of linear equations* for the unknowns x_1, \dots, x_n .

Here n is the number of unknowns, m is the number of equations, and the system of linear equations is called an $m \times n$ system.

Introduce the $m \times n$ matrix \underline{A} with (i, j) element a_{ij} , and let \underline{x} be the n -vector with j^{th} element x_j . Then equations (3.2) can be summarized as

$$\underline{Ax} = \underline{b}. \tag{3.3}$$

This compact representation can be verified by noticing that for all $i = 1, 2, \dots, m$, the i^{th} element of the left-hand side is the product of the i^{th} row of matrix \underline{A} by the column vector \underline{x} :

$$(a_{i1}, a_{i2}, \dots, a_{in}) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n,$$

which is the left-hand side of the i^{th} equation of system (3.2). The i^{th} element of the right-hand side of equation (3.3) is b_i , which is the right-hand side of the i^{th} equation of system (3.2). Hence equations (3.1)-(3.3) are equivalent to each other. The construction of matrix \underline{A} implies that its columns are the vectors $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$, respectively.

Definition 3.2. If $\underline{b} = \underline{0}$, then equation (3.3) is called *homogeneous*, otherwise the equation is called *inhomogeneous*.

For example,

$$2x_1 + x_2 = 0$$

$$x_1 - x_2 = 0$$

is a 2×2 homogeneous system of linear equations, and the system

$$2x_1 + x_2 = 1$$

$$x_1 - x_2 = 0$$

is inhomogeneous, since there is at least one nonzero right-hand side number.

3.2 Existence and Uniqueness of a Solution

In this section we will find necessary and sufficient conditions for the existence of solutions for systems of linear equations. In the case of the

existence of a solution we will also find conditions for the uniqueness of the solution. In the cases of unique and multiple solutions practical algorithms will be introduced to find the solution or to characterize all solutions.

Our first example shows that a system of linear equations might not have any solution at all.

Example 3.1. Consider the 2×2 system

$$x_1 + x_2 = 0$$

$$x_1 + x_2 = 1,$$

which must not have solution, since $x_1 + x_2$ must not have two different values at the same time. ♦

The following example shows a case, when infinitely many solutions exist.

Example 3.2. Consider system

$$x_1 + x_2 = 0$$

$$x_1 + x_2 = 0,$$

which has infinitely many solutions: x_1 is arbitrary, and $x_2 = -x_1$. ♦

In certain cases, a unique solution might exist, as it is illustrated next.

Example 3.3. Consider now the system

$$\begin{aligned}x_1 + x_2 &= 2 \\x_1 - x_2 &= 0,\end{aligned}$$

which has a unique solution: $x_1 = x_2 = 1$, which can be obtained by adding and subtracting the two equations.

◆

We will first show that the above examples have covered all possibilities concerning the number of solutions of linear equations by verifying that in the case of multiple solutions infinitely many solutions exist. Consider the linear equations represented in the compact form (3.3), and assume that \underline{x}_1 and \underline{x}_2 ($\neq \underline{x}_1$) are two solutions. For arbitrary real (or complex) number t , consider the vector

$$\underline{x} = t\underline{x}_1 + (1-t)\underline{x}_2 = \underline{x}_2 + t(\underline{x}_1 - \underline{x}_2).$$

Since $\underline{x}_1 - \underline{x}_2 \neq \underline{0}$, different values of t give different \underline{x} vectors. Substitute \underline{x} into equation (3.3) to see that

$$\underline{A}\underline{x} = \underline{A}(t\underline{x}_1 + (1-t)\underline{x}_2) = t\underline{A}\underline{x}_1 + (1-t)\underline{A}\underline{x}_2 = t\underline{b} + (1-t)\underline{b} = \underline{b}.$$

That is, for all t , \underline{x} solves the equation. Hence infinitely many solutions exist.

From equation (3.1) and Theorem 2.4 we immediately obtain the following important result.

Theorem 3.1. A system of linear equations has at least one solution if and only if the right-hand side vector belongs to the subspace generated by the columns of the coefficient matrix.

The assertion of the theorem is illustrated in the following example.

Example 3.4. Consider the system of linear equations

$$x_1 + 2x_2 + x_3 = 2$$

$$2x_1 + 4x_2 + 2x_3 = 5.$$

The columns of the coefficient matrix are

$$\underline{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \underline{a}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \text{ and } \underline{a}_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Notice that $\underline{a}_3 = \underline{a}_1$ and $\underline{a}_2 = 2\underline{a}_1$, therefore $\{\underline{a}_1\}$ is a basis of the subspace generated by the columns. Since it is generated by only one vector, it consists of all scalar multiples of \underline{a}_1 . For any scalar c ,

$$c\underline{a}_1 = c \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} c \\ 2c \end{pmatrix},$$

where the second element is twice the first element. Since the right-hand side vector $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ does not have this property, it does not belong to the subspace generated by \underline{a}_1 . Therefore the above system of linear equations has no solution.

Using the notation $V(\underline{a}_1, \dots, \underline{a}_n)$ for the subspace generated by the columns of matrix \underline{A} , the condition of Theorem 3.1 can be formulated as

$$\underline{b} \in V(\underline{a}_1, \dots, \underline{a}_n).$$

For easier reference $V(\underline{a}_1, \dots, \underline{a}_n)$ is often called the column space of matrix \underline{A} . (Similarly, the subspace generated by the rows of a matrix \underline{A} is called the row space of \underline{A} .) Combining Theorems 3.1 and 2.7 leads to the following condition for the uniqueness of a solution.

Theorem 3.2. A system of linear equations has a unique solution if and only if the right-hand side vector belongs to the subspace generated by the columns of the coefficient matrix, and the columns are linearly independent.

The statements of the above two theorems will be restated in a more convenient way by using the following concepts:

Definition 3.3. Let \underline{A} be an $m \times n$ matrix with column vectors $\underline{a}_1, \dots, \underline{a}_n$. The rank of matrix \underline{A} is defined as the dimension of the column space $V(\underline{a}_1, \dots, \underline{a}_n)$.

Definition 3.4. The *augmented matrix* of the system of linear equations (3.3) is the $m \times (n + 1)$ matrix with columns $\underline{a}_1, \dots, \underline{a}_n, \underline{b}$.

The next result follows immediately from these definitions and Theorems 3.1 and 3.2.

Theorem 3.3. A system of linear equations has a solution if and only if the rank of its coefficient matrix equals the rank of the augmented matrix. The solution is unique if and only if in addition, the common rank is n , where n is the number of unknowns.

The statement of the theorem is illustrated next.

Example 3.5. Consider again the system of linear equations

$$\begin{aligned}x_1 + 2x_2 + x_3 &= 2 \\2x_1 + 4x_2 + 2x_3 &= 5,\end{aligned}$$

which was the subject of our previous example, where we have seen that the column space of the coefficient matrix is generated by only one vector, that is, the rank of the coefficient matrix equals 1. The columns of the augmented matrix are the vectors

$$\underline{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \underline{a}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \underline{a}_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{and} \quad \underline{b} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

We will next prove that the rank of the augmented matrix is 2, which differs from the rank of the coefficient matrix implying that no solution exists. Let

$$\underline{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

be an arbitrary vector in R^2 . We will verify that \underline{z} can be expressed as a linear combination of the columns of the augmented matrix showing that its rank equals 2. We will prove that \underline{z} is a linear combination of \underline{a}_1 and \underline{b} which means that with some scalars

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

Comparing the corresponding elements of the left-, and right-hand sides we get the equations

$$\begin{aligned}z_1 &= c_1 + 2c_2 \\z_2 &= 2c_1 + 5c_2.\end{aligned}$$

From the first equation we have

$$c_1 = z_1 - 2c_2,$$

and substitution of this relation into the second equation gives equation

$$z_2 = 2z_1 - 4c_2 + 5c_2,$$

which implies that

$$c_2 = z_2 - 2z_1,$$

and hence

$$c_1 = z_1 - 2c_2 = 5z_1 - 2z_2.$$

Hence, for arbitrary vector \underline{z} , there is a unique pair of scalars c_1 and c_2 .

◆

In the case of multiple solutions it is very important to characterize all solutions. The following theorem provides such a characterization.

Theorem 3.4. Assume that the inhomogeneous system of linear equations $\underline{A} \cdot \underline{x} = \underline{b}$ has a solution \underline{x}_0 . Then for all solutions \underline{x}^* of the corresponding homogeneous system $\underline{A} \cdot \underline{x} = \underline{0}$, $\underline{x}_0 + \underline{x}^*$ is a solution of the

inhomogeneous system, and all solutions of the inhomogeneous system can be represented in this form.

Proof. Assume first that \underline{x}_0 solves the inhomogeneous system and \underline{x}^* is a solution of the corresponding homogeneous system. Then $\underline{z} = \underline{x}_0 + \underline{x}^*$ solves the inhomogeneous system, since

$$\underline{Az} = \underline{A}(\underline{x}_0 + \underline{x}^*) = \underline{Ax}_0 + \underline{Ax}^* = \underline{b} + \underline{0} = \underline{b}.$$

Assume next that \underline{z} and \underline{x}_0 are solutions of the inhomogeneous system. Then $\underline{x}^* = \underline{z} - \underline{x}_0$ is a solution of the homogeneous system, since

$$\underline{Ax}^* = \underline{A}(\underline{z} - \underline{x}_0) = \underline{Az} - \underline{Ax}_0 = \underline{b} - \underline{b} = \underline{0}.$$

♣

Corollary. An inhomogeneous system of linear equations has a unique solution if and only if $\underline{x} = \underline{0}$ is the only solution of the corresponding homogeneous system.

The assertion of Theorem 3.3 can be reformulated by saying that the general solution of an inhomogeneous system can be obtained as the sum of a particular solution of the inhomogeneous system and the general solution of the corresponding homogeneous system. Therefore the complete characterization of all solutions of homogeneous systems is the first step before the same for inhomogeneous systems can be examined.

3.3 Systems of Homogeneous Linear Equations

In this section the set of all solutions of systems of homogeneous equations will be characterized. Our first result gives a characterization based on the concept of orthogonal complementary subspaces (which was introduced earlier in Definition 2.17).

Theorem 3.5. The set of all solutions of an $m \times n$ homogeneous system $\underline{A}\underline{x} = \underline{0}$ is a subspace of R^n , which coincides with the orthogonal complementary subspace of the row-space of matrix \underline{A} .

Proof. Let $\underline{r}_1^T, \underline{r}_2^T, \dots, \underline{r}_m^T$ denote the rows of \underline{A} , then the homogeneous system $\underline{A}\underline{x} = \underline{0}$ is equivalent to the set of equations:

$$\begin{aligned} \underline{r}_1^T \underline{x} &= 0 \\ \underline{r}_2^T \underline{x} &= 0 \\ &\dots\dots \\ \underline{r}_m^T \underline{x} &= 0. \end{aligned} \tag{3.4}$$

Assume first that a vector $\underline{z} \in R^n$ is orthogonal to all vectors of the subspace generated by the rows of matrix \underline{A} . Then it is orthogonal to all rows of \underline{A} , since these rows belong to the row-space of the matrix. Therefore, vector \underline{z} satisfies equations (3.4), and so, it is a solution of the homogeneous equation $\underline{A}\underline{x} = \underline{0}$.

Assume next that \underline{z} solves the homogeneous equation, then it satisfies equations (3.4). We will next prove that \underline{z} is orthogonal to all elements of the row-space of matrix \underline{A} . Let \underline{r}^T be an arbitrary element of the row-space, then with some scalars $c_i (1 \leq i \leq m)$,

$$\underline{r}^T = c_1 \underline{r}_1^T + c_2 \underline{r}_2^T + \dots + c_m \underline{r}_m^T,$$

and therefore,

$$\underline{r}^T \underline{z} = c_1 \underline{r}_1^T \underline{z} + c_2 \underline{r}_2^T \underline{z} + \dots + c_m \underline{r}_m^T \underline{z} = 0,$$

since each term is zero. Hence, \underline{z} is orthogonal to \underline{r} , which completes the proof.

♣

Our second result on characterizing the solutions of homogeneous systems of linear equations provides a basis for the subspace of all solutions. Before formulating the theorem, some notations are introduced. Let $\underline{a}_1, \dots, \underline{a}_n$ denote the columns of the coefficient matrix of the $m \times n$ homogeneous system of linear equations $\underline{A}\underline{x} = \underline{0}$. Assume that the columns (as well as the unknowns) are ordered so that $\{\underline{a}_1, \dots, \underline{a}_r\}$ is a basis of the column space. Then vectors $\underline{a}_{r+1}, \dots, \underline{a}_n$ are linear combinations of the basis elements. That is, for $k = r+1, \dots, n$,

$$\underline{a}_k = c_{k1} \underline{a}_1 + \dots + c_{kr} \underline{a}_r, \quad (3.5)$$

with some scalars c_{k1}, \dots, c_{kr} . For each $k = r+1, \dots, n$, introduce the vectors

$$\underline{z}_k = \begin{pmatrix} c_{k1} \\ \vdots \\ c_{kr} \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$k \rightarrow$

where the first r elements are the coefficients of the linear combination (3.5), the k^{th} element is -1 , and all other elements are equal to zero.

Theorem 3.6. Set $\{\underline{z}_{r+1}, \dots, \underline{z}_n\}$ is a basis of the subspace of all solutions of the homogeneous system of linear equations $\underline{A}\underline{x} = \underline{0}$.

Proof. First we prove that for all $k = r+1, \dots, n$, \underline{z}_k is a solution of the homogeneous system. Simple substitution and equation (3.5) imply that

$$\underline{A}\underline{z}_k = c_{k1}\underline{a}_1 + \dots + c_{kr}\underline{a}_r + (-1)\underline{a}_k = \underline{0}.$$

Next we prove that vectors $\underline{z}_{r+1}, \dots, \underline{z}_n$ are linearly independent. Consider a zero valued linear combination of these vectors:

$$b_{r+1} \begin{pmatrix} c_{r+1,1} \\ \vdots \\ c_{r+1,r} \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix} + b_{r+2} \begin{pmatrix} c_{r+2,1} \\ \vdots \\ c_{r+2,r} \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix} + \dots + b_n \begin{pmatrix} c_{n1} \\ \vdots \\ c_{nr} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Comparing the $(r+1)^{\text{th}}, (r+2)^{\text{th}}, \dots, n^{\text{th}}$ elements we see that

$$-b_{r+1} = -b_{r+2} = \dots = -b_n = 0,$$

that is, all coefficients are necessarily equal to zero, proving the linear independence of these vectors.

We will finally prove that vectors $\underline{z}_{r+1}, \dots, \underline{z}_n$ generate the entire solution set. Let $\underline{x} = (x_i)$ be a solution. The definition of vectors $\underline{z}_{r+1}, \dots, \underline{z}_n$ implies that vector

$$\underline{x}^* = \underline{x} + x_{r+1} \underline{z}_{r+1} + \dots + x_n \underline{z}_n \quad (3.6)$$

satisfies the following properties:

(a) \underline{x}^* solves the homogeneous system, since it is a linear combination of solutions;

(b) for $k = r+1, \dots, n$, the k^{th} element of \underline{x}^* equals

$$x_k + x_{r+1} \cdot 0 + \dots + x_k(-1) + \dots + x_n \cdot 0 = 0.$$

Therefore

$$\underline{x}^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_r^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with some elements x_1^*, \dots, x_r^* . Property (a) implies that

$$x_1^* \underline{a}_1 + \dots + x_r^* \underline{a}_r = \underline{0},$$

and since vectors $\underline{a}_1, \dots, \underline{a}_r$ are linearly independent, necessarily, $x_1^* = \dots = x_r^* = 0$. That is, $\underline{x}^* = \underline{0}$, and from equation (3.6) we conclude that

$$\underline{x} = -x_{r+1} \underline{z}_{r+1} - \dots - x_n \underline{z}_n.$$

Hence \underline{x} can be expressed as a linear combination of vectors $\underline{z}_{r+1}, \dots, \underline{z}_n$. Thus the proof is complete. ♣

Example 3.6. Consider the homogeneous system

$$x_1 + 2x_2 + 4x_3 = 0$$

$$x_1 - x_2 + x_3 = 0.$$

The columns of the coefficient matrix are

$$\underline{a}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \underline{a}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \text{and} \quad \underline{a}_3 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

Vectors \underline{a}_1 and \underline{a}_2 are linearly independent and $\underline{a}_3 = 2\underline{a}_1 + \underline{a}_2$, therefore $\{\underline{a}_1, \underline{a}_2\}$ is a basis for the column space of the coefficient matrix. In our case, $n = 3$ and $r = 2$. Since $\underline{a}_3 = 2\underline{a}_1 + \underline{a}_2$, from equation (3.5) we have

$$c_{31} = 2 \quad \text{and} \quad c_{32} = 1,$$

and therefore the basis of the solution space consists of only one element:

$$\underline{z}_3 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}.$$

Hence the general solution has the form:

$$\underline{x} = \begin{pmatrix} 2c \\ c \\ -c \end{pmatrix}$$

with arbitrary scalar c .



3.4 Systems of Inhomogeneous Linear Equations

In this sector the solutions of systems of inhomogeneous linear equations of the form

$$\underline{Ax} = \underline{b}$$

will be discussed, where \underline{A} is an $m \times n$ real (or complex) matrix, and \underline{b} is a real (or complex) n -vector such that $\underline{b} \neq \underline{0}$. In Theorem 3.4 we have proved that the solution set of the inhomogeneous system is given as

$$S = \{ \underline{x}_0 + \underline{x}^* \mid \underline{x}_0 \text{ is a particular solution and } \underline{x}^* \in V_1 \}$$

where V_1 is the set of all solutions of the corresponding homogeneous system. We have also shown in Theorem 3.5 that V_1 is a subspace. That is, the solutions can be obtained as the sum of a fixed vector and all elements of a subspace. This structure can be generally defined as follows.

Definition 3.5. Let V be a vector space, let V_1 be a subspace of V , and assume that a given element $\underline{x}_0 \in V$ is not in V_1 . The set

$$S = \{ \underline{x}_0 + \underline{x} \mid \underline{x} \in V_1 \} \tag{3.7}$$

is called a *linear manifold*. Subspace V_1 is called the *directing space* of S . The dimension of S is defined as the dimension of V_1 , and is denoted by $\dim(S)$.

We know from Theorem 3.5 that the set of all solutions of an inhomogeneous system of linear equation forms a linear manifold. We can

also prove that for any linear manifold $S \subseteq R^n$, there exists a certain system of linear equations such that its solution set coincides with S .

Theorem 3.7. Let S be an r -dimensional linear manifold in R^n . Then there exists an $(n-r) \times n$ system of linear equations such that the rank of the coefficient matrix is r , and its solution set coincides with S .

Proof. Assume that

$$S = \{ \underline{x}_0 + \underline{z} \mid \underline{z} \in V_1 \},$$

where $\dim(V_1) = r$. Let $\underline{a}_1, \dots, \underline{a}_{n-r}$ be a basis of the orthogonal complementary subspace of V_1 . (From Theorem 2.15 we know that the dimension of the orthogonal complementary subspace is $n-r$.) Define matrix \underline{A} as the matrix with rows $\underline{a}_1^T, \dots, \underline{a}_{n-r}^T$. Let $\underline{b} = \underline{A}\underline{x}_0$. We will next prove that the system $\underline{A}\underline{x} = \underline{b}$ satisfies the assertion, that is, its solution set coincides with S .

Assume first that $\underline{x} \in S$. Then $\underline{x} = \underline{x}_0 + \underline{z}$ with some $\underline{z} \in V_1$. Then

$$\underline{A}\underline{x} = \underline{A}\underline{x}_0 + \underline{A}\underline{z} = \underline{b},$$

since each row of \underline{A} is orthogonal to \underline{z} , that is, $\underline{A}\underline{z} = \underline{0}$. Therefore, \underline{x} is a solution of system $\underline{A}\underline{x} = \underline{b}$.

Assume next that \underline{x} is a solution of the system $\underline{A}\underline{x} = \underline{b}$. We will prove that $\underline{x} = \underline{x}_0 + \underline{z}$ with some $\underline{z} \in V_1$. Consider the vector $\underline{z} = \underline{x} - \underline{x}_0$. Then

$$\underline{A}\underline{z} = \underline{A}\underline{x} - \underline{A}\underline{x}_0 = \underline{b} - \underline{b} = \underline{0}.$$

Consequently, for all $k = 1, 2, \dots, n-r$, $a_k^T \underline{z} = 0$. Hence \underline{z} is orthogonal to all basis vectors of V_1 , consequently it is orthogonal to all elements of V_1 .

♣

Theorem 3.8. Let a linear manifold be defined as given in (3.7). Then

$$V_1 = \{ \underline{y} - \underline{z} \mid \underline{y} \text{ and } \underline{z} \in S \}. \quad (3.8)$$

Proof. Let W denote the right-hand side of relation (3.8). We will prove that $W \subseteq V_1$ and $V_1 \subseteq W$.

Assume first that $\underline{u} \in W$ is an arbitrary element. Then $\underline{u} = \underline{y} - \underline{z}$ with some \underline{y} and $\underline{z} \in S$. Therefore $\underline{y} = \underline{x}_0 + \underline{x}$ and $\underline{z} = \underline{x}_0 + \underline{x}^*$, where \underline{x} and \underline{x}^* are from V_1 . Then

$$\underline{u} = \underline{y} - \underline{z} = (\underline{x}_0 + \underline{x}) - (\underline{x}_0 + \underline{x}^*) = \underline{x} - \underline{x}^* \in V_1.$$

Assume next that $\underline{x} \in V_1$ is an arbitrary element. Then both $\underline{y} = \underline{x}_0 + 2\underline{x}$ and $\underline{z} = \underline{x}_0 + \underline{x}$ belong to S (since $\underline{x} \in V_1$ implies that $2\underline{x} \in V_1$), therefore

$$\underline{x} = (\underline{x}_0 + 2\underline{x}) - (\underline{x}_0 + \underline{x}) = \underline{y} - \underline{z} \in W.$$

In summary, $V_1 = W$, which completes the proof.

♣

Notice that the assertion of the theorem implies that the directing subspace V_1 does not depend on the selection of vector \underline{x}_0 .

Example 3.7. Assume that the vector space is R^2 , and V_1 is generated by a nonzero 2-element vector \underline{a} . Then all elements of V_1 can be expressed as $c \cdot \underline{a}$ with some scalar c . As it is illustrated in Figure 3.1, the vectors of V_1 form a straight line passing through the origin, and the direction of the line is the same as the direction of vector \underline{a} . Adding \underline{x}_0 to all vectors of this line geometrically means that the line is shifted by the vector \underline{x}_0 . Hence, the linear manifold is parallel to the one-dimensional subspace V_1 .

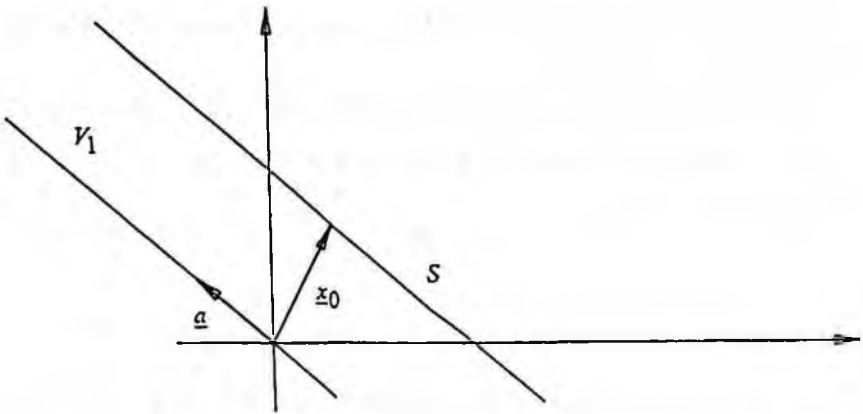


Figure 3.1. Illustration of a one-dimensional manifold.

In analyzing particular linear structures a new concept is often applied, which is a common generalization of straight lines of two-dimensional spaces and planes of three-dimensional spaces.

Definition 3.6. Let V be a finite dimensional inner-product space and let $\underline{a} \in V$ be a given nonzero element. Assume that b is a given scalar. The set

$$H = \{\underline{x} \mid \underline{x} \in V, (\underline{a}, \underline{x}) = b\} \quad (3.9)$$

is called a *hyperplane* in V .

In order to show that every hyperplane is a linear manifold, let V_1 denote the orthogonal complementary subspace of the subspace generated by vector \underline{a} . That is,

$$V_1 = \{\underline{x} \mid (\underline{a}, \underline{x}) = 0\},$$

and $\dim(V_1) = n - 1$, where n is the dimension of V . Select the vector

$$\underline{x}_0 = \frac{b}{(\underline{a}, \underline{a})} \cdot \underline{a},$$

and consider the linear manifold

$$S = \{\underline{x}_0 + \underline{x} \mid \underline{x} \in V_1\}.$$

Theorem 3.9. The hyperplane H coincides with the linear manifold S .

Proof. We will prove that $H \subseteq S$ and $S \subseteq H$. Assume first that $\underline{z} \in H$. We will show that necessarily $\underline{z} \in S$ by verifying that $\underline{x} = \underline{z} - \underline{x}_0 \in V_1$. Simple calculation shows that

$$\begin{aligned} (\underline{a}, \underline{x}) &= (\underline{a}, \underline{z} - \underline{x}_0) = (\underline{a}, \underline{z}) - (\underline{a}, \underline{x}_0) = b - \left(\underline{a}, \frac{b}{(\underline{a}, \underline{a})} \underline{a} \right) = b - \frac{b}{(\underline{a}, \underline{a})} (\underline{a}, \underline{a}) \\ &= 0, \end{aligned}$$

which implies that $\underline{x} \in V_1$, and therefore $\underline{z} \in S$.

Assume next that $\underline{z} \in S$. Then $\underline{z} - \underline{x}_0 \in V_1$, which means that $(\underline{a}, \underline{z} - \underline{x}_0) = 0$. Since

$$(\underline{a}, \underline{z} - \underline{x}_0) = (\underline{a}, \underline{z}) - \left(\underline{a}, \frac{b}{(\underline{a}, \underline{a})} \underline{a} \right) = (\underline{a}, \underline{z}) - \frac{b}{(\underline{a}, \underline{a})} (\underline{a}, \underline{a}) = (\underline{a}, \underline{z}) - b,$$

necessarily $(\underline{a}, \underline{z}) = b$ which implies that $\underline{z} \in H$.

Thus, the proof is complete. ♣

In the special case when $V = R^n$, the assertion of the theorem is obvious, since H is the solution set of the single linear equation

$$\underline{a}^T \underline{x} = b,$$

and hence, from Theorem 3.4 we know that the solution set is a linear manifold. Its dimension is the dimension of the solution set of the homogeneous equation

$$\underline{a}^T \underline{x} = 0,$$

which is $n-1$, since vector \underline{a} generates a one-dimensional subspace, the solution set is its orthogonal complementary subspace with dimension $n-1$.

3.5 Rank of Matrices

In Section 3.2 we have already introduced the concept of the rank of an $m \times n$ matrix as the dimension of the subspace generated by its columns. We have also seen that conditions for the existence of solutions of systems of linear equations could be conveniently formulated by using this concept. In this brief section the fundamental properties of matrix ranks will be introduced. These results can be summarized as follows.

Theorem 3.10. The following relations are true:

- (i) $\text{rank}(\underline{A}) = \text{rank}(\underline{A}^T)$;
- (ii) $\text{rank}(\underline{A} + \underline{B}) \leq \text{rank}(\underline{A}) + \text{rank}(\underline{B})$;
- (iii) $\text{rank}(\underline{AB}) \leq \min\{\text{rank}(\underline{A}); \text{rank}(\underline{B})\}$

where we assume that \underline{A} and \underline{B} are arbitrary matrices such that the indicated operations are defined.

Proof. (i) Assume that matrix \underline{A} is $m \times n$. Consider the homogeneous equation $\underline{Ax} = \underline{0}$. From Theorem 3.5 we know that the set of all solutions of this equation is the orthogonal complementary subspace of the subspace generated by the rows of \underline{A} . The rows of \underline{A} are the columns of \underline{A}^T , the dimension of the solution set is therefore $n - \text{rank}(\underline{A}^T)$. Theorem 3.6 implies that the dimension of the solution set is $n - \text{rank}(\underline{A})$. Hence,

$$n - \text{rank}(\underline{A}^T) = n - \text{rank}(\underline{A}),$$

from which the assertion follows.

(ii) Let $\underline{u}_1, \dots, \underline{u}_r$ and $\underline{v}_1, \dots, \underline{v}_s$ be bases of the column spaces of \underline{A} and \underline{B} , respectively. It means that all columns of \underline{A} can be expressed as a linear combination of vectors $\underline{u}_1, \dots, \underline{u}_r$:

$$\underline{a}_k = c_{k1}\underline{u}_1 + \dots + c_{kr}\underline{u}_r.$$

Similarly, all columns of \underline{B} can also be expressed as

$$\underline{b}_k = d_{k1}\underline{v}_1 + \dots + d_{ks}\underline{v}_s.$$

Since

$$\underline{a}_k + \underline{b}_k = c_{k1}\underline{u}_1 + \dots + c_{kr}\underline{u}_r + d_{k1}\underline{v}_1 + \dots + d_{ks}\underline{v}_s,$$

all columns of $\underline{A} + \underline{B}$ are in the subspace generated by the set $\{\underline{u}_1, \dots, \underline{u}_r, \underline{v}_1, \dots, \underline{v}_s\}$. Therefore the dimension of the column-space of matrix $\underline{A} + \underline{B}$ is not greater than the number $r + s$ of vectors in the generating system. Notice that $r = \text{rank}(\underline{A})$, $s = \text{rank}(\underline{B})$, from which the assertion follows.

(iii) We will first prove that $\text{rank}(\underline{AB}) \leq \text{rank}(\underline{A})$ by verifying that each column of \underline{AB} belongs to the column-space of \underline{A} . The k^{th} column of \underline{AB} equals the product of matrix \underline{A} and the k^{th} column \underline{b}_k of \underline{B} . Let $b_{1k}, b_{2k}, \dots, b_{nk}$ denote the elements of this column, then

$$\underline{Ab}_k = b_{1k}\underline{a}_1 + b_{2k}\underline{a}_2 + \dots + b_{nk}\underline{a}_n,$$

where $\underline{a}_1, \dots, \underline{a}_n$ are the columns of \underline{A} . This equation shows that \underline{Ab}_k is a linear combination of the columns of \underline{A} , consequently it belongs to the column-space of \underline{A} .

Next we verify that $\text{rank}(\underline{AB}) \leq \text{rank}(\underline{B})$. By using properties (i) and (ii) we have

$$\text{rank}(\underline{AB}) = \text{rank}((\underline{AB})^T) = \text{rank}(\underline{B}^T \underline{A}^T) \leq \text{rank}(\underline{B}^T) = \text{rank}(\underline{B}).$$

Thus the proof is complete. ♣

3.6 Matrix Equations and Inverses of Matrices

In many applications, matrix equations of the form

$$\underline{A}\underline{X} = \underline{B} \tag{3.10}$$

have to be solved, where \underline{A} is a given $m \times n$ matrix, and \underline{B} is an $m \times p$ given matrix. The solution of equation (3.10) has to be $n \times p$, otherwise either the product $\underline{A}\underline{X}$ is not defined, or the product does not have the type $m \times p$. Let $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p$ denote the unknown columns of \underline{X} , and assume that $\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p$ are the columns of \underline{B} . We may therefore use the symbol

$$\underline{X} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p) \text{ and } \underline{B} = (\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p),$$

where we wish to show the columns of these matrices. Equation (3.10) can be rewritten as

$$\underline{A}(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p) = (\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p),$$

and comparing the corresponding columns of the left-hand and right-hand side matrices we obtain the following set of equations:

$$\begin{aligned}
 \underline{Ax}_1 &= \underline{b}_1 \\
 \underline{Ax}_2 &= \underline{b}_2 \\
 &\dots \\
 \underline{Ax}_p &= \underline{b}_p.
 \end{aligned}
 \tag{3.11}$$

Hence the matrix equation (3.10) is equivalent to p systems of linear equations. Equation (3.10) has at least one solution if and only if all systems (3.11) of linear equations have solutions. The solution of equation (3.10) is unique if and only if all systems (3.11) have unique solutions. Based on Theorems 3.1, 3.2 and 3.3 we immediately obtain the following results.

Theorem 3.11. The following conditions are equivalent to each other:

- (a) Matrix equation (3.10) has at least one solution;
- (b) All vectors $\underline{b}_1, \underline{b}_2, \dots, \underline{b}_p$ belong to the subspace generated by the columns of matrix \underline{A} ;
- (c) $\text{rank}(\underline{A}) = \text{rank}(\underline{(A, B)})$, where the first n columns of matrix $\underline{(A, B)}$ are those of \underline{A} and the next p columns of $\underline{(A, B)}$ are the columns of \underline{B} .

Theorem 3.12. Equation (3.10) has a unique solution if and only if condition (b) or (c) of Theorem 3.11 holds and the columns of matrix \underline{A} are linearly independent.

Example 3.8. Consider the matrix equation

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 4 & 5 \\ 4 & 6 \end{pmatrix}.$$

Here \underline{A} is 3×3 , \underline{X} and \underline{B} are 3×2 . Equations (3.11) have now the forms:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 4 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

Simple calculation shows that the solutions are:

$$x_{11} = x_{21} = x_{31} = 1 \quad \text{and} \quad x_{12} = x_{22} = 1, x_{32} = 2.$$

Hence the solution of the matrix equation is as follows:

$$\underline{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In Section 1.4 we have already introduced the concept of the inverse of a matrix. If \underline{A} is an $n \times n$ square matrix, then the inverse \underline{A}^{-1} of \underline{A} is an $n \times n$ matrix which satisfies equations

$$\underline{A} \cdot \underline{A}^{-1} = \underline{A}^{-1} \cdot \underline{A} = \underline{I}_n,$$

where \underline{I}_n is the $n \times n$ identity matrix. If \underline{A}^{-1} exists, then matrix \underline{A} is called *invertible* or *nonsingular*, otherwise it is called *singular*. We may speak about the inverse of a matrix \underline{A} , since, as we have verified in Section 1.4, the inverse of any matrix, if exists, is necessarily unique.

Our first result gives a sufficient and necessary condition for the existence of the inverse of an $n \times n$ matrix.

Theorem 3.13. Let \underline{A} be a given $n \times n$ matrix. The following conditions are equivalent to each other:

- (a) \underline{A}^{-1} exists;
- (b) The columns of \underline{A} are linearly independent;
- (c) $\text{rank}(\underline{A}) = n$.

Proof. Using Theorem 3.11 we will prove that condition (b) is equivalent to the existence of at least one solution of matrix equation

$$\underline{A} \cdot \underline{X} = \underline{I}_n.$$

Notice first that the columns of \underline{I}_n are the natural basis vectors $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$, and therefore these vectors are in the column-space of \underline{A} if and only if the columns of \underline{A} generate the entire space R^n (or C^n). Since \underline{A} has n columns, this is the case if and only if condition (b) holds. The definition of matrix ranks implies that conditions (b) and (c) are equivalent. Notice in addition that condition (c) implies that the solution of the matrix equation $\underline{A} \cdot \underline{X} = \underline{I}_n$ is unique. In order to complete the proof we have to show that this unique solution also satisfies equation $\underline{X} \cdot \underline{A} = \underline{I}_n$. Observe first that this equation also has a unique solution, since it can be rewritten as

$$\underline{A}^T \underline{X}^T = \underline{I}_n$$

with $\text{rank}(\underline{A}^T) = \text{rank}(\underline{A}) = n$, therefore condition (c) holds. Let \underline{Y} be the unique matrix such that $\underline{Y}\underline{A} = \underline{I}_n$. Then

$$\underline{X} = \underline{I}_n \underline{X} = (\underline{Y}\underline{A})\underline{X} = \underline{Y}(\underline{A}\underline{X}) = \underline{Y}\underline{I}_n = \underline{Y},$$

that is, $\underline{X} = \underline{Y}$, which complete the proof.

♣

Corollary. Let \underline{A} be an $n \times n$ matrix. Then $\underline{X} = \underline{A}^{-1}$ if and only if $\underline{A}\underline{X} = \underline{I}_n$. Since the columns of \underline{I}_n are the natural basis vectors $\underline{e}_1, \dots, \underline{e}_n$, the columns of \underline{X} are the solutions of the following systems of linear equations:

$$\begin{aligned} \underline{A}x_1 &= \underline{e}_1 \\ \underline{A}x_2 &= \underline{e}_2 \\ &\dots \\ \underline{A}x_n &= \underline{e}_n. \end{aligned} \tag{3.12}$$

Example 3.9. Let

$$\underline{A} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix}.$$

If $\underline{x}_1, \underline{x}_2$, and \underline{x}_3 denote the columns of \underline{A}^{-1} , then they are the solutions of equations

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

and

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

where x_{ik} denotes the i^{th} element of \underline{x}_k . By substitution one may easily check that the solutions are:

$$\underline{x}_1 = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{16}{11} \\ -\frac{8}{3} \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} \frac{3}{8} \\ \frac{3}{4} \\ -\frac{1}{2} \end{pmatrix}, \underline{x}_3 = \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \end{pmatrix} = \begin{pmatrix} -\frac{5}{16} \\ \frac{1}{8} \\ \frac{1}{4} \end{pmatrix},$$

therefore

$$\underline{A}^{-1} = (\underline{x}_1, \underline{x}_2, \underline{x}_3) = \begin{pmatrix} \frac{1}{16} & \frac{3}{8} & -\frac{5}{16} \\ -\frac{11}{8} & \frac{3}{4} & -\frac{1}{8} \\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

Equations (3.12) show a strong relation between matrix inversion and the solution of systems of linear equations. If one has a solver for systems of linear equations, then the repeated application of the solver provides matrix inverses. We will next show that the reverse is also true: if one has an algorithm for inverting matrices, then the matrix inverses can be used to get the solutions of systems of linear equations with nonsingular coefficient matrices. Consider the system of linear equations

$$\underline{A}\underline{x} = \underline{b}$$

where \underline{A} is an $n \times n$ nonsingular matrix. Multiply both sides from the left by \underline{A}^{-1} to see that

$$\underline{A}^{-1}(\underline{A}\underline{x}) = \underline{A}^{-1}\underline{b}.$$

Notice that

$$\underline{A}^{-1}(\underline{A}\underline{x}) = (\underline{A}^{-1}\underline{A})\underline{x} = \underline{I}_n \underline{x} = \underline{x},$$

that is, the unique solution is obtained as $\underline{x} = \underline{A}^{-1}\underline{b}$.

Example 3.10. Consider the system of linear equations

$$2x_1 + x_2 + 3x_3 = 3$$

$$4x_1 + 4x_2 + 7x_3 = 8$$

$$2x_1 + 5x_2 + 9x_3 = 7,$$

where the coefficient matrix and right-hand side vector are as follows:

$$\underline{A} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix} \quad \text{and} \quad \underline{b} = \begin{pmatrix} 3 \\ 8 \\ 7 \end{pmatrix}.$$

In the previous example we have seen that

$$\underline{A}^{-1} = \begin{pmatrix} \frac{1}{16} & \frac{3}{8} & -\frac{5}{16} \\ -\frac{11}{8} & \frac{3}{4} & -\frac{1}{8} \\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix},$$

therefore the solution can be obtained as

$$\underline{x} = \underline{A}^{-1} \underline{b} = \begin{pmatrix} \frac{1}{16} & \frac{3}{8} & -\frac{5}{16} \\ -\frac{11}{8} & \frac{3}{4} & -\frac{1}{8} \\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 3 \\ 8 \\ 7 \end{pmatrix} = \begin{pmatrix} \frac{3}{16} & +3 & -\frac{35}{16} \\ -\frac{33}{8} & +6 & -\frac{7}{8} \\ \frac{9}{4} & -4 & +\frac{7}{4} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

That is, $x_1 = x_2 = 1$ and $x_3 = 0$.



3.7 The Elimination Method

In this section the most frequently used method for solving systems of linear equations, matrix equations, and for finding matrix inverses will be introduced. By "elimination" we refer to the important property of the method that by adding constant multiples of an equation to other equations we can make the coefficients of a given variable in other equations equal to zero. This procedure is repeated until the resulting system of equations has a diagonal coefficient matrix. Before giving the mathematical formulation of the method a simple example is presented.

Example 3.11. Consider the system of three linear equations

$$\begin{aligned}x_1 + x_2 + x_3 &= 3 \\2x_1 + 3x_2 - x_3 &= 4 \\3x_1 + 5x_2 + x_3 &= 9.\end{aligned}$$

The term x_1 of the first equation is used to eliminate x_1 from the other two equations. Multiply the first equation by 2, and subtract the resulting equation from the original second equation:

$$(2x_1 + 3x_2 - x_3) - (2x_1 + 2x_2 + 2x_3) = 4 - 6,$$

that is,

$$x_2 - 3x_3 = -2.$$

Notice that this new equation does not contain the unknown x_1 . Multiply next the first equation by 3 and subtract the resulting equation from the original third equation:

$$(3x_1 + 5x_2 + x_3) - (3x_1 + 3x_2 + 3x_3) = 9 - 9,$$

which can be simplified as

$$2x_2 - 2x_3 = 0.$$

Notice that x_1 is not contained in this new equation. By replacing the original second and third equations by these "new" equations we get the so-called *first derived system*:

$$\begin{aligned}x_1 + x_2 + x_3 &= 3 \\x_2 - 3x_3 &= -2 \\2x_2 - 2x_3 &= 0.\end{aligned}\tag{3.13}$$

Select next the first term x_2 of the new second equation to eliminate x_2 from the first and third equations. Subtract the second equation from the first one to have

$$(x_1 + x_2 + x_3) - (x_2 - 3x_3) = 3 - (-2),$$

that is,

$$x_1 + 4x_3 = 5.$$

To eliminate x_2 from the third equation, subtract the 2-multiple of the second equation from the third equation:

$$(2x_2 - 2x_3) - (2x_2 - 6x_3) = 0 - (-4),$$

which simplifies as

$$4x_3 = 4.$$

By replacing the first and third equations of the first derived system by these new versions, we have

$$\begin{aligned}x_1 + 4x_3 &= 5 \\x_2 - 3x_3 &= -2 \\4x_3 &= 4,\end{aligned}\tag{3.14}$$

which is called the *second derived system*. As the last stage of the method, we eliminate x_3 from the first and second equations. Divide the last equation by 4 then the coefficient of x_3 becomes 1. This small step makes the elimination of x_3 from the other two equations easy, since we have to multiply this new third equation

$$x_3 = 1\tag{3.15}$$

by the coefficients of x_3 of the other equations. Subtract the 4-multiple of equation (3.15) from the first equation of (3.14) to have

$$x_1 + 4x_3 - (4x_3) = 5 - 4,$$

that is,

$$x_1 = 1.$$

Multiply equation (3.15) by -3 and subtract the resulting equation from the second equation of (3.14):

$$x_2 - 3x_3 - (-3x_3) = -2 - (-3),$$

which can be simplified as

$$x_2 = 1.$$

The resulted third derived system becomes:

$$x_1 = 1$$

$$x_2 = 1$$

$$x_3 = 1,$$

which provides the solution of the original system.

Following the procedure of the above example some general observations can be made: ◆

1. If we wish to eliminate a variable from all but one equation, then a nonzero coefficient of that variable is first selected. This coefficient is called the *pivot*, and it must be nonzero. If we divide the equation of the pivot coefficient by the pivot, then this coefficient becomes 1. (This was done by dividing the third equation of (3.14) by 4.)

2. Pivot element can be selected only once from an equation, otherwise we bring back already eliminated variables into the equations. To illustrate this point assume that in system (3.13) the term x_2 of the first equation is selected as the pivot. That is, pivot is selected from the first equation for the second time. Subtract the first equation from the second one, and subtract the 2-multiple of the first equation from the third equation to have

$$(x_2 - 3x_3) - (x_1 + x_2 + x_3) = -2 - 3$$

and

$$(2x_2 - 2x_3) - (2x_1 + 2x_2 + 2x_3) = 0 - 6,$$

that is,

$$-x_1 - 4x_3 = -5$$

$$-2x_1 - 4x_3 = -6.$$

As we see, variable x_1 has "come back" into the second and third equations.

3. In the above example always diagonal elements were selected as pivots. It does not need to be in this way all the times. In case of the selection of a non-diagonal pivot element, the pivot can be placed into the diagonal by interchanging equations and variables, if necessary. To illustrate this point, assume that in the original system of the previous example the second term $5x_2$ of the last equation is selected as the first pivot. We can bring this element to the first diagonal position by interchanging the first and third equations, and then by interchanging variables x_1 and x_2 :

$$5x_2 + 3x_1 + x_3 = 9$$

$$3x_2 + 2x_1 - x_3 = 4$$

$$x_2 + x_1 + x_3 = 3.$$

4. Notice that at each step of the elimination process we subtract a constant multiple of an equation from another equation. It is done by subtracting the constant multiples of each coefficients of an equation from the corresponding coefficients of another equation. This is the same as subtracting a constant multiple of a row from another row. This observation makes the calculation faster, since there is no need to copy entire equations with all the variables. It is sufficient to copy only the coefficients and the right-hand side numbers into the elimination table and manipulate with the entire rows of the table. It is very important to indicate the variables at the top of the table, since in cases when variables are interchanged, we have to indicate these changes there.

In Table 3.1 we show the entire calculation of solving the system of the previous example. In each derived system the new pivot row is first constructed by dividing the row containing the pivot element by the pivot coefficient. The new pivot element (which is circled in the table) is therefore always equal to one. Elimination of the pivot variable from the other equations is the next step by subtracting constant multiples of the pivot row from the other rows.

Table 3.1. Elimination Process

	x_1	x_2	x_3	
	1	1	1	3
	2	3	-1	4
	3	5	1	9
Original system				
pivot row	1	1	1	3
(row 2) $-2 \times$ (pivot row)		1	-3	-2
(row 3) $-3 \times$ (pivot row)		2	-2	0
First derived system				
(row 1) $-1 \times$ (pivot row)	1		4	
pivot row		1	-3	2
(row 3) $-2 \times$ (pivot row)			4	
Second derived system				
(row 1) $-4 \times$ (pivot row)	1			
(row 2) $-(-3) \times$ (pivot row)		1		
Third derived system				
pivot row			1	

5. Each derived system is equivalent to the original system of linear equations, which can be proved as follows. Notice first that by adding back the constant multiples of the pivot row to the other rows in any derived system, the corresponding equations of the previous derived system are obtained, and by multiplying back the new pivot row by the pivot coefficient, the original pivot row of the previous derived system is obtained. Therefore the solutions of any derived system necessarily satisfy all previous derived systems, hence they satisfy the original system. The solutions of the original system also satisfy all derived systems since all derived systems are the consequences of the original system of equations. In the above example the last derived system had the identity matrix as coefficient matrix, and therefore the unique solution is given by the corresponding right hand side vector. This is not always the case, since solution might not exist and in many cases, there are infinitely many solutions. We will discuss these cases later in this section.

The elimination process illustrated in the above example can be generally formulated in the following way. Assume that the coefficient matrix $\underline{A} = (a_{ij})$ is $m \times n$, that is, there are m equations for n unknowns. For the sake of simplicity the right-hand side numbers are denoted by $a_{1,n+1}, a_{2,n+1}, \dots, a_{m,n+1}$ (that is, they form the $(n+1)^{\text{st}}$ column of \underline{A}).

Step 1. Set $k = 1$.

Step 2. Consider all elements a_{ij} for $i = k, k+1, \dots, m$ and $j = k, k+1, \dots, n$. If they are all zeros, then elimination terminates. Otherwise, select an $a_{ij} \neq 0$ as the pivot coefficient. If $i \neq k$, then interchange rows k and i . If $j \neq k$, then interchange variables j and k by interchanging columns j and k in the elimination table as well as interchanging the symbols of the j^{th} and k^{th} variables on the top of the table. The pivot element is now located in the k^{th} diagonal position.

Step 3. Divide all coefficients of the k^{th} row by a_{kk} :

$$a_{kj} \leftarrow \frac{a_{kj}}{a_{kk}} \quad (j = k+1, \dots, n+1)$$

$$a_{kk} \leftarrow 1,$$

and then eliminate the k^{th} variable from all other equations:

$$a_{ik} \leftarrow 0 \quad (i \neq k, i = 1, \dots, m),$$

$$a_{ij} \leftarrow a_{ij} - a_{ik} \cdot a_{kj} \quad (i \neq k, i = 1, \dots, m; j = k+1, \dots, n+1).$$

After this elimination step, set $k \leftarrow k+1$ and go back to Step 2.

After the process terminates, the resulted final derived system has one of the following structures:

Case 1. The coefficient matrix is the identity matrix. It occurs only if $m = n$, and the final table is as follows:

x_{i_1}	x_{i_2}	\dots	x_{i_n}	
1				$a_{1,n+1}$
	1			$a_{2,n+1}$
		\ddots		\vdots
			1	$a_{n,n+1}$

In this case a unique solution exists:

$$x_{i_1} = a_{1,n+1}, x_{i_2} = a_{2,n+1}, \dots, x_{i_n} = a_{n,n+1}.$$

By using matrix notation the table and solution can be given as

\underline{x}^T	
\underline{I}_n	\underline{a}_{n+1}

and

$$\underline{x} = \underline{a}_{n+1}. \quad (3.16)$$

We have to mention here, that the order of the elements of the solution vector \underline{x} might differ from the original order, if we had to interchange columns during the elimination process.

Case 2. All variables have been eliminated, but there are some additional equations. This case may occur only if $m > n$. The final table has the form:

x_{i_1}	x_{i_2}	\dots	x_{i_n}	
1				$a_{1,n+1}$
	1			$a_{2,n+1}$
		\ddots		\vdots
			1	$a_{n,n+1}$
				$a_{n+1,n+1}$
				\vdots
				$a_{m,n+1}$

Since all variables are eliminated, all other elements except the right hand side numbers and the indicated 1's are equal to zero. Therefore no more nonzero elements are in the table. If all of the "extra" right-hand side numbers $a_{n+1,n+1}, \dots, a_{m,n+1}$ are equal to zero, then a unique solution exists:

$$x_{i_1} = a_{1,n+1}, x_{i_2} = a_{2,n+1}, \dots, x_{i_n} = a_{n,n+1}.$$

Otherwise, there is at least one nonzero value $a_{k,n+1}$ among the numbers $a_{n+1,n+1}, \dots, a_{m,n+1}$. Then the corresponding equation

$$0 \cdot x_{i_1} + 0 \cdot x_{i_2} + \dots + 0 \cdot x_{i_n} = a_{k,n+1}$$

is a contradiction, since the right-hand side is nonzero, and the left-hand side is zero. Hence, no solution exists. Using matrix notation the final table can be written as

\underline{x}^T	
	$\underline{a}_{n+1}^{(1)}$
\underline{I}_n	$\underline{a}_{n+1}^{(2)}$

where $\underline{a}_{n+1}^{(1)}$ and $\underline{a}_{n+1}^{(2)}$ denote the vectors containing the first n and last $(m-n)$ elements of \underline{a}_{n+1} . If $\underline{a}_{n+1}^{(2)} \neq \underline{0}$, then no solution exists, and if $\underline{a}_{n+1}^{(2)} = \underline{0}$, then the unique solution is

$$\underline{x} = \underline{a}_{n+1}^{(1)}. \quad (3.17)$$

Case 3. We could not eliminate all variables, because there are no more equations to select a new pivot element from. This case may occur only if $m < n$. In this case the final table is the following:

x_{i_1}	x_{i_2}	\dots	x_{i_m}	$x_{i_{m+1}}$	\dots	x_{i_n}	
1				$a_{1,m+1}$	\dots	$a_{1,n}$	$a_{1,n+1}$
	1			$a_{2,m+1}$	\dots	$a_{2,n}$	$a_{2,n+1}$
		\ddots		\vdots		\vdots	\vdots
			1	$a_{m,m+1}$	\dots	$a_{m,n}$	$a_{m,n+1}$

In this case infinitely many solutions exist. Variables $x_{i_{m+1}}, \dots, x_{i_n}$ can have arbitrary values, therefore they are called the *free variables*. The values of

$x_{i_1}, x_{i_2}, \dots, x_{i_m}$ depend on the particular selection of the values of the free variables. Solve the first equation for x_{i_1} , the second equation for x_{i_2} , and so on, to get the *general solution*:

$$\begin{aligned}x_{i_1} &= a_{1,n+1} - \sum_{j=m+1}^n a_{1j} x_{i_j} \\x_{i_2} &= a_{2,n+1} - \sum_{j=m+1}^n a_{2j} x_{i_j} \\&\dots \\x_{i_m} &= a_{m,n+1} - \sum_{j=m+1}^n a_{mj} x_{i_j}.\end{aligned}$$

Using matrix notation, the final table can be given as

$\underline{x}^{(1)T}$	$\underline{x}^{(2)T}$	
\underline{I}_m	$\underline{A}^{(2)}$	\underline{a}_{n+1}

and the general solution has the form:

$$\underline{x}^{(1)} = \underline{a}_{n+1} - \underline{A}^{(2)} \underline{x}^{(2)}, \quad (3.18)$$

where vector $\underline{x}^{(2)}$ is arbitrary.

Case 4. We could not eliminate all variables, because all coefficients equal zero in the additional equations. The final table is now the following:

x_{i_1}	x_{i_2}	\dots	x_{i_k}	$x_{i_{k+1}}$	\dots	x_{i_n}	
1				$a_{1,k+1}$	\dots	$a_{1,n}$	$a_{1,n+1}$
	1			$a_{2,k+1}$	\dots	$a_{2,n}$	$a_{2,n+1}$
		\ddots		\vdots		\vdots	\vdots
			1	$a_{k,k+1}$	\dots	$a_{k,n}$	$a_{k,n+1}$
							$a_{k+1,n+1}$
							\vdots
							$a_{m,n+1}$

Notice that this table is a combination of Cases 2 and 3. If at last one of the "extra" right-hand side numbers $a_{k+1,n+1}, \dots, a_{m,n+1}$ has nonzero value, then no solution exists, since the corresponding equation is a contradiction. If all of these numbers equal zero, then simply ignore the additional equations and the resulting table coincides with the final table of Case 3 with $m = k$. Therefore infinitely many solutions exist, and the general solution is obtained as it was shown in Case 3. In matrix notation this case results in a table of the form:

$\underline{x}^{(1)T}$	$\underline{x}^{(2)T}$	
		$\underline{a}_{n+1}^{(1)}$
		$\underline{a}_{n+1}^{(2)}$
\underline{I}_k	$\underline{A}^{(2)}$	

If $\underline{a}_{n+1}^{(2)} \neq \underline{0}$, then no solution exists. Otherwise, there are infinitely many solutions, and the general solution has the form:

$$\underline{x}^{(1)} = \underline{a}_{n+1}^{(1)} - \underline{A}^{(2)} \underline{x}^{(2)}, \quad (3.19)$$

where vector $\underline{x}^{(2)}$ is arbitrary.

The following examples illustrate the above four cases and the ways how the solutions are obtained.

Example 3.12. We will first solve the system

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ 2x_1 + x_2 - x_3 &= -1 \\ x_1 + 2x_2 + x_3 &= 4. \end{aligned}$$

The elimination procedure is given in Table 3.2.

Table 3.2. Elimination of Example 3.12.

	x_1	x_2	x_3	
	1	-1	1	1
	2	1	-1	-1
	1	2	1	4
	1	-1	1	1
Pivot row		3	-3	-3
(row 2) $-2 \times$ (pivot row)		3	0	3
(row 3) $-1 \times$ (pivot row)	1		0	0
(row 1) $-(-1) \times$ (pivot row)		1	-1	-1
Pivot row			3	6
(row 3) $-3 \times$ (pivot row)	1			0
row 1		1		1
(row 2) $+1 \times$ (pivot row)			1	2
Pivot row				

original system

first derived system

second derived system

third derived system

Notice that in the first equation of the second derived system the coefficient of x_3 is zero. We simply copied this equation into the third derived system, since zero coefficient does not need to be eliminated. The pivot elements (after dividing the pivot terms by the pivot coefficients) are circled in all derived systems. The last system gives the unique solution:

$$x_1 = 0, x_2 = 1, x_3 = 2.$$

Example 3.13. Consider next the system ◆

$$\begin{aligned}x_1 - x_2 + x_3 &= 1 \\2x_1 + x_2 - x_3 &= -1 \\x_1 + 2x_2 + x_3 &= 4 \\3x_1 + 2x_2 - x_3 &= 0 \\x_1 + x_2 - 2x_3 &= -5.\end{aligned}$$

The elimination process is shown in Table 3.3. The last two equations have all zero coefficients, and in the case of the last equation the right-hand side is nonzero. Therefore no solution exists. Modify the original last equation as

$$x_1 + x_2 - 2x_3 = -3,$$

and repeat the elimination process. The calculations are also shown in Table 3.3, where the new right-hand side numbers of the last equations are given next to their original values.

Table 3.3. Elimination of Example 3.13.

	x_1	x_2	x_3	
	1	-1	1	1
	2	1	-1	-1
	1	2	1	4
	3	2	-1	0
	1	1	-2	-5
				-3
pivot row	1	-1	1	1
(row 2)-2×(pivot row)		3	-3	-3
(row 3)-1×(pivot row)		3	0	3
(row 4)-3×(pivot row)		5	-4	-3
(row 5)-1×(pivot row)		2	-3	-6
				-4
(row 1)+1×(pivot row)	1	0	0	0
pivot row		1	-1	-1
(row 3)-3×(pivot row)			3	6
(row 4)-5×(pivot row)			1	2
(row 5)-2×(pivot row)			-1	-4
				-2
row 1	1			0
(row 2)+1×(pivot row)		1		1
pivot row			1	2
(row 4)-1×(pivot row)				0
(row 5)-(-1)×(pivot row)				-2
				0

Notice that only the last right-hand side numbers become different, since we did not select pivot from the last equation. The right-hand side numbers of the last two equations become zero showing that a unique solution exists:

$$x_1 = 0, x_2 = 1, x_3 = 2.$$

Example 3.14. We solve now the system

$$\begin{aligned}x_1 + x_2 + x_3 + x_4 &= 2 \\x_1 + x_2 + 2x_3 + 2x_4 &= 2 \\x_1 + x_2 + 2x_3 + 4x_4 &= 2,\end{aligned}$$

which has only three equations for four unknowns. The elimination process is shown in Table 3.4.

Table 3.4. Elimination of Example 3.14.

	x_1	x_3	x_4	x_2		
		x_2	x_2	x_4		
			x_3			
	1	1	1	1	2	
	1	1	2	2	2	original system
	1	1	2	4	2	
pivot row	1	1	1	1	2	
(row 2)–1×(pivot row)		0	1	1	0	first derived system
(row 3)–1×(pivot row)		0	1	3	0	
	1	1	1	1	2	
		1	0	1	0	columns 2 and 3 are interchanged in the first derived system
		1	0	3	0	
(row 1)–1×(pivot row)	1		1	0	2	
pivot row		1	0	1	0	second derived system
(row 3)–1×(pivot row)			0	2	0	
	1		0	1	2	
		1	1	0	0	columns 3 and 4 are interchanged in the second derived system
			2	0	0	
row 1	1			1	2	
(row 2)–1×(pivot row)		1		0	0	third derived system
pivot row			1	0	0	

Since in the first derived system, the coefficient of x_2 is zero in the second and third equations, no pivot element could be selected from the second column. We had therefore selected the x_3 term from the second equation as the pivot term. We had to interchange the second and third columns in order to bring the pivot element into the diagonal position. In the same time we had to interchange x_2 and x_3 on the top of the table. Since the third coefficient in the third equation of the second derived system is also zero, the last pivot element had to be selected as the fourth coefficient of this equation. We had to interchange the third and fourth columns in order to bring the pivot element into the diagonal position again. In addition, we had to interchange x_2 and x_4 on the top of the table. From the third derived system we see that x_2 is the only free variable, and

$$x_1 = 2 - x_2, x_3 = 0, x_4 = 0$$

is the general solution. ♦

Example 3.15. Consider next the system

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 2 \\ 2x_1 + x_2 + 2x_3 + 2x_4 &= 3 \\ 3x_1 + 2x_2 + 3x_3 + 3x_4 &= 5 \\ x_1 &+ x_3 + x_4 = 1. \end{aligned}$$

Table 3.5 shows the elimination process. From the last table we see that infinitely many solutions exist.

Table 3.5. Elimination of Example 3.15.

	x_1	x_2	x_3	x_4		
	1	1	1	1	2	original system
	2	1	2	2	3	
	3	2	3	3	5	
	1	0	1	1	1	
pivot row	1	1	1	1	2	first derived system
(row 2) $-2 \times$ (pivot row)		-1	0	0	-1	
(row 3) $-3 \times$ (pivot row)		-1	0	0	-1	
(row 4) $-1 \times$ (pivot row)		-1	0	0	-1	
(row 1) $-1 \times$ (pivot row)	1		1	1	1	should be deleted as meaningless equations
pivot row		1	0	0	1	
(row 3) $-(-1) \times$ (pivot row)						
(row 4) $-(-1) \times$ (pivot row)						

The free variables are x_3 and x_4 , and the general solution is:

$$x_1 = 1 - x_3 - x_4, \quad x_2 = 1.$$

The solution of matrix equations and inverting matrices are based on the repeated application of the elimination method. Since the coefficient matrix is common in the systems to be solved, the elimination process has to be performed only once, but several right-hand side vectors should be considered simultaneously. This idea is illustrated in the next two examples.

Example 3.16. We will first solve the matrix equation

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \cdot \underline{X} = \begin{pmatrix} 3 & 4 \\ 4 & 5 \\ 4 & 6 \end{pmatrix},$$

which was the subject earlier in Example 3.8. In the elimination table the entire right hand side matrix has to replace the right hand side vector. The process is shown in Table 3.6.

Table 3.6. Elimination of Example 3.16.

	x_1	x_2	x_3			
	1	1	1	3	4	original system
	1	2	1	4	5	
	1	1	2	4	6	
pivot row	1	1	1	3	4	first derived system
(row 2)−1×(pivot row)		1	0	1	1	
(row 3)−1×(pivot row)		0	1	1	2	
(row 1)−1×(pivot row)	1		1	2	3	second derived system
pivot row		1	0	1	1	
row 3			1	1	2	
(row 1)−1×(pivot row)	1			1	1	third derived system
row 2		1		1	1	
pivot row			1	1	2	

The "right-hand side matrix" gives the solution:

$$\underline{X} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Example 3.17. We will next find the inverse of matrix

$$\underline{A} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix}.$$

The elimination procedure is given in Table 3.7. The identity matrix is placed as the right-hand matrix. From the last derived system we see that

$$\underline{A}^{-1} = \begin{pmatrix} \frac{1}{16} & \frac{3}{8} & -\frac{5}{16} \\ \frac{11}{8} & \frac{3}{4} & -\frac{1}{8} \\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

Table 3.7. Elimination of Example 3.17.

	x_1	x_2	x_3				
	2	1	3	1	0	0	original system
	4	4	7	0	1	0	
	2	5	9	0	0	1	
pivot row	1	1/2	3/2	1/2	0	0	First Derived System
(row 2)–4×(pivot row)		2	1	–2	1	0	
(row 3)–2×(pivot row)		4	6	–1	0	1	Second derived system
(row 1)–1/2×(pivot row)	1		5/4	1	–1/4	0	
pivot row			1/2	–1	1/2	0	
(row 3)–4×(pivot row)			4	3	–2	1	Third derived system
(row 1)–5/4×(pivot row)	1			1/16	3/8	–5/16	
(row 2)–1/2×(pivot row)		1		–11/8	3/4	–1/8	
pivot row			1	3/4	–1/2	1/4	

We conclude this section with the example of a matrix which has no inverse.

Example 3.18. Consider matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 4 & 4 \end{pmatrix}.$$

The elimination process is shown in Table 3.8. The process terminates at the second derived system, since the last coefficient of the third equation is zero, so no more pivot element can be selected. The right hand side numbers have nonzero values in the last equation, which implies that no solution exists. We have seen earlier that a matrix inverse is unique if it exists. Therefore no such case may occur during matrix inversion that would indicate the existence of multiple solutions.

Table 3.8. Elimination of Example 3.18.

	x_1	x_2	x_3				
	1	1	1	1	0	0	original system
	2	3	3	0	1	0	
	3	4	4	0	0	1	
pivot row	1	1	1	1	0	0	first derived system
(row 2)–2×(pivot row)		1	1	–2	1	0	
(row 3)–3×(pivot row)		1	1	–3	0	1	
(row 1)–1×(pivot row)	1		0	3	–1	0	second derived system
pivot row		1	1	–2	1	0	
(row 3)–1×(pivot row)			0	–1	–1	1	

In the discussions on elimination as well as in the above examples we always assumed that the pivot element was selected from the diagonal of the coefficient matrix, otherwise by interchanging rows and columns (which is equivalent to interchanging equations and unknowns) the pivot element was "moved" into a diagonal position. This interchange has certain advantages in simplifying the matrix notation of the solution, however in practical cases it is not necessary. If a non-diagonal pivot element is selected, then the coefficients of the same variable have to be eliminated from all other equations. When the elimination process terminates, the final coefficient matrix will not necessarily contain the identity matrix, but it will always contain a matrix that can be obtained from the identity matrix by column interchanges. The solution of the system can then be obtained in the same way as shown before with the minor difference that the order in which the unknowns are determined might differ from the natural order. This modified procedure is illustrated in the following example.

Example 3.19. We will now solve equations

$$2x_1 + x_2 - x_3 = 3$$

$$x_1 + x_2 + x_3 = 2$$

$$x_1 - x_2 + 2x_3 = 0.$$

The elimination process is shown in Table 3.9.

Table 3.9. Elimination of Example 3.19.

	x_1	x_2	x_3	
	2	1	-1	3
	1	1	1	2
	1	-1	2	0
Original system				
(row 1) - 2 × (pivot row)	0	-1	-3	-1
pivot row	1	1	1	2
First derived system				
(row 3) - (pivot row)	0	-2	1	-2
pivot row	0	1	3	1
Second derived system				
(row 2) - (pivot row)	1	0	-2	1
(row 3) - (-2) × (pivot row)	0	0	7	0
Third derived system				
(row 1) - 3 × (pivot row)	0	1	0	1
(row 2) - (-2) × (pivot row)	1	0	0	1
pivot row	0	0	1	0

The first equation shows that $x_2 = 1$, the second equation implies that $x_1 = 1$, and from the third equation we obtain $x_3 = 0$.



3.8 Applications

In this section some particular problems will be introduced and solved by using the methodology of this chapter.

1. The first model is known as *Leontief's input-output system*. Consider an economy where n kinds of goods are produced, traded, and consumed. The goods are labeled by $i = 1, 2, \dots, n$. It is assumed that each sector produces a single kind of good, so no joint production prevails. Distinct sectors produce distinct kinds of goods, therefore there is a one-to-

one correspondence between goods and the sectors. The sector producing the i^{th} good is also denoted by i . In each sector, production transforms several (possibly all) kinds of goods in some quantities into a single kind of good in some amount. It is also assumed that to produce one unit of the j^{th} good, a_{ij} units of the i^{th} good ($i = 1, 2, \dots, n$) are needed as inputs for sector j . The a_{ij} quantities are called the *input coefficients*. Let x_j denote the output of the i^{th} good per unit of time, say, per annum. This x_j is the gross output, and part of it is consumed as input needed for production activities. The net output y_i of good i can be obtained as the difference of the gross output and the total amount of the good consumed as input:

$$y_i = x_i - \sum_{j=1}^n a_{ij} x_j \quad (i = 1, 2, \dots, n). \quad (3.20)$$

Introduce the following notation:

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Vectors \underline{x} and \underline{y} are called the gross and net output vectors, respectively, and matrix \underline{A} is called the input coefficient matrix. Equations (3.20) can be written in an equivalent form:

$$\underline{x} - \underline{A}\underline{x} = \underline{y},$$

that is,

$$(\underline{I}_n - \underline{A})\underline{x} = \underline{y}. \quad (3.21)$$

If the gross production vector is known, then the corresponding net production vector is obtained by multiplying matrix $\underline{I}_n - \underline{A}$ by vector \underline{x} . It is an important question in economic planning to determine the gross production vector for given net productions. That is, vector \underline{x} has to be determined for a given matrix \underline{A} and a net production vector \underline{y} . Notice that (3.21) is a system of linear equations with coefficient matrix $\underline{I}_n - \underline{A}$, right-hand side vector \underline{y} , and unknown vector \underline{x} . Hence, the elimination method introduced in the previous section can be directly used to find vector \underline{x} .

Example 3.20. As an illustration, consider a three-sector economy with input coefficient matrix

$$\underline{A} = \begin{pmatrix} \frac{1}{10} & \frac{2}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{1}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{2}{10} & \frac{1}{10} \end{pmatrix}.$$

Assume that an economic plan specifies the net output vector

$$\underline{y} = \begin{pmatrix} 310 \\ 200 \\ 90 \end{pmatrix}.$$

Then the gross output vector is the solution of equation (3.21), which now has the form

$$\begin{aligned}\frac{9}{10}x_1 - \frac{2}{10}x_2 - \frac{2}{10}x_3 &= 310 \\ -\frac{2}{10}x_1 + \frac{9}{10}x_2 - \frac{2}{10}x_3 &= 200 \\ -\frac{2}{10}x_1 - \frac{2}{10}x_2 + \frac{9}{10}x_3 &= 90.\end{aligned}$$

The elimination process can be applied, and the solution

$$x_1 = 500, \quad x_2 = 400, \quad \text{and} \quad x_3 = 300$$

can be found. The details of elimination are left as an exercise.

We mention here that a comprehensive analysis of linear input-output models is given in Nikaido (1968). ♦

2. The second model to be discussed here is known as *polynomial interpolation*. Assume that x_1, x_2, \dots, x_n are given different real numbers, and f_1, f_2, \dots, f_n are arbitrary real values. The f_i values are not necessarily different. The x_i values can be interpreted as the values of the independent variable where a certain unknown real function is measured, and for $i = 1, 2, \dots, n$, f_i is the measured function value at x_i . The unknown function is approximated by a polynomial of least degree that is consistent with the measurements. That is, a polynomial p of degree as small as possible is determined such that for $i = 1, 2, \dots, n$, $p(x_i) = f_i$. We have therefore n equations. In order to have the same number of unknowns as equations, a polynomial of degree $n - 1$ is selected:

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}.$$

The unknown coefficients must satisfy the equations:

$$a_0 + a_1x_i + a_2x_i^2 + \dots + a_{n-1}x_i^{n-1} = f_i \quad (i = 1, 2, \dots, n),$$

which is a system of n linear equations for n unknowns. Using matrix and vector notation, these equations can be summarized as follows:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}. \quad (3.22)$$

It can be proved that this system has a unique solution, if the x_i values are different (that is, when no repeated measurement is used). For details, see for example, Szidarovszky and Yakowitz (1978). The resulting polynomial is known as the *interpolation polynomial*.

Example 3.21. Consider four measurements:

$$x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$$

and

$$f_1 = 0, f_2 = 2, f_3 = 10, f_4 = 30.$$

The unique cubic interpolation polynomial can be determined by solving system (3.22), which now has the form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 10 \\ 30 \end{pmatrix}.$$

Easy calculation (which is left as an exercise) shows that the solution is the following:

$$a_0 = 0, \quad a_1 = 1, \quad a_2 = 0, \quad \text{and} \quad a_3 = 1.$$

Thus, the interpolation polynomial is the following:

$$p(x) = x + x^3.$$

We mention here that alternative methods to find interpolation polynomials can be found in the numerical analysis literature. See, for example, Szidarovszky and Yakowitz (1978). ◆

3. From the previous case study we know that by increasing the number of measurement points the degree of the interpolation polynomial also increases. In many applications, instead of finding the high degree interpolation polynomial with exact fit, a polynomial of lower degree is determined which is not required to fit the data points exactly, but which fits the data as well as possible. Depending on the selection of the measure of the goodness of fit, several concepts have been developed. The most popular approach is the *least squares method*, which can be described as follows. Assume again that x_1, x_2, \dots, x_n are different real numbers, and f_1, f_2, \dots, f_n are arbitrary real values. A polynomial

$$p(x) = a_0 + a_1x + \dots + a_mx^m \quad (3.23)$$

of fixed degree $m \leq n-1$ has to be determined which minimizes the sum of the squared residuals:

$$Q = \sum_{i=1}^n (f_i - p(x_i))^2 = \sum_{i=1}^n \left(f_i - \sum_{j=0}^m a_j x_i^j \right)^2.$$

Since for all i , x_i and f_i are known, this quantity depends only on the $m+1$ unknown coefficients a_0, a_1, \dots, a_m . It can be proved (see, for example, Ross, 1987) that this function has a unique minimizer, and it can be obtained by solving the system

$$(\underline{X}^T \underline{X}) \underline{a} = \underline{X}^T \underline{f} \quad (3.24)$$

where

$$\underline{X} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{pmatrix}, \quad \underline{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}, \quad \underline{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}.$$

The solution of this system provides the values of the unknown coefficients a_0, a_1, \dots, a_m of the least squares polynomial (3.23).

Example 3.22. As the illustration of the application of system (3.24), we will consider again the data of the previous example, and we will

find the quadratic least squares polynomial based on the given data set. Since $m = 2$, in this case,

$$\underline{X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}, \quad \underline{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}, \quad \underline{f} = \begin{pmatrix} 0 \\ 2 \\ 10 \\ 30 \end{pmatrix}.$$

Simple calculation shows that

$$\underline{X}^T \underline{X} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix},$$

$$\underline{X}^T \underline{f} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 10 \\ 30 \end{pmatrix} = \begin{pmatrix} 42 \\ 112 \\ 312 \end{pmatrix},$$

therefore equation (3.24) can be written as

$$\begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 42 \\ 112 \\ 312 \end{pmatrix}.$$

The application of the elimination process (the details are left as an exercise) shows that

$$a_0 = \frac{3}{10}, \quad a_1 = -\frac{37}{10}, \quad \text{and} \quad a_2 = \frac{45}{10}.$$

Hence, the least squares polynomial is given as

$$p(x) = \frac{3}{10} - \frac{37}{10}x + \frac{9}{2}x^2.$$

Finally we mention that further details on the least squares method (including alternative procedures) can be found in the numerical analysis as well as in the statistical literature. See for example, Szidarovszky and Yakowitz (1978) or Ross (1987). ♦

4. In application 7 of Section 1.5 we have seen that the material cost of the three kinds of biscuits can be obtained from the ingredient prices as

$$\begin{aligned} c_1 &= 0.5p_1 + 0.2p_2 + 0.3p_3 \\ c_2 &= 0.55p_1 + 0.25p_2 + 0.2p_3 \\ c_3 &= 0.6p_1 + 0.15p_2 + 0.25p_3. \end{aligned}$$

If the material costs are given, then the prices p_1, p_2, p_3 can be determined by solving the above system of linear equations. As an example assume that

$$c_1 = 2.3, \quad c_2 = 2.2, \quad \text{and} \quad c_3 = 2.25,$$

then the application of the elimination method shows that

$$p_1 = 2, \quad p_2 = 2, \quad \text{and} \quad p_3 = 3.$$

5. Assume now that at any time $t = 0, 1, 2, \dots$, a system is in one of the states S_1, S_2, \dots, S_n . For example, the weather in a city could be in one of the following possible states: sunny and cold, sunny and warm, cloudy and cold, cloudy and warm, raining and cold, raining and warm. It is also assumed that the transition from one state to another is not deterministic, it can be specified in terms of probabilities that depend on the previous history of the system. In the special case, when the probability of each state depends only on the immediate history of the system, then the process is called a Markov chain. The mathematical model is based on the transition matrix P which is defined in the following way. The matrix is $n \times n$, and its (i, j) -element P_{ij} gives the probability that if the system is in state j at any time period t , then it will be in state i at the next time period $t+1$. Clearly,

$$P_{1j} + P_{2j} + \dots + P_{nj} = 1,$$

that is, the sum of the elements of each column is one, and all matrix elements are nonnegative. At each time period t , let $x_1(t), x_2(t), \dots, x_n(t)$ denote the probability of each state. That is

$$P(\text{state at time period } t = S_k) = x_k(t) \quad (1 \leq k \leq n).$$

Clearly,

$$x_1(t) + x_2(t) + \dots + x_n(t) = 1$$

and each number $x_k(t)$ is nonnegative. It is easy to see that for all $t \geq 0$,

$$\underline{x}(t+1) = \underline{P}\underline{x}(t), \quad (3.25)$$

that is, the probabilities of different states can be recursively computed.

Example 3.23. Assume that

$$\underline{P} = \begin{pmatrix} 0.1 & 0.5 & 0.9 \\ 0.4 & 0.2 & 0 \\ 0.5 & 0.3 & 0.1 \end{pmatrix},$$

and at a certain time period $\underline{x}(t) = (0.1, 0.5, 0.4)^T$.

Then

$$\underline{x}(t+1) = \begin{pmatrix} 0.1 & 0.5 & 0.9 \\ 0.4 & 0.2 & 0 \\ 0.5 & 0.3 & 0.1 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0.5 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 0.62 \\ 0.14 \\ 0.24 \end{pmatrix},$$

that is, at time period $t+1$, the system will be in S_1 with probability 0.62, in S_2 with probability 0.14, and in state S_3 with probability 0.24. Predicting the probabilities of different state for the next time period is a simple task: a matrix-vector multiplication has to be performed. We can also raise the question of determining the probabilities of the different state for previous time periods. For example, we may compute $\underline{x}(t)$ from $\underline{x}(t+1)$ by solving a system of linear equations with coefficient matrix \underline{P} and right hand side vector $\underline{x}(t+1)$. Assume now that $\underline{x}(4) = (0.62, 0.14, 0.24)^T$, then $\underline{x}(3)$ is the solution of linear equations

$$\begin{pmatrix} 0.1 & 0.5 & 0.9 \\ 0.4 & 0.2 & 0 \\ 0.5 & 0.3 & 0.1 \end{pmatrix} \begin{pmatrix} x_1(3) \\ x_2(3) \\ x_3(3) \end{pmatrix} = \begin{pmatrix} 0.62 \\ 0.14 \\ 0.24 \end{pmatrix},$$

which has the usual form:

$$0.1 \cdot x_1(3) + 0.5 \cdot x_2(3) + 0.9 \cdot x_3(3) = 0.62$$

$$0.4 \cdot x_1(3) + 0.2 \cdot x_2(3) = 0.14$$

$$0.5 \cdot x_1(3) + 0.3 \cdot x_2(3) + 0.1 \cdot x_3(3) = 0.24.$$

Simple calculation shows that the solution is

$$x_1(3) = 0.1, x_2(3) = 0.5, \quad \text{and} \quad x_3(3) = 0.4.$$

The repeated application of equation (3.25) shows that for all $t \geq 0$ and $n \geq 1$,

$$\underline{x}(t+n) = \underline{P}^n \underline{x}(t) \tag{3.26}$$

giving a simple way to predict future state probabilities. In addition, by solving equation (3.26) for $\underline{x}(t)$, earlier state probabilities can be obtained.

Notice also, that an efficient method to compute matrix powers \underline{P}^n for large values of n is therefore very useful in examining Markov chains.

The steady state probability vector of a Markov chain is a nonnegative vector \underline{x}^* such that the sum of its elements is one, furthermore

$$\underline{x}^* = \underline{P} \underline{x}^* .$$

Hence \underline{x}^* can be obtained by solving the linear equations

$$\begin{aligned}
 (\underline{I} - \underline{P})\underline{x}^* &= \underline{0} \\
 x_1^* + x_2^* + \dots + x_n^* &= 0.
 \end{aligned}
 \tag{3.27}$$

It can be proved that the rows of $\underline{I} - \underline{P}$ are linearly dependent, so we have at most n linearly independent equations for the n unknowns. Since the linear equations are homogeneous, there is at least one nonzero solution. In section 7.7 we will also see that all components of \underline{x}^* can be selected nonnegative.

3.9 Exercises

1. Check if vector \underline{x} can be expressed as a linear combination of $\underline{x}_1, \underline{x}_2$, and \underline{x}_3 if

$$\text{a) } \underline{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \underline{x}_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \underline{x} = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix};$$

$$\text{b) } \underline{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \underline{x}_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \underline{x} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}.$$

2. Find the solutions, if they exist, of the following systems of equations:

a) $2x_1 - x_2 + x_3 = 3$

$x_1 + x_2 + x_3 = 2$

$3x_1 + x_2 - x_3 = 2$

b) $x_1 + x_2 + x_3 + x_4 = 2$

$x_1 - x_2 + x_3 + 2x_4 = 0$

$2x_1 + x_2 + x_3 - x_4 = 3$

c) $x_1 + x_2 = 2$

$2x_1 - x_2 = 1$

$2x_1 + x_2 = 3$

d) $x_1 + x_2 + x_3 + x_4 = 2$

$x_1 - x_2 + x_3 + 2x_4 = 0.$

3. Rewrite the systems of the previous problem in matrix form $\underline{Ax} = \underline{b}$.

4. Determine which of the following systems of linear equations are solvable:

a) $x_1 + 2x_2 + x_3 + x_4 = 3$

$-x_1 + x_2 - x_3 - x_4 = 0$

b) $x_1 + x_2 + x_4 = 2$

$x_1 + x_3 + x_4 = 2$

$x_2 + x_3 + x_4 = 1$

$x_1 + x_2 + x_3 + x_4 = 2$

c) $x_1 + 2x_2 + x_3 = 4$

$x_1 + x_2 + 2x_3 = 4$

$2x_1 + x_2 + x_3 = 4$

$x_1 + x_2 + x_3 = 2$

d) $x_1 + x_2 + x_3 + x_4 = 1$

$x_1 - x_2 + x_3 - x_4 = 1$

$x_1 + 2x_2 + x_3 + 2x_4 = 2.$

5. Check the solvability of the equations of Problem 2 by using Theorem 3.3.
6. Check the solvability of the equations of Problem 4 by using Theorem 3.3.
7. Find a basis of the solution space of the following homogeneous systems:

a) $x_1 + x_2 + x_3 = 0$

$2x_1 + x_2 + 2x_3 = 0$

b) $x_1 + x_2 - x_3 - x_4 = 0$

$2x_1 + x_2 + x_3 - 2x_4 = 0$

$3x_1 + x_2 - x_3 - 3x_4 = 0$

c) $x_1 + 2x_2 + x_3 + x_4 = 0$

$-x_1 + x_2 - x_3 - x_4 = 0.$

8. What is the dimension of the solution space of the homogeneous system $\underline{A}x = \underline{0}$, where \underline{A} is a $2 \times n$ matrix?
9. Find the ranks of the coefficient matrices of the systems of Problems 2 and 4.
10. Let \underline{A} be an $m \times n$ real matrix, and assume the \underline{U} has m columns and \underline{V} has n rows. Are the following statements true?
- a) If the rows of \underline{U} are linearly independent, then
$$\text{rank}(\underline{A}) = \text{rank}(\underline{UA});$$
- b) If the columns of \underline{V} are linearly independent, then
$$\text{rank}(\underline{A}) = \text{rank}(\underline{AV}).$$
11. Find the ranks of the coefficient matrices of the systems of Problem 7.
12. Solve the following systems of equations by using the elimination method
- a) $x_1 + x_2 + x_3 = 3$
 $x_1 - x_2 - x_3 = 3$
 $2x_1 + x_2 + 7x_3 = 6$
- b) $x_1 + 2x_2 + 3x_3 = 4$
 $2x_1 + 6x_2 + 10x_3 = 14$
 $-x_1 - 3x_2 - 6x_3 = 9$
- c) $x_1 + 2x_2 + 3x_3 = 4$
 $2x_1 + 4x_2 + 9x_3 = 4$
 $4x_1 + 3x_2 + 2x_3 = 1.$

13. Solve the systems of the previous problem by the formula $\underline{x} = \underline{A}^{-1} \underline{b}$, where \underline{A} is the coefficient matrix and \underline{b} is the right-hand side vector.

14. Illustrate Theorem 3.2 with the systems of Problem 12.

15. Illustrate Theorem 3.13 with the coefficient matrices of Problem 12.

16. Solve the following matrix equations:

$$\text{a) } \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 2 & 1 & 7 \end{pmatrix} \underline{X} = \begin{pmatrix} 3 & 0 \\ 3 & 0 \\ 6 & 0 \end{pmatrix}$$

$$\text{b) } \begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ -1 & -3 & -6 \end{pmatrix} \underline{X} = \begin{pmatrix} 4 & 8 \\ 14 & 28 \\ 9 & 18 \end{pmatrix}$$

$$\text{c) } \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 9 \\ 4 & 3 & 2 \end{pmatrix} \underline{X} = \begin{pmatrix} 4 & 12 \\ 4 & 12 \\ 1 & 3 \end{pmatrix}.$$

17. Solve the previous problems by using the solution formula $\underline{X} = \underline{A}^{-1} \underline{B}$, where \underline{A} is the coefficient matrix and \underline{B} is the right-hand side matrix.

18. Invert the coefficient matrices of Problem 12.

19. Illustrate Theorem 3.11 with the matrix equations of Problem 16.

20. Repeat Example 3.20 with input coefficient matrix

$$\underline{A} = \begin{pmatrix} 0.05 & 0.10 & 0.05 \\ 0.10 & 0.20 & 0.10 \\ 0.05 & 0.05 & 0.20 \end{pmatrix}$$

and output vector

$$\underline{y} = \begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix}.$$

21. Find the interpolation polynomial based on data values

$$x_1 = -2, x_2 = -1, x_3 = 0, x_4 = 1$$

and

$$f_1 = 5, f_2 = 2, f_3 = 1, f_4 = 2.$$

22. It is known from numerical analysis that the unique interpolation polynomial based on data set $(x_i, f_i), i = 1, 2, \dots, n$, can be obtained as

$$p(x) = f_1 l_1(x) + f_2 l_2(x) + \dots + f_n l_n(x),$$

where

$$l_i(x) = \frac{(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}$$

(see, for example, Szidarovszky and Yakowitz, 1978). Solve Problem 3.21 by using this formula.

23. (Continuation of Problem 21). Explain why the existence of a unique interpolation polynomial implies that the coefficient matrix of system (3.22) is nonsingular if the values x_1, x_2, \dots, x_n are distinct.
24. Repeat Example 3.22 with data values of Problem 21. Explain why the same result is obtained as in Problem 21.
25. The least squares problem discussed in Section 3.8 can be generalized as follows. Let $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ be given n -element real vectors, and let $\underline{x} \in R^n$ be given. Find the linear combination $c_1 \underline{x}_1 + c_2 \underline{x}_2 + \dots + c_k \underline{x}_k$ that has minimal distance from \underline{x} . This problem is usually solved as follows. Define matrix $\underline{X} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k)$, then the vector of coefficients $\underline{c} = (c_1, c_2, \dots, c_k)^T$ is obtain as the solution of equation

$$(\underline{X}^T \underline{X}) \underline{c} = \underline{X}^T \underline{x}.$$

Repeat Example 2.33 and Problem 2/25 by using this method.

Chapter 4

Determinants

4.1 Introduction

The most simple "system" of linear equations has only one equation and one unknown:

$$a_{11}x = b_1,$$

where a_{11} and b_1 are real or complex scalars. If $a_{11} \neq 0$, then a unique solution exists:

$$x = \frac{b_1}{a_{11}}. \quad (4.1)$$

If $a_{11} = 0$ and $b_1 \neq 0$, then there is no solution; and if $a_{11} = b_1 = 0$, then x can take any arbitrary real or complex value.

Consider next the case of two equations with two unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

which can be summarized as

$$\underline{Ax} = \underline{b}$$

with the 2×2 matrix $\underline{A} = (a_{ij})$ and the 2-element vector $\underline{b} = (b_i)$.

The application of the elimination method gives the solution:

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}} \quad \text{and} \quad x_2 = \frac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}} \quad (4.2)$$

if the denominator $a_{11} a_{22} - a_{12} a_{21}$ is nonzero. Otherwise, there is no solution if at least one of the numerators is nonzero, and there are infinitely many solutions if both numerators are equal to zero. Notice that the denominator depends on only the coefficient matrix. The numerator of x_1 can be obtained from the denominator if the first column of the coefficient matrix is replaced by the right-hand side vector. Similarly, the numerator of x_2 can be obtained from the denominator by replacing the second column of the coefficient matrix by the right-hand side vector. Since the numerator and denominator have the same structure, it will be useful in analyzing systems of linear equations to examine these quantities and their main properties in detail. Notice first that the denominators are scalars, and are defined for square (1×1 , 2×2) matrices. If \underline{A} denotes a square matrix, then we will denote this quantity by $\det(\underline{A})$ and call it the *determinant* of matrix \underline{A} . From equations (4.1) and (4.2) we see that for 1×1 and 2×2 matrices

$$\det(a_{11}) = a_{11}$$

and

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21}.$$

Consider next the case of three linear equations with three unknowns. Similar expressions are obtained for the solutions with a common

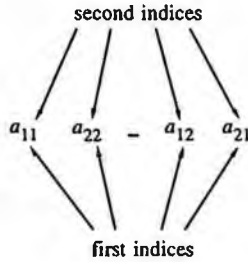
denominator, which has to be therefore the definition of the determinant of 3×3 matrices:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} a_{22} a_{33} + a_{13} a_{21} a_{32} + a_{12} a_{23} a_{31} \\ - a_{13} a_{22} a_{31} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33}.$$

In this chapter the concept of determinants of square matrices will be first introduced for the general $n \times n$ case, and the main properties of determinants will be examined including their applications in solving systems of linear equations.

The general definition of the determinants of $n \times n$ matrices will be based on the common properties and characteristics of the above special cases.

Notice first that each term has n factors. In the 1×1 case only one element, a_{11} , gives the determinant; in the 2×2 case there are two terms, each term is a product of two matrix elements; and in the 3×3 case, there are six terms and each of them is the product of 3 matrix elements. Two questions have to be answered here: what rule decides which matrix elements are multiplied in each term, and how to determine the sign of each term. Consider first an arbitrary term of the above special determinant expressions. In the 1×1 case, the first index of the only term is 1, the second index is also 1. In the 2×2 case two matrix elements are multiplied in each term. The first indices of the two elements are in order 1 and 2 in both terms. The second indices in the first (positive) term are 1 and 2, and those in the second (negative) term are 2 and 1 as shown in the following diagram:



A similar structure is observed in the 3·3 case. In each term, the first indices are always 1, 2, and 3 in this natural order. In the positive terms, the second indices are

$$1,2,3; \quad 3,1,2; \quad \text{and} \quad 2,3,1,$$

and in the negative terms they are

$$3,2,1; \quad 1,3,2; \quad \text{and} \quad 2,1,3.$$

From the above description we realize that in each term, the first indices are the integers 1, or 1,2 or 1,2,3 always in increasing order, however the second indices form all permutations of these integers. In the 2·2 case, the two possible permutations $\langle 1,2 \rangle$ and $\langle 2,1 \rangle$ form the sequences of the second indices. In the 3·3 case there are $3! = 6$ possible permutations of the numbers 1,2,3, and the above six sequences of the second indices coincide with these permutations. We can summarize the above discussion by noticing that in the determinants of 2·2 and 3·3 matrices there are $2!$ and $3!$ terms, respectively, and each term has the form

$$a_{1i_1} a_{2i_2} \quad \text{or} \quad a_{1i_1} a_{2i_2} a_{3i_3},$$

where $\langle i_1, i_2 \rangle$ or $\langle i_1, i_2, i_3 \rangle$ is a permutation of the integers 1,2 or 1,2,3, respectively. In addition, each permutation of the integers 1,2 or 1,2,3 can be found exactly once as a sequence of the second indices. In order to see how the sign of each term is selected, we have to examine the set of permutations associated to positive terms as well as that associated to the negative terms. In the $2 \cdot 2$ case, the permutation $\langle 1,2 \rangle$ gives the positive term and $\langle 2,1 \rangle$ corresponds to the negative term. Notice that the first permutation is the natural (increasing) order of the integers 1,2. In the second permutation we have to interchange the elements 2 and 1 in order to obtain the natural order. In any permutation, the interchange of any two elements is called *inversion*. So, in the positive term no inversion is needed to obtain the natural order, and in the negative term one inversion is required. Consider next the positive terms in the $3 \cdot 3$ case. The first such permutation is $\langle 1,2,3 \rangle$ with no inversion required, since the elements are already in the natural order. The second such permutation is $\langle 3,1,2 \rangle$. We need one inversion to move 1 to the first position. After interchanging 1 and 3, the resulting permutation is $\langle 1,3,2 \rangle$, and an additional inversion (interchanging 3 and 2) is needed to obtain the natural order. In the case of the third such permutation two inversions are needed again:

$$\langle 2,3,1 \rangle \text{ (interchange 2 and 1)} \rightarrow \langle 1,3,2 \rangle \text{ (interchange 3 and 2)} \rightarrow \langle 1,2,3 \rangle.$$

Consider next the negative terms. Permutation $\langle 3,2,1 \rangle$ needs only one inversion (interchanging 3 and 1) to obtain the natural order. The same number of inversions are needed in the case of the other two such permutations:

$$\langle 1,3,2 \rangle \text{ (interchange 3 and 2)} \rightarrow \langle 1,2,3 \rangle$$

and

$\langle 2, 1, 3 \rangle$ (interchange 2 and 1) $\rightarrow \langle 1, 2, 3 \rangle$.

Notice that in the case of positive terms 0 or 2 inversions are needed, and for negative terms only one is required. This notion can be generally formulated as follows.

Definition 4.1. Let i_1, i_2, \dots, i_n be a permutation of the integers $1, 2, \dots, n$. Let π denote this permutation and let $I(\pi)$ be the minimum number of inversions needed to obtain the natural order from the given permutation. The *sign* of permutation π is denoted by $\text{sign}(\pi)$ and is defined as

$$\text{Sign}(\pi) = \begin{cases} 1 & \text{if } I(\pi) \text{ is even} \\ -1 & \text{if } I(\pi) \text{ is odd.} \end{cases}$$

We mention that if $I(\pi)$ is even, then any (not necessarily minimal) rearrangement of permutation π into the increasing order always requires even number of inversions, and if $I(\pi)$ is odd, then all rearrangements require an odd number of inversions. Therefore in determining the sign of a permutation we do not need to find the minimum number of inversions required to rearrange the numbers into natural order. It is sufficient to find one of such rearrangements and check if the number of inversions is even or odd.

Definition 4.2. The *determinant* of an $n \times n$ real or complex matrix $\underline{A} = (a_{ij})$ is defined as

$$\det(\underline{A}) = \sum_{\pi = \langle i_1, \dots, i_n \rangle \in P_n} \text{sign}(\pi) a_{1i_1} a_{2i_2} \dots a_{ni_n}, \quad (4.3)$$

where P_n is the set all $n!$ permutations of the integers $1, 2, \dots, n$.

This definition is illustrated by three examples.

Example 4.1. Assume that \underline{A} is a lower triangular matrix:

$$\underline{A} = \begin{pmatrix} a_{11} & 0 & 0 & \dots & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \dots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}.$$

We will show that $\det(\underline{A})$ is the product of the diagonal elements. Consider an arbitrary term of the sum (4.3):

$$\text{sign}(\pi) a_{1i_1} a_{2i_2} \dots a_{ni_n}.$$

Notice first that $a_{12} = a_{13} = \dots = a_{1n} = 0$, therefore a_{1i_1} is zero for $i_1 \neq 1$. Therefore all such terms can be omitted. We may therefore assume that $i_1 = 1$. Consider next the factor a_{2i_2} . Since $i_2 \neq i_1 = 1$ and with $i_2 = 3, 4, \dots, n$ this matrix element is zero, all terms for $i_2 \neq 2$ are equal to zero. We may therefore assume that $i_2 = 2$. Continue this argument for the third factor, then for the fourth factor, and so on, until the last factor to see that the only nonzero term is

$$a_{11} a_{22} \dots a_{nn},$$

since $\text{sign}(\pi) = 1$ for the natural order $\langle 1, 2, \dots, n \rangle$.

One may similarly prove that the determinant of an upper triangular matrix is also the product of the diagonal elements. In this case we have to

consider the factors in the reverse order $a_{n_i}, a_{n-1, i_{n-1}}, \dots, a_{1i_1}$. (This case will also be a simple consequence of Theorem 4.1 to be given in the next section, since the transpose of an upper triangular matrix is lower triangular.)

◆

Example 4.2. Assume that \underline{A} is a diagonal matrix. Then

$$\det(\underline{A}) = a_{11} a_{22} \cdots a_{nn},$$

since \underline{A} is a special lower triangular matrix with zero elements under the diagonal. For example, $\det(\underline{I}_n) = 1$, where \underline{I}_n denotes the $n \times n$ identity matrix.

◆

Example 4.3. Assume that the k^{th} row of matrix \underline{A} contains only zero elements. Then in each term of (4.3) the factor a_{k_i} is zero, therefore the entire sum is zero. Hence, $\det(\underline{A}) = 0$. One may similarly prove that $\det(\underline{A}) = 0$, if the k^{th} column of \underline{A} consists of only zero elements. Then the factor a_{i_i} (with $i_i = k$) is zero in each term. (We may also apply Theorem 4.1 of the next section.)

◆

The main properties of determinants will be discussed next and their main applications will be outlined.

4.2 Properties of Determinants

The first result to be discussed in this section can be stated as follows. If we interchange the columns of a matrix by its rows then the value of the determinant remains the same.

Theorem 4.1. $\det(\underline{A}) = \det(\underline{A}^T)$.

Proof. Let $\pi = \langle i_1, i_2, \dots, i_n \rangle$ be a permutation of the integers $1, 2, \dots, n$. The inverse of permutation π is defined as the reordering of the numbers i_1, i_2, \dots, i_n into the natural order $1, 2, \dots, n$, and it is denoted by π^{-1} . For example the inverse of $\pi = \langle 3, 2, 1 \rangle$ is itself. The inverse of $\pi = \langle 2, 3, 1 \rangle$ is $\langle 3, 1, 2 \rangle$, since in π , 1 is at the 3rd place, 2 is at the 1st place, and 3 is at the 2nd place. In general, the inverse of permutation $\langle i_1, i_2, \dots, i_n \rangle$ is $\langle j_1, j_2, \dots, j_n \rangle$, where $i_{j_1} = 1, i_{j_2} = 2, \dots, i_{j_n} = n$. It is easy to see that if π runs through all permutations P_n , then the same is true for π^{-1} , furthermore $\text{sign}(\pi) = \text{sign}(\pi^{-1})$. The first statement follows from the facts that the inverses of different permutations are different, and there are exactly $n!$ different permutations of the integers $1, 2, \dots, n$. The second statement is the consequence of the following simple observation. Consider a sequence of pair-wise inversions that transforms permutation π into natural order. If we apply the same sequence in the opposite order to the inverse permutation π^{-1} , then the resulting order of the integers becomes increasing again. Therefore

$$\det(\underline{A}^T) = \sum_{\pi = \langle i_1, i_2, \dots, i_n \rangle \in P_n} \text{sign}(\pi) a_{i_1 1} a_{i_2 2} \dots a_{i_n n},$$

since the (k, i_k) element of \underline{A}^T is $a_{i_k k}$. Let $\langle j_1, \dots, j_n \rangle$ denote the inverse of π , then by rearranging the summation with respect to π^{-1} we have

$$\det(\underline{A}^T) = \sum_{\pi^{-1} = \langle j_1, \dots, j_n \rangle \in P_n} \text{sign}(\pi^{-1}) a_{1j_1} a_{2j_2} \dots a_{nj_n} = \det(\underline{A}).$$

✦

In many cases, as it will be illustrated in the next theorem, it is convenient to manipulate with the rows or the columns of the matrix the determinant of which is under consideration. Let \underline{A} be an $n \times n$ real or complex matrix. Let $\underline{r}_1^T, \dots, \underline{r}_n^T$ be the rows of \underline{A} , and let $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$

denote the columns of \underline{A} . Then instead of $\det(\underline{A})$ we may write $\det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_n^T \end{pmatrix}$

or $\det(\underline{a}_1, \dots, \underline{a}_n)$.

Theorem 4.2. The following properties of $\det(\underline{A})$ are valid:

- (a) If two rows (or columns) of matrix \underline{A} are interchanged, then the determinant of the new matrix is $-\det(\underline{A})$;
- (b) If two rows (or columns) of matrix \underline{A} are identical, then $\det(\underline{A})=0$;
- (c) The value of $\det(\underline{A})$ is multiplied by c if a row (or column) of \underline{A} is multiplied by c ;

(d)

$$\det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_{k-1}^T \\ \underline{r}_k^T + \underline{r}_k^T \\ \underline{r}_{k+1}^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = \det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_{k-1}^T \\ \underline{r}_k^T \\ \underline{r}_{k+1}^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} + \det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_{k-1}^T \\ \underline{r}_k^T \\ \underline{r}_{k+1}^T \\ \dots \\ \underline{r}_n^T \end{pmatrix},$$

and

$$\begin{aligned} & \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k + \underline{a}', \underline{a}_{k+1}, \dots, \underline{a}_n) \\ &= \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k, \underline{a}_{k+1}, \dots, \underline{a}_n) + \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}', \underline{a}_{k+1}, \dots, \underline{a}_n); \end{aligned}$$

(e) The value of $\det(\underline{A})$ does not change, if a constant multiple of a row (or column) is added to another row (or column);

(f) $\det(\underline{A}) \neq 0$ if and only if the rows (or columns) of \underline{A} are linearly independent.

Proof. From Theorem 4.1 we know that interchanging the rows of a matrix by its columns the value of the determinant does not change. Therefore it is sufficient to prove the assertions for rows or for columns only. We will next present the proofs for rows.

(a) If two rows in \underline{A} are interchanged, then each term of sum (4.3) becomes

$$\text{sign}(\pi) a_{l_1} \dots a_{l_k} \dots a_{k_l} \dots a_{n_{l_n}},$$

where $l > k, \pi = \langle i_1, i_2, \dots, i_n \rangle$, and permutation $\langle 1, \dots, k, \dots, l, \dots, n \rangle$ is obtained from the increasing order $\langle 1, 2, \dots, n \rangle$ by interchanging k and l . Since one pair-wise inversion changes the sign of the permutation, this term can be rewritten as

$$- \text{sign}(\pi') a_{i_1} \dots a_{k_i} \dots a_{l_i} \dots a_{n_i},$$

where $\pi' = \langle i_1, \dots, i_l, \dots, i_k, \dots, i_n \rangle$ (that is, π' is obtained by interchanging i_l and i_k in π). Adding up all such terms the value of $-\det(\underline{A})$ is obtained.

(b) By interchanging the identical rows, the matrix, as well as the value of its determinant remains the same. However, from property (a) we know that the value of the determinant changes sign. Therefore

$$\det(\underline{A}) = -\det(\underline{A}),$$

that is, $\det(\underline{A}) = 0$.

(c) If row k of matrix \underline{A} is multiplied by c , then each term of the sum (4.3) is multiplied by c , since the factor a_{k_i} is replaced by $c \cdot a_{k_i}$. Therefore the entire sum is multiplied by c .

(d) Consider an arbitrary term of the sum (4.3):

$$\begin{aligned} & \text{sign}(\pi) a_{i_1} \dots (a_{k_i} + a'_{k_i}) \dots a_{n_i} \\ &= \text{sign}(\pi) a_{i_1} \dots a_{k_i} \dots a_{n_i} + \text{sign}(\pi) a_{i_1} \dots a'_{k_i} \dots a_{n_i}. \end{aligned}$$

Adding up the first terms for all $\pi \in P_n$ gives

$$\det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_k^T \\ \dots \\ \underline{r}_n^T \end{pmatrix},$$

and the sum of the second terms for all $\pi \in P_n$ equals

$$\det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_k^T \\ \dots \\ \underline{r}_n^T \end{pmatrix}.$$

(e) Using properties (d) and (b) we see that

$$\det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_k^T + c\underline{r}_l^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = \det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_k^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} + c \cdot \det \begin{pmatrix} \underline{r}_1^T \\ \dots \\ \underline{r}_l^T \\ \dots \\ \underline{r}_n^T \end{pmatrix}.$$

The first term is the determinant of the original matrix. In the second term, \underline{r}_l^T is in the k^{th} row however the same vector \underline{r}_l^T is also present in the l^{th} row, therefore two rows are identical. Then property (b) implies that the second term is zero.

(f) Assume first that the rows of \underline{A} are linearly dependent. Then with suitable coefficients

$$c_1 \underline{r}_1^T + c_2 \underline{r}_2^T + \dots + c_n \underline{r}_n^T = \underline{0}^T,$$

where at least one coefficient, say c_1 , is nonzero. Then

$$\underline{r}_1^T + d_2 \underline{r}_2^T + \dots + d_n \underline{r}_n^T = \underline{0}^T,$$

where $d_k = c_k / c_1$, ($k = 2, \dots, n$). Then repeated application of property (e) implies that

$$\begin{aligned} \det(\underline{A}) &= \det \begin{pmatrix} \underline{r}_1^T + d_2 \underline{r}_2^T \\ \underline{r}_2^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = \det \begin{pmatrix} \underline{r}_1^T + d_2 \underline{r}_2^T + d_3 \underline{r}_3^T \\ \underline{r}_2^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = \dots \\ &= \det \begin{pmatrix} \underline{r}_1^T + d_2 \underline{r}_2^T + \dots + d_n \underline{r}_n^T \\ \underline{r}_2^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = \det \begin{pmatrix} \underline{0}^T \\ \underline{r}_2^T \\ \dots \\ \underline{r}_n^T \end{pmatrix} = 0, \end{aligned}$$

since in each term of the sum (4.3) applied to the last determinant, the factor a_{1i} is always zero.

Assume next that the rows of matrix \underline{A} are linearly independent. Then $\text{rank}(\underline{A}) = n$, and the system of homogeneous linear equations $\underline{A}\underline{x} = \underline{0}$ has the unique solution $\underline{x} = \underline{0}$. Apply the elimination procedure of Section 3.7 to

solve this system. At each elimination step one of the following operations is performed:

(i) Interchanging two rows and/or columns may only change the sign of the determinant.

(ii) Dividing any row by the pivot coefficient results in dividing the value of the determinant by the nonzero pivot coefficient.

(iii) Adding a constant-multiple of a row to another row does not change the value of the determinant.

Therefore $\det(\underline{A})$ is the nonzero-multiple of the determinant of the last coefficient matrix, which is upper triangular with nonzero diagonal elements. Then Example 4.1 implies that the determinant is nonzero.

♣

Consider next a scalar-valued function f which assigns to each n -tuple $(\underline{a}_1, \dots, \underline{a}_n)$ of n -element vectors a scalar, and the following conditions are satisfied:

$$(A) f(\underline{e}_1, \dots, \underline{e}_n) = 1 \text{ where } \underline{e}_1, \dots, \underline{e}_n \text{ are the natural basis vectors;}$$

(B) $f(\underline{a}_1, \dots, \underline{a}_n)$ changes sign if two of the vectors \underline{a}_i and \underline{a}_j are interchanged;

$$(C) f(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k + \underline{a}'_k, \underline{a}_{k+1}, \dots, \underline{a}_n) \\ = f(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k, \underline{a}_{k+1}, \dots, \underline{a}_n) + f(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}'_k, \underline{a}_{k+1}, \dots, \underline{a}_n);$$

$$(D) f(\underline{a}_1, \dots, \underline{a}_{k-1}, c\underline{a}_k, \underline{a}_{k+1}, \dots, \underline{a}_n) = cf(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k, \underline{a}_{k+1}, \dots, \underline{a}_n).$$

Theorem 4.3. If conditions (A), (B), (C) and (D) are satisfied, then for all n -element vectors $\underline{a}_1, \dots, \underline{a}_n$,

$$f(\underline{a}_1, \dots, \underline{a}_n) = \det(\underline{a}_1, \dots, \underline{a}_n).$$

Proof. Consider the function

$$\Delta(\underline{a}_1, \dots, \underline{a}_n) = f(\underline{a}_1, \dots, \underline{a}_n) - \det(\underline{a}_1, \dots, \underline{a}_n).$$

Then $\Delta(\underline{e}_1, \dots, \underline{e}_n) = 0$ and function Δ satisfies properties (B), (C) and (D). The proof of part (b) of Theorem 4.2 implies that $\Delta(\underline{a}_1, \dots, \underline{a}_n) = 0$ if $\underline{a}_i = \underline{a}_j$ for some $i \neq j$. Let a_{ij} denote the i th element of vector \underline{a}_j ($j = 1, 2, \dots, n; i = 1, 2, \dots, n$), then

$$\underline{a}_j = a_{1j}\underline{e}_1 + \dots + a_{nj}\underline{e}_n,$$

therefore

$$\begin{aligned} \Delta(\underline{a}_1, \dots, \underline{a}_n) &= \Delta\left(\sum_{l_1} a_{l_1 1} \underline{e}_{l_1}, \underline{a}_2, \dots, \underline{a}_n\right) \\ &= \sum_{l_1} a_{l_1 1} \Delta(\underline{e}_{l_1}, \underline{a}_2, \dots, \underline{a}_n) = \sum_{l_1} a_{l_1 1} \Delta\left(\underline{e}_{l_1}, \sum_{l_2} a_{l_2 2} \underline{e}_{l_2}, \underline{a}_3, \dots, \underline{a}_n\right) \\ &= \sum_{l_1} a_{l_1 1} \sum_{l_2} a_{l_2 2} \Delta(\underline{e}_{l_1}, \underline{e}_{l_2}, \underline{a}_3, \dots, \underline{a}_n) = \dots \\ &= \sum_{l_1} \sum_{l_2} \dots \sum_{l_n} a_{l_1 1} a_{l_2 2} \dots a_{l_n n} \Delta(\underline{e}_{l_1}, \dots, \underline{e}_{l_n}). \end{aligned} \tag{4.4}$$

We will next prove that each term is zero. Notice first that $\Delta(\underline{e}_{l_1}, \dots, \underline{e}_{l_n}) = 0$ if two numbers l_i and l_j are equal. If the numbers l_1, l_2, \dots, l_n are different, then they form a permutation of the integers $1, 2, \dots, n$, therefore

$$\Delta(\underline{e}_1, \dots, \underline{e}_n) = \begin{cases} \Delta(\underline{e}_1, \dots, \underline{e}_n) & \text{if even number of inversions is} \\ & \text{needed to bring permutation } \langle l_1, \dots, l_n \rangle \text{ to natural} \\ & \text{order} \\ -\Delta(\underline{e}_1, \dots, \underline{e}_n) & \text{if an odd number of inversions is needed} \\ & \text{to bring permutation } \langle l_1, \dots, l_n \rangle \text{ to natural order} \end{cases}$$

Since $\Delta(\underline{e}_1, \dots, \underline{e}_n) = 0$, all terms in (4.4) are equal to zero. Therefore $\Delta(\underline{a}_1, \dots, \underline{a}_n) = 0$, that is

$$f(\underline{a}_1, \dots, \underline{a}_n) = \det(\underline{a}_1, \dots, \underline{a}_n).$$

Thus the proof is completed.

Corollary. Assume that function f satisfies only conditions (B), (C), and (D). Then for all n -vectors $\underline{a}_1, \dots, \underline{a}_n$,

$$f(\underline{a}_1, \dots, \underline{a}_n) = \det(\underline{a}_1, \dots, \underline{a}_n) \cdot f(\underline{e}_1, \dots, \underline{e}_n). \quad (4.5)$$

Proof. If $f(\underline{e}_1, \dots, \underline{e}_n) = 1$, then the assertion follows from the theorem. Otherwise, consider the function

$$g(\underline{a}_1, \dots, \underline{a}_n) = \frac{\Delta(\underline{a}_1, \dots, \underline{a}_n)}{f(\underline{e}_1, \dots, \underline{e}_n) - 1},$$

which satisfies all conditions (A), (B), (C), and (D). Therefore

$$g(\underline{a}_1, \dots, \underline{a}_n) = \det(\underline{a}_1, \dots, \underline{a}_n),$$

and from the last equation we see that

$$\det(\underline{a}_1, \dots, \underline{a}_n) = \frac{f(\underline{e}_1, \dots, \underline{e}_n) - \det(\underline{a}_1, \dots, \underline{a}_n)}{f(\underline{e}_1, \dots, \underline{e}_n) - 1},$$

which is equivalent to the assertion. *

One of the most frequently used properties of determinants is known as the multiplication theorem, which is formulated next.

Theorem 4.4. Let \underline{B} and \underline{A} be $n \times n$ real or complex matrices. Then

$$\det(\underline{B}\underline{A}) = \det(\underline{B}) \cdot \det(\underline{A}).$$

Proof. Define the function

$$f(\underline{a}_1, \dots, \underline{a}_n) = \det(\underline{B}\underline{a}_1, \dots, \underline{B}\underline{a}_n)$$

which obviously satisfies properties (B), (C), and (D). Therefore the Corollary of Theorem 4.3 implies that

$$\det(\underline{B}\underline{a}_1, \dots, \underline{B}\underline{a}_n) = \det(\underline{a}_1, \dots, \underline{a}_n) \cdot \det(\underline{B}\underline{e}_1, \dots, \underline{B}\underline{e}_n).$$

Let now $\underline{a}_1, \dots, \underline{a}_n$ be the columns of \underline{A} . Then

$$\begin{aligned}(\underline{a}_1, \dots, \underline{a}_n) &= \underline{A}, \\(\underline{B}\underline{a}_1, \dots, \underline{B}\underline{a}_n) &= \underline{B} \cdot (\underline{a}_1, \dots, \underline{a}_n) = \underline{B}\underline{A}, \\(\underline{B}\underline{e}_1, \dots, \underline{B}\underline{e}_n) &= \underline{B} \cdot (\underline{e}_1, \dots, \underline{e}_n) = \underline{B} \cdot \underline{I} = \underline{B}\end{aligned}$$

and hence, the proof is complete. ♦

Corollary. If \underline{A}^{-1} exists, then

$$\det(\underline{A}^{-1}) = \frac{1}{\det(\underline{A})},$$

since

$$1 = \det(\underline{I}_n) = \det(\underline{A}^{-1} \cdot \underline{A}) = \det(\underline{A}^{-1}) \det(\underline{A}).$$

Remark. Recall that $\underline{A}\underline{B}$ and $\underline{B}\underline{A}$ are usually different matrices. The assertion of the theorem implies that their determinants are the same even if they are different.

4.3 Cofactors and Expansion by Cofactors

The value of $\det(\underline{A})$ for $2 \cdot 2$ or $3 \cdot 3$ matrices can be obtained easily from the definition. In the $2 \cdot 2$ case,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

where the two terms can be obtained by the method illustrated in Figure 4.1.

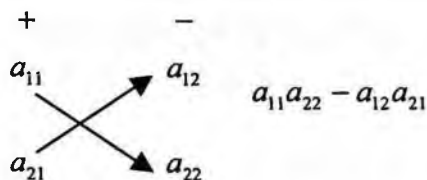


Figure 4.1. Diagram for evaluating $2 \cdot 2$ determinants.

The six terms of the determinants of a $3 \cdot 3$ matrix can be obtained by a similar, but more complicated diagram, which is shown in Figure 4.2. The reader should be cautioned, since there is no similar way to find the values of determinants of higher order. In this section a general method will be introduced to find the values of $n \times n$ determinants in general. Our method will be based on the following concept.

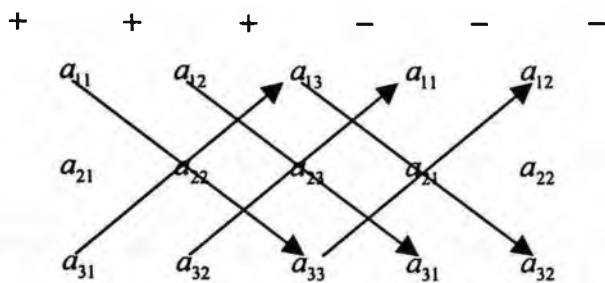


Figure 4.2. Diagram for evaluating $3 \cdot 3$ determinants.

Definition 4.3. Let $\underline{A} = (a_{ij})$ be an $n \times n$ matrix. The (i, j) cofactor A_{ij} of \underline{A} is defined as

$$A_{ij} = (-1)^{i+j} D_{ij},$$

where D_{ij} is the determinant of the $(n-1) \times (n-1)$ matrix that is obtained by deleting the i^{th} row and j^{th} column of matrix \underline{A} .

The main result of this section can be formulated as follows.

Theorem 4.5. Let $\underline{A} = (a_{ij})$ be an $n \times n$ matrix. Then for $i = 1, 2, \dots, n$,

$$\det(\underline{A}) = \sum_{j=1}^n a_{ij} A_{ij}, \quad (4.6)$$

where A_{ij} is the (i, j) cofactor of \underline{A} .

Proof. Fix the value of i and define function

$$f(\underline{a}_1, \dots, \underline{a}_n) = \sum_{j=1}^n a_{ij} A_{ij}$$

where a_{ij} is the i^{th} element of vector \underline{a}_j and A_{ij} is the (i, j) cofactor of matrix $(\underline{a}_1, \dots, \underline{a}_n)$. It is easy to prove that f satisfies all conditions of Theorem 4.3:

(A) If $\underline{A} = \underline{I}_n$, then from (4.6) we see that

$$\det(\underline{A}) = a_{ii} A_{ii} = 1,$$

because all other terms are equal to zero (since $a_{ij} = 0$ for $i \neq j$) and $a_{ii} = 1$ and $A_{ii} = \det(\underline{I}_{n-1}) = 1$.

(B) Assume that columns k and l ($k < l$) are interchanged. Let a'_{ij} and A'_{ij} denote the elements and cofactors of the new matrix. Then it is easy to see that

$$a'_{ij} = \begin{cases} a_{ij} & \text{if } j = k \\ a_{ik} & \text{if } j = l \\ a_{ij} & \text{otherwise,} \end{cases}$$

and

$$A'_{ij} = \begin{cases} -A_{ij} & \text{if } j = k \\ -A_{ik} & \text{if } j = l \\ -A_{ij} & \text{otherwise.} \end{cases}$$

Therefore each term changes sign.

(C) Assume that column \underline{a}_k is replaced by $\underline{a}_k + \underline{a}'_k$. If $j = k$, then for all i , the A_{ij} cofactor remains the same but the element a_{ij} has to be replaced by the sum $a_{ij} + a'_{ij}$. If $j \neq k$, then the element a_{ij} remains the same and the cofactor is replaced by $A_{ij} + A'_{ij}$, where A'_{ij} is obtained from A_{ij} by changing the elements a_{ik} to a'_{ik} for $l \neq i$. In both cases each term breaks up as the sum of the corresponding terms with k^{th} column \underline{a}_k and \underline{a}'_k , respectively.

(D) Assume next that column \underline{a}_k is multiplied by a scalar c . Then each term in the definition of f is multiplied by c : for $j = k$, a_{ik} is multiplied by c ; and for $j \neq k$, the cofactor A_{ij} is multiplied by c , since one of its columns is multiplied by this scalar.

These properties and Theorem 4.3 implies that f coincides with the determinant function.

Corollary 1. Combining Theorems 4.1 and 4.5 we conclude that for all j ,

$$\det(\underline{A}) = \sum_{i=1}^n a_{ij} A_{ij}. \quad (4.7)$$

Formulas (4.6) and (4.7) are called the *row* and *column expansions* of the determinant, respectively.

Corollary 2. For $i \neq k$

$$\sum_{j=1}^n a_{ij} A_{kj} = 0 \quad (4.8)$$

and for $j \neq l$,

$$\sum_{i=1}^n a_{ij} A_{il} = 0. \quad (4.9)$$

The left-hand side of equation (4.8) is the determinant of the matrix which is obtained from \underline{A} by replacing its k^{th} row by its i^{th} row. The i^{th} and k^{th} rows of the resulting matrix therefore become identical, and the determinant of this matrix is necessarily zero. The left-hand side of equation (4.9) is the determinant of the matrix which is obtained from \underline{A} by replacing its l^{th} column by its j^{th} column, and since this matrix has two identical columns, its determinant is zero.

The expansion formulas are illustrated next.

Example 4.4. We will expand the determinant of the 4×4 matrix

$$\underline{A} = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

by repeated application of the row expansion formula. Selecting always the first row,

$$\begin{aligned} \det(\underline{A}) &= 2 \cdot (-1)^{1+1} \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} + 1 \cdot (-1)^{1+2} \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \\ &+ 1 \cdot (-1)^{1+3} \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} + 1 \cdot (-1)^{1+4} \det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \end{aligned}$$

The 3×3 determinants are obtained in the same way:

$$\begin{aligned} \det \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} &= 2 \cdot (-1)^{1+1} \det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} + 1 \cdot (-1)^{1+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &+ 1 \cdot (-1)^{1+3} \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = 2 \cdot (4-1) - 1 \cdot (2-1) + 1 \cdot (1-2) = 6-1-1=4, \\ \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} &= 1 \cdot (-1)^{1+1} \det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} + 1 \cdot (-1)^{1+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &+ 1 \cdot (-1)^{1+3} \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = 1 \cdot (4-1) - 1 \cdot (2-1) + 1 \cdot (1-2) = 3-1-1=1, \end{aligned}$$

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} = 1 \cdot (-1)^{1+1} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + 2 \cdot (-1)^{1+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ + 1 \cdot (-1)^{1+3} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1 \cdot (2-1) - 2 \cdot (2-1) + 1 \cdot (1-1) = 1 - 2 + 0 = -1$$

(which result can be immediately obtained from the previous determinant, since the first two rows are interchanged),

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = 1 \cdot (-1)^{1+1} \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} + 2 \cdot (-1)^{1+2} \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \\ + 1 \cdot (-1)^{1+3} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1 \cdot (1-2) - 2 \cdot (1-2) + 1 \cdot (1-1) = -1 + 2 + 0 = 1$$

(which result is obvious from the value of the previous determinant, since the last two rows are interchanged). Therefore

$$\det(\underline{A}) = 2 \cdot 4 - 1 \cdot 1 + 1 \cdot (-1) - 1 \cdot 1 = 5.$$

◆

4.4 Determinants and Systems of Linear Equations

We start this section with a theorem, which is the restatement and combination of Theorems 3.3 and 3.13.

Theorem 4.6. Let \underline{A} be an $n \times n$ matrix. The following conditions are equivalent to each other:

- (a) The system $\underline{Ax} = \underline{b}$ has a unique solution for all n -element vectors \underline{b} ;
- (b) \underline{A}^{-1} exists;
- (c) The columns of \underline{A} are linearly independent;
- (d) The rows of \underline{A} are linearly independent;
- (e) $\text{rank}(\underline{A}) = n$;
- (f) $\det(\underline{A}) \neq 0$.

Notice that condition (f) gives a sufficient and necessary condition for the existence of a unique solution of system $\underline{Ax} = \underline{b}$. By using the concept of determinants, we can even solve this system. Assume that $\det(\underline{A}) \neq 0$. For $k = 1, 2, \dots, n$, introduce the notation

$$D_k = \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{b}, \underline{a}_{k+1}, \dots, \underline{a}_n)$$

Here D_k is the determinant of the matrix which is obtained from \underline{A} by replacing its k^{th} column by the right-hand side vector. The columns of \underline{A} are denoted again by $\underline{a}_1, \dots, \underline{a}_n$.

Theorem 4.7. If $\det(\underline{A}) \neq 0$, then for $k = 1, 2, \dots, n$, the k^{th} element of the solution of system $\underline{Ax} = \underline{b}$ is obtained as

$$x_k = \frac{D_k}{\det(\underline{A})}. \quad (4.10)$$

Proof. If $\det(\underline{A}) \neq 0$ then there is a unique solution, $\underline{x} = (x_i)$. Using the properties of determinant, for $k = 1, 2, \dots, n$, we have

$$\begin{aligned} D_k &= \det\left(\underline{a}_1, \dots, \underline{a}_{k-1}, \sum_{i=1}^n x_i \underline{a}_i, \underline{a}_{k+1}, \dots, \underline{a}_n\right) \\ &= \sum_{i=1}^n x_i \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_i, \underline{a}_{k+1}, \dots, \underline{a}_n). \end{aligned}$$

If $i \neq k$, then the i^{th} and k^{th} columns of matrix $(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_i, \underline{a}_{k+1}, \dots, \underline{a}_n)$ are identical, so its determinant is zero. Therefore

$$D_k = x_k \det(\underline{a}_1, \dots, \underline{a}_{k-1}, \underline{a}_k, \underline{a}_{k+1}, \dots, \underline{a}_n) = x_k \cdot \det(\underline{A}),$$

which implies the assertion. ♣

Remark. Relations (4.10) are called the *Cramer's rule*.

Example 4.5. We will apply Cramer's rule to solve the system

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 2x_1 + 3x_2 - x_3 &= 4 \\ 3x_1 + 5x_2 + x_3 &= 9, \end{aligned}$$

which was the subject of our earlier Example 3.11. Expanding the $3 \cdot 3$ determinants with respect to their first rows we have

$$\det(\underline{A}) = \det \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ 3 & 5 & 1 \end{pmatrix} = 1 \cdot (3+5) - 1 \cdot (2+3) + 1 \cdot (10-9) = 4,$$

$$D_1 = \det \begin{pmatrix} 3 & 1 & 1 \\ 4 & 3 & -1 \\ 9 & 5 & 1 \end{pmatrix} = 3 \cdot (3+5) - 1 \cdot (4+9) + 1 \cdot (20-27) = 4,$$

$$D_2 = \det \begin{pmatrix} 1 & 3 & 1 \\ 2 & 4 & -1 \\ 3 & 9 & 1 \end{pmatrix} = 1 \cdot (4+9) - 3 \cdot (2+3) + 1 \cdot (18-12) = 4,$$

and

$$D_3 = \det \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 9 \end{pmatrix} = 1 \cdot (27-20) - 1 \cdot (18-12) + 3 \cdot (10-9) = 4.$$

Therefore

$$x_1 = \frac{D_1}{\det(\underline{A})} = 1, \quad x_2 = \frac{D_2}{\det(\underline{A})} = 1, \quad \text{and} \quad x_3 = \frac{D_3}{\det(\underline{A})} = 1.$$

◆

In Theorem 4.6 we observed a strong relation between the ranks and determinants of $n \times n$ matrices by noticing that $\text{rank}(\underline{A}) = n$ if and only if

$\det(\underline{A}) \neq 0$. We can largely extend this result for general $m \times n$ matrices by using a new concept known as minor.

Definition 4.4. Let \underline{A} be an $m \times n$ real or complex matrix. An r -rowed minor (determinant) of \underline{A} is the determinant of an $r \times r$ matrix obtained from \underline{A} by deleting $m - r$ rows and $n - r$ columns.

As $m \times n$ matrix therefore has $\binom{m}{r} \cdot \binom{n}{r}$ r -rowed minors, since the remaining rows can be selected in $\binom{m}{r}$ different ways, and the remaining columns can be selected in $\binom{n}{r}$ different ways. The construction of minors is illustrated first, then we will prove an important theorem relating matrix ranks to basic properties of minors.

Example 4.6. Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \end{pmatrix}$$

has three 2-rowed minors

$$\det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \det \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}, \quad \text{and} \quad \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

Similarly, matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

has nine 2-rowed minors:

$$\begin{aligned} & \det \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}, \det \begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix}, \det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \\ & \det \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix}, \det \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}, \det \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \\ & \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \det \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}, \det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}. \end{aligned}$$

Notice that the numbers of the 2×2 minors are given as

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 1 \cdot 3 = 3$$

and

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot 3 = 9.$$

Theorem 4.8. Let A be an $m \times n$ nonzero matrix. The rank of A is r if and only if there is at least one nonzero r -rowed minor, and all higher-rowed minors are zero. ♦

Proof. Assume that $\text{rank}(\underline{A}) = r$, and the size of the largest nonzero minor is $s \times s$. We will prove that $r = s$. Since $\text{rank}(\underline{A}) = r$, there are r linearly independent rows of \underline{A} . Let \underline{A}' denote the submatrix of \underline{A} after deleting all other rows. The fact that $\text{rank}(\underline{A}') = r$ implies that there are r linearly independent columns of \underline{A}' . Let \underline{A}'' denote the submatrix of \underline{A}' after deleting all other columns. Then \underline{A}'' is $r \times r$ and its columns are linearly independent implying that $\det(\underline{A}'') \neq 0$. Since s denotes the size of the largest nonzero minor, $s \geq r$. Consider next the nonzero $s \times s$ minor. Its rows are linearly independent which implies that the corresponding entire rows of \underline{A} are also linearly independent. Hence $r \geq s$, which completes the proof. +

Example 4.7. We will now determine the rank of matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

by applying the above theorem. The only 3-rowed minor is the determinant of \underline{A} :

$$\det(\underline{A}) = 1 \cdot (6 - 4) - 1 \cdot (6 - 2) + 1 \cdot (4 - 2) = 2 - 4 + 2 = 0,$$

where we used expansion with respect to the first row. Therefore $\text{rank}(\underline{A}) < 3$. However there is at least one 2-rowed nonzero minor

$$\det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 2 - 1 = 1 \neq 0,$$

which is obtained by deleting the last column and the second row of \underline{A} . Hence, $\text{rank}(\underline{A}) = 2$.

◆

4.5 Further Examples and Applications

The most important application of determinants in solving systems of linear equations and in finding matrix ranks was already discussed in this chapter. Another application will be introduced later in Section 6.2, where a parametric determinant will be expanded in order to find the characteristics equation of a square matrix. In Section 7.4, several special matrix classes will be introduced based on the signs of the principal minors (which are the determinants of certain parts of the matrices in question). In this section some additional applications will be introduced.

1. We give first a *closed form representation* to find the *inverse* of a nonsingular $n \times n$ matrix \underline{A} :

$$\underline{A}^{-1} = \frac{1}{\det(\underline{A})} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}. \quad (4.11)$$

Notice that the matrix of the right-hand side can be obtained from \underline{A} by replacing each element by its cofactor and transposing the resulting matrix.

This formula can be proved in the following way. Consider the following matrix product:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

The (i, j) element of the product can be obtained as

$$\sum_{k=1}^n a_{ik} A_{jk} = \begin{cases} 0 & \text{if } i \neq j \\ \det(\underline{A}) & \text{if } i = j, \end{cases}$$

which follows from relations (4.6) and (4.8). Therefore the above matrix product is $\det(\underline{A}) \cdot \underline{I}_n$, which implies equality (4.11).

Example 4.8. Consider first a $2 \cdot 2$ nonsingular matrix

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then Definition 4.3 implies that

$$\begin{aligned} A_{11} &= (-1)^{1+1} a_{22} = a_{22}, \quad A_{12} = (-1)^{1+2} a_{21} = -a_{21}, \\ A_{21} &= (-1)^{2+1} a_{12} = -a_{12}, \quad \text{and } A_{22} = (-1)^{2+2} a_{11} = a_{11}. \end{aligned}$$

Therefore equation (4.11) shows that

$$\underline{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

◆

Example 4.9. We will next find the inverse of the $3 \cdot 3$ matrix

$$\underline{A} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & 7 \\ 2 & 5 & 9 \end{pmatrix}$$

which was the subject of our earlier Example 3.9. Simple calculation shows that

$$A_{11} = \det \begin{pmatrix} 4 & 7 \\ 5 & 9 \end{pmatrix} = 36 - 35 = 1, \quad A_{12} = -\det \begin{pmatrix} 4 & 7 \\ 2 & 9 \end{pmatrix} = -22,$$

$$A_{13} = \det \begin{pmatrix} 4 & 4 \\ 2 & 5 \end{pmatrix} = 12, \quad A_{21} = -\det \begin{pmatrix} 1 & 3 \\ 5 & 9 \end{pmatrix} = 6,$$

$$A_{22} = \det \begin{pmatrix} 2 & 3 \\ 2 & 9 \end{pmatrix} = 12, \quad A_{23} = -\det \begin{pmatrix} 2 & 1 \\ 2 & 5 \end{pmatrix} = -8,$$

$$A_{31} = \det \begin{pmatrix} 1 & 3 \\ 4 & 7 \end{pmatrix} = -5, \quad A_{32} = -\det \begin{pmatrix} 2 & 3 \\ 4 & 7 \end{pmatrix} = -2,$$

and

$$A_{33} = \det \begin{pmatrix} 2 & 1 \\ 4 & 4 \end{pmatrix} = 4.$$

Expanding $\det(\underline{A})$ with respect to the first row we have

$$\det(\underline{A}) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = 2(1) + 1(-22) + 3(12) = 16,$$

and therefore,

$$\underline{A}^{-1} = \frac{1}{16} \begin{pmatrix} 1 & 6 & -5 \\ -22 & 12 & -2 \\ 12 & -8 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{16} & \frac{3}{8} & -\frac{5}{16} \\ -\frac{11}{8} & \frac{3}{4} & -\frac{1}{8} \\ \frac{3}{4} & -\frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

This result coincides with that obtained in Example 3.9 as it should.

◆

2. Consider a single-input discrete dynamic system with state transition equation

$$\underline{x}(t+1) = \underline{A}\underline{x}(t) + \underline{b}u(t), \quad \underline{x}(0) = \underline{x}_0 \quad (4.12)$$

where t denotes the time periods $0, 1, 2, \dots$, $\underline{x}(t)$ and $u(t)$ are the n -element real state vector and single-dimensional real input at time period t , \underline{A} is a given $n \times n$ constant real matrix and \underline{b} is a given real n -vector. An important problem of systems theory is the controllability of this system. We say that system (4.12) is completely controllable if for arbitrary time period $T \geq n$ and vector $\underline{x}_T \in R^n$ there exists an input sequence $u(0), u(1), \dots, u(T-1)$ such that at time period T , the state of the system coincides with the given vector \underline{x}_T . It is known from the theory of linear systems (see, for example, Szidarovszky and Bahill, 1992) that the above system is completely controllable if and only if the Kalman matrix

$$\underline{K} = (\underline{b}, \underline{A}\underline{b}, \underline{A}^2\underline{b}, \dots, \underline{A}^{n-1}\underline{b})$$

is nonsingular. We know from Theorem 4.6 that \underline{K} is nonsingular if and only if $\det(\underline{K}) \neq 0$. This observation gives a simple controllability check, which is illustrated in the following example.

Example 4.10. We will now show that system

$$\underline{x}(t+1) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} u(t)$$

is completely controllable.

Simple calculation shows that

$$\underline{A}\underline{b} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$\underline{A}^2 \underline{b} = \underline{A}(\underline{A}\underline{b}) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Therefore

$$\underline{K} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

which is an upper triangular matrix. From Example 4.1 we know that $\det(\underline{K})$ equals the product of the diagonal elements of \underline{K} , that is, $\det(\underline{K}) = 1 \neq 0$. Consequently, the system is completely controllable. ♦

3. The controllability of continuous single-input linear systems

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{b}u(t), \underline{x}(0) = \underline{x}_0 \quad (4.13)$$

also depends on the Kalman matrix

$$\underline{K} = (\underline{b}, \underline{A}\underline{b}, \underline{A}^2\underline{b}, \dots, \underline{A}^{n-1}\underline{b})$$

It is also known from systems theory that the system is completely controllable if and only if \underline{K} is nonsingular. Hence $\det(\underline{K}) \neq 0$ is a sufficient and necessary condition for complete controllability.

4. In Chapter 6 we will see that the eigenvalues λ of an $n \times n$ matrix \underline{A} are defined by the equation

$$\underline{A}\underline{x} = \lambda\underline{x} \quad (4.14)$$

where $\underline{x} \neq \underline{0}$. Since this equation can be rewritten as

$$(\underline{A} - \lambda\underline{I})\underline{x} = \underline{0}$$

and this homogeneous equation has nonzero if and only if the determinant of the coefficient matrix is zero, we conclude that all eigenvalues must satisfy the determinant equation

$$\det(\underline{A} - \lambda\underline{I}) = 0. \quad (4.15)$$

Expanding this determinant, an n -degree polynomial is obtained, which can be solved by computer methods. This idea is illustrated next.

Example 4.11. Consider matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ 3 & 4 & 4 \end{pmatrix}.$$

Since \underline{I} is the identity matrix with unit diagonal elements, equation (4.15) can be written now as

$$\det \begin{pmatrix} 1-\lambda & 2 & 1 \\ 2 & 2-\lambda & 3 \\ 3 & 4 & 4-\lambda \end{pmatrix} = 0.$$

By expanding the determinant by the cofactors of the first column we have

$$\begin{aligned} \det \begin{pmatrix} 1-\lambda & 2 & 1 \\ 2 & 2-\lambda & 3 \\ 3 & 4 & 4-\lambda \end{pmatrix} &= (1-\lambda) \det \begin{pmatrix} 2-\lambda & 3 \\ 4 & 4-\lambda \end{pmatrix} - 2 \det \begin{pmatrix} 2 & 3 \\ 3 & 4-\lambda \end{pmatrix} \\ &+ 1 \cdot \det \begin{pmatrix} 2 & 2-\lambda \\ 3 & 4 \end{pmatrix} = (1-\lambda)[(2-\lambda)(4-\lambda)-12] - 2[2(4-\lambda)-9] \\ &+ [8-3(2-\lambda)] = -\lambda^3 + 7\lambda^2 + 5\lambda = 0. \end{aligned}$$

The roots of this cubic polynomial are

$$\lambda_1 = 0, \lambda_2 = \frac{7 + \sqrt{69}}{2}, \text{ and } \lambda_3 = \frac{7 - \sqrt{69}}{2}.$$

4.6 Exercises

1. Evaluate the following determinants:

a) $\det \begin{pmatrix} a+b & a-b \\ a-b & a+b \end{pmatrix}$

b) $\det \begin{pmatrix} 1 & 5 & 2 \\ 4 & 7 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

c) $\det \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

d) $\det \begin{pmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{pmatrix}$

by using the definition of determinants.

2. Solve the previous problem by using cofactor expansions.

3. Show that with distinct real numbers a_1, a_2, \dots, a_n ,

$$\det \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix} \neq 0.$$

4. Solve equation

$$\det \begin{pmatrix} x & 2 & 1 \\ x & 1 & 2 \\ x & 1 & 3 \end{pmatrix} = 1.$$

5. Evaluate

$$\det \begin{pmatrix} x & 1 & 2 \\ 0 & y & 3 \\ 0 & 0 & z \end{pmatrix}.$$

6. Find

$$\det \begin{pmatrix} a & b & 0 & 0 \\ b & a & b & 0 \\ 0 & b & a & b \\ 0 & 0 & b & a \end{pmatrix}.$$

7. Is $\det(\underline{A} + \underline{B}) = \det(\underline{A}) + \det(\underline{B})$?

8. Is $\det(a\underline{A}) = a \det(\underline{A})$?

9. Prove that if two columns of a square matrix \underline{A} are proportional, then $\det(\underline{A}) = 0$.

10. Prove that the converse of the statement of the previous problem is false.

11. Use determinants to solve the following systems of linear equations:

a) $x + y = 2$

$$x - y = 3$$

b) $2x - y = 2$

$$x + y = 1$$

c) $x + y + z = 2$

$$-y + 2z = -1$$

$$2x + 2y = 2$$

d) $x + 2y + z = 1$

$$x - 2y + z = 1$$

$$x + y - z = -1$$

12. Find the inverses of the coefficient matrices of the previous problem by solving the matrix equation $\underline{A}\underline{X} = \underline{I}$ column-wise, and solving each system by the Cramer's rule.

13. Let \underline{A} be a block-diagonal matrix:

$$\underline{A} = \begin{pmatrix} \underline{A}_1 & & & \mathbf{O} \\ & \underline{A}_2 & & \\ & & \ddots & \\ \mathbf{O} & & & \underline{A}_k \end{pmatrix},$$

where all blocks \underline{A}_i ($i = 1, 2, \dots, k$) are square matrices. Prove that $\det(\underline{A}) = \det(\underline{A}_1) \cdot \det(\underline{A}_2) \cdot \dots \cdot \det(\underline{A}_k)$.

14. Assume that for an $n \times n$ real matrix, $\underline{A} \cdot \underline{A}^T = \underline{I}_n$ (such matrices are called orthogonal). Prove that $\det(\underline{A}) = \pm 1$.

15. Assume that for an $n \times n$ real matrix, $\underline{A}^k = \underline{0}$ (such matrices are called nilpotent). Prove that $\det(\underline{A}) = 0$.

16. Check the determinant multiplication theorem for 2×2 real matrices.

17. Apply the determinant multiplication theorem to prove that $\det(\alpha \underline{A}) = \alpha^n \det(\underline{A})$, if \underline{A} is an $n \times n$ real matrix and α is a real scalar.

18. Illustrate Theorem 4.1 and parts (a), (c), (e), and (f) of Theorem 4.2 with matrices

$$\underline{A} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \quad \underline{B} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix}.$$

19. Illustrate relations (4.8) and (4.9) for matrix \underline{A} of the previous problem.

20. Check the conditions of Theorem 4.6 for the linear systems of Problem 11.

21. Find the rank of the following matrices by using the definition of ranks and also by using Theorem 4.8:

a)
$$\begin{pmatrix} 1 & 2 & 2 & 3 \\ 2 & 1 & 1 & 0 \end{pmatrix}$$

b)
$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 0 & 1 & 4 \end{pmatrix}$$

c)
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 4 & 3 & 2 & 1 \end{pmatrix}.$$

22. Repeat Example 4.9 with matrices

a)
$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 2 & 1 & 7 \end{pmatrix}$$

b)
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ -1 & -3 & -6 \end{pmatrix}$$

$$c) \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 9 \\ 4 & 3 & 2 \end{pmatrix}.$$

Compare your result with those of Problem 3/18.

23. If \underline{A} is an $n \times n$ matrix, and A_{ij} denotes its (i, j) cofactor, then show that relation (4.11) implies that

$$\det \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} = [\det(\underline{A})]^{n-1}.$$

24. Illustrate the relation of the previous problem for matrices of Problem 22.

25. Examine the complete controllability of system

$$\underline{x}(t+1) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} u(t).$$

Chapter 5

Linear Mappings and Matrices

5.1 Introduction

The interrelation of different mathematical systems of similar types is usually studied by examining some mappings from one system to another. Such mappings are especially important to be examined that preserves some important characteristics of the systems. In earlier chapters the linear structure of vector spaces played the most significant role, therefore mappings preserving the linear structure will be the most important for future studies. This chapter is devoted to the examination of such mappings.

5.2 Vector Coordinates

Assume that V is a finitely generated vector space and let $B = \{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n\}$ be a basis in V . From Theorem 2.7 we know that an arbitrary vector $\underline{x} \in V$ can be uniquely represented as the linear combination of the basis vectors. That is, there are unique scalars x_1, \dots, x_n such that

$$\underline{x} = x_1 \underline{u}_1 + x_2 \underline{u}_2 + \dots + x_n \underline{u}_n. \quad (5.1)$$

The coefficients x_1, \dots, x_n are called the *coordinates* of \underline{x} in basis B . We can therefore identify vector \underline{x} with the n -element column vector

$$\underline{x}_B = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix},$$

which is called the *coordinate-vector* of \underline{x} in basis B . Notice that in the case when a different basis is selected, the coefficients x_i as well as the coordinate-vector \underline{x}_B may become different. This observation is illustrated in the following example.

Example 5.1. Select first the natural basis $B = \{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ in R^3 , and let

$$\underline{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then obviously,

$$\underline{x} = 1 \cdot \underline{e}_1 + 1 \cdot \underline{e}_2 + 1 \cdot \underline{e}_3,$$

therefore,

$$\underline{x}_B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Assume next that the basis $B' = \{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$ is selected, with

$$\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Then $\underline{x} = \underline{u}_1$, that is, $\underline{x} = 1 \cdot \underline{u}_1 + 0 \cdot \underline{u}_2 + 0 \cdot \underline{u}_3$, consequently,

$$\underline{x}_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

♦

The uniqueness of the scalars x_1, \dots, x_n in representation (5.1) implies that mapping $\underline{x} \mapsto \underline{x}_B$ is one-to-one and it maps V into the space of n -element real (or complex) vectors. We will next show that the mapping preserves the linear structures of these vector spaces by showing that for arbitrary $\underline{x}, \underline{y} \in V$,

$$(\underline{x} + \underline{y})_B = \underline{x}_B + \underline{y}_B, \quad (5.2)$$

and for arbitrary $\underline{x} \in V$ and scalar a ,

$$(a \cdot \underline{x})_B = a \cdot \underline{x}_B. \quad (5.3)$$

The first relation can be proven as follows. Assume that $\underline{x}_B = (x_i)$ and $\underline{y}_B = (y_i)$, then

$$\underline{x} = x_1 \underline{u}_1 + \dots + x_n \underline{u}_n \quad \text{and} \quad \underline{y} = y_1 \underline{u}_1 + \dots + y_n \underline{u}_n,$$

and so,

$$\underline{x} + \underline{y} = (x_1 + y_1)\underline{u}_1 + \dots + (x_n + y_n)\underline{u}_n.$$

Hence, the coordinates of $\underline{x} + \underline{y}$ in basis B are $x_1 + y_1, \dots, x_n + y_n$. The second relation can be proved in similar way. If $\underline{x} = (x_i)$, then

$$a\underline{x} = a(x_1\underline{u}_1 + \dots + x_n\underline{u}_n) = (ax_1)\underline{u}_1 + \dots + (ax_n)\underline{u}_n.$$

That is, the coordinates of \underline{x} in basis B are ax_1, \dots, ax_n , which proves the assertion. Equations (5.2) and (5.3) show that there is a special relation between V and the vector space of n -dimensional real (or complex) vectors. In the next section this concept will be introduced in general and examined in detail.

5.3 Linear Mappings

Let V and V' be two (not necessarily different) vector spaces, and assume that either both are real or both are complex, furthermore mapping A maps V into V' . That is, $A(\underline{x})$ is defined for all $\underline{x} \in V$, and the images $A(\underline{x})$ are in V' for all $\underline{x} \in V$. This fact is denoted as $A: V \mapsto V'$. It is not required that each element of V' can be obtained as an image $A(\underline{x})$ with some $\underline{x} \in V$. The set of all images is denoted by $R(A)$ and called the *range* of mapping A :

$$R(A) = \{\underline{x}' \in V' \mid \text{there exists an } \underline{x} \in V \text{ such that } \underline{x}' = A(\underline{x})\}.$$

In this section a special class of mappings between vector spaces will be discussed.

Definition 5.1. Mapping $A:V \mapsto V'$ is called a *linear mapping*, if for all $\underline{x}, \underline{y} \in V$ and scalars a ,

$$(i) \ A(\underline{x} + \underline{y}) = A(\underline{x}) + A(\underline{y});$$

$$(ii) \ A(a\underline{x}) = a \cdot A(\underline{x}).$$

Notice that in property (ii), elements of V as well as elements of V' are multiplied by scalar a . This is the reason why we assumed that either both vector spaces are real (that is, only real numbers are considered as scalars), or both are complex (when complex numbers are considered to be scalars).

Notice that the mapping $\underline{x} \mapsto \underline{x}_B$ satisfies conditions (i) and (ii) which are equivalent to requirements (5.2) and (5.3). Thus, it is a linear mapping. As it was mentioned earlier, the range of a linear mapping does not need to be the entire space V' . For example, if one defines $A(\underline{x}) = \underline{0}'$ for all $\underline{x} \in V$, where $\underline{0}'$ is the zero element of V' , then this is a special linear mapping, the range of which consists of only one element. In all cases, the range of a linear mapping has to be a subspace of V' , as it is stated in the following theorem.

Theorem 5.1. Let $A:V \mapsto V'$ be a linear mapping. Then $R(A)$ is a subspace of V' .

Proof. By applying Theorem 2.2 we have to show that if \underline{x}' and \underline{y}' are in $R(A)$, then with arbitrary scalars a and b , $a\underline{x}' + b\underline{y}' \in R(A)$. Since \underline{x}' and \underline{y}' are in $R(A)$, there exist elements \underline{x} and \underline{y} in V such that $\underline{x}' = A(\underline{x})$ and $\underline{y}' = A(\underline{y})$. Since V is a vector space, $a\underline{x} + b\underline{y} \in V$, and the linearity of mapping A implies that

$$A(\underline{ax} + \underline{by}) = A(\underline{ax}) + A(\underline{by}) = aA(\underline{x}) + bA(\underline{y}) = \underline{ax}' + \underline{by}' \in R(A).$$

+

In verifying that a mapping is linear we have to check if conditions (i) and (ii) are satisfied. The following result gives only one condition as the sufficient and necessary condition for a mapping to be linear. So, in practical cases only one condition has to be checked.

Theorem 5.2. A mapping $A: V \rightarrow V'$ is linear if and only if for all $\underline{x}, \underline{y} \in V$ and scalars a and b ,

$$A(\underline{ax} + \underline{by}) = a \cdot A(\underline{x}) + b \cdot A(\underline{y}). \quad (5.4)$$

Proof. Assume first that conditions (i) and (ii) hold. Then

$$A(\underline{ax} + \underline{by}) = A(\underline{ax}) + A(\underline{by}) = a \cdot A(\underline{x}) + b \cdot A(\underline{y}),$$

so (5.4) is satisfied. Suppose next that equation (5.4) holds for all $\underline{x}, \underline{y}, a$ and b . First we show that $A(\underline{0}) = \underline{0}'$, where $\underline{0}$ and $\underline{0}'$ are the zero elements in V and V' , respectively.

Select $\underline{x} = \underline{y} = \underline{0}$ and $a = b = 1$. Then

$$A(\underline{0}) = A(\underline{0} + \underline{0}) = A(\underline{0}) + A(\underline{0})$$

and adding $-A(\underline{0})$ to both sides we have

$$\underline{0}' = A(\underline{0}).$$

Select next $a = b = 1$. Then

$$A(\underline{x} + \underline{y}) = A(1 \cdot \underline{x} + 1 \cdot \underline{y}) = 1 \cdot A(\underline{x}) + 1 \cdot A(\underline{y}) = A(\underline{x}) + A(\underline{y})$$

that is, condition (i) is satisfied. Select next $\underline{y} = \underline{0}$, then

$$A(a\underline{x}) = A(a\underline{x} + b\underline{0}) = aA(\underline{x}) + bA(\underline{0}) = aA(\underline{x}) + b\underline{0}' = aA(\underline{x}) + \underline{0}' = aA(\underline{x}).$$

Hence condition (ii) also holds showing that A is a linear mapping. ♣

In the special case when $R(A) = V'$, we say that A is a mapping of V onto V' . A special but important class of such linear mappings is defined next.

Definition 5.2. Let A be a one-to-one linear mapping of V onto V' . Then A is called an *isomorphism*, and vector spaces V and V' are called *isomorphic*.

We will first show that finitely generated vector spaces are isomorphic if and only if they have the same dimension.

Theorem 5.3. Let V and V' be vector spaces, and assume that V is finitely generated. Then they are isomorphic if and only if V' is also finitely generated and $\dim(V) = \dim(V')$.

Proof. We first show that an isomorphism preserves linear independence and linear dependence by proving that the images of linearly independent elements are also linearly independent, and the images of linearly dependent elements is also linearly dependent. It is sufficient to show that elements $\underline{x}_1, \dots, \underline{x}_k \in V$ are linearly dependent if and only if $A(\underline{x}_1), \dots, A(\underline{x}_k)$ are linearly dependent in V' . Assume first that $\underline{x}_1, \dots, \underline{x}_k$ are linearly dependent, then with some scalars,

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k = \underline{0},$$

where at least one coefficient is nonzero. Then

$$\begin{aligned} a_1 A(\underline{x}_1) + \dots + a_k A(\underline{x}_k) &= A(a_1 \underline{x}_1) + \dots + A(a_k \underline{x}_k) = A(a_1 \underline{x}_1 + \dots + a_k \underline{x}_k) \\ &= A(\underline{0}) = \underline{0}', \end{aligned}$$

therefore $A(\underline{x}_1), \dots, A(\underline{x}_k)$ are also linearly dependent. Assume next that vectors $A(\underline{x}_1), \dots, A(\underline{x}_k)$ are linearly dependent, then with some scalars

$$c_1 A(\underline{x}_1) + \dots + c_k A(\underline{x}_k) = \underline{0}'$$

and at least one coefficient is nonzero. This equation implies that

$$\begin{aligned} A(c_1 \underline{x}_1 + \dots + c_k \underline{x}_k) &= A(c_1 \underline{x}_1) + \dots + A(c_k \underline{x}_k) = c_1 A(\underline{x}_1) + \dots + c_k A(\underline{x}_k) \\ &= \underline{0}'. \end{aligned}$$

Since A is one-to-one and $A(\underline{0}) = \underline{0}'$, necessarily

$$c_1 \underline{x}_1 + \dots + c_k \underline{x}_k = \underline{0}$$

showing that elements $\underline{x}_1, \dots, \underline{x}_k$ are linearly dependent.

The above observation immediately implies that if V and V' are isomorphic and V is finitely generated, then the same is true for V' , and $\dim(V') = \dim(V)$. We will next prove that if V and V' are finitely generated vector spaces and $\dim(V') = \dim(V)$, then they are isomorphic.

Let $\{\underline{x}_1, \dots, \underline{x}_n\}$ be a basis of V and let $\{\underline{z}_1, \dots, \underline{z}_n\}$ be a basis in V' . If \underline{x} is an arbitrary element of V , then \underline{x} can be written as

$$\underline{x} = a_1 \underline{x}_1 + \dots + a_n \underline{x}_n$$

with some scalars a_i ($i = 1, 2, \dots, n$). Define a mapping $A: V \mapsto V'$ as

$$A(\underline{x}) = a_1 \underline{z}_1 + \dots + a_n \underline{z}_n.$$

In order to show that mapping A is an isomorphism between V and V' we have to show that A is a one-to-one linear mapping from V onto V' . It is clear that mapping A is defined for all $\underline{x} \in V$, since arbitrary element of V can be expressed as a linear combination of the basis elements. To show that A is one-to-one assume that for some \underline{x} and \underline{y} , $A(\underline{x}) = A(\underline{y})$. With some scalars a_1, \dots, a_n and b_1, \dots, b_n ,

$$\underline{x} = a_1 \underline{x}_1 + \dots + a_n \underline{x}_n \quad \text{and} \quad \underline{y} = b_1 \underline{x}_1 + \dots + b_n \underline{x}_n.$$

Then

$$A(\underline{x}) = a_1 \underline{z}_1 + \dots + a_n \underline{z}_n \quad \text{and} \quad A(\underline{y}) = b_1 \underline{z}_1 + \dots + b_n \underline{z}_n,$$

and since $A(\underline{x}) = A(\underline{y})$,

$$a_1 \underline{z}_1 + \dots + a_n \underline{z}_n = b_1 \underline{z}_1 + \dots + b_n \underline{z}_n.$$

That is,

$$(a_1 - b_1) \underline{z}_1 + \dots + (a_n - b_n) \underline{z}_n = \underline{0}'.$$

The linear independence of the basis elements implies that $a_1 = b_1, \dots, a_n = b_n$, consequently $\underline{x} = \underline{y}$, and therefore A is one-to-one. The linearity of mapping A can be proved by using Theorem 5.2. Let a and b be two scalars. Then

$$a\underline{x} + b\underline{y} = a(a_1\underline{x}_1 + \dots + a_n\underline{x}_n) + b(b_1\underline{x}_1 + \dots + b_n\underline{x}_n) = (aa_1 + bb_1)\underline{x}_1 + \dots + (aa_n + bb_n)\underline{x}_n,$$

therefore the definition of mapping A implies that

$$\begin{aligned} A(a\underline{x} + b\underline{y}) &= (aa_1 + bb_1)\underline{z}_1 + \dots + (aa_n + bb_n)\underline{z}_n \\ &= a(a_1\underline{z}_1 + \dots + a_n\underline{z}_n) + b(b_1\underline{z}_1 + \dots + b_n\underline{z}_n) = aA(\underline{x}) + bA(\underline{y}) \end{aligned}$$

We will finally show that $R(A) = V'$, that is A is a mapping from V onto V' . Let $\underline{z} \in V'$ be arbitrary, then with some scalars c_i ($i = 1, 2, \dots, n$),

$$\underline{z} = c_1\underline{z}_1 + \dots + c_n\underline{z}_n.$$

Consider the element $\underline{x} = c_1\underline{x}_1 + \dots + c_n\underline{x}_n$. Then the definition of A implies that $A(\underline{x}) = c_1\underline{z}_1 + \dots + c_n\underline{z}_n = \underline{z}$, therefore $\underline{z} \in R(A)$, which completes the proof. \clubsuit

Some special linear mappings are examined in the next few examples.

Example 5.2. Select $V = R^n$ and $V' = R^m$, and let \underline{A} be an $m \times n$ real matrix. For all $\underline{x} \in V$, define mapping A as $A(\underline{x}) = \underline{A} \cdot \underline{x}$. It is clear that $A(\underline{x})$ is defined for all $\underline{x} \in V$, since the product $\underline{A} \cdot \underline{x}$ exists for all $\underline{x} \in R^n$. Notice that $R(A)$ is not necessarily the entire space R^m . For example, if \underline{A} is the zero matrix, then $\underline{A} \cdot \underline{x} = \underline{0}$ for all $\underline{x} \in V$, therefore in

this case, $R(A) = \{0\}$. Mapping A is linear, since for all $\underline{x}, \underline{y} \in V$ and scalars a and b ,

$$\begin{aligned} A(\underline{ax} + \underline{by}) &= \underline{A} \cdot (\underline{ax} + \underline{by}) \\ &= \underline{A} \cdot (\underline{ax}) + \underline{A} \cdot (\underline{by}) = a\underline{Ax} + b\underline{Ay} = aA(\underline{x}) + bA(\underline{y}). \end{aligned}$$

Next we show that $R(A) = V'$ if and only if $\text{rank}(\underline{A}) = m$. This is the consequence of the simple fact that if $\underline{x} = (x_i)$, then

$$A(\underline{x}) = \underline{A} \cdot \underline{x} = x_1 \underline{a}_1 + \dots + x_n \underline{a}_n$$

where $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$ are the columns of A . Therefore $R(A)$ coincides with the set of all linear combinations of the columns of \underline{A} , which is the subspace generated by the columns of \underline{A} . Since the dimension of this subspace equals $\text{rank}(\underline{A})$,

$$\dim(R(A)) = \text{rank}(\underline{A}).$$

Hence, $R(A) = V'$ if and only if $\dim(R(A)) = m$, which is equivalent to the condition that $\text{rank}(\underline{A}) = m$.

Assume next that \underline{A} is $n \times n$ and nonsingular. Then $\text{rank}(\underline{A}) = n$, therefore $R(A) = R^n = V$. That is, A is a mapping of R^n onto itself. It can also be shown that the mapping is one-to-one. In contrary to this assertion assume that with some $\underline{x}, \underline{y} \in R^n$, $A(\underline{x}) = A(\underline{y})$. Then $\underline{A} \cdot \underline{x} = \underline{A} \cdot \underline{y}$, which is equivalent to equality

$$\underline{A} \cdot (\underline{x} - \underline{y}) = \underline{0}.$$

Multiply both sides from the left by \underline{A}^{-1} to have

$$\underline{x} - \underline{y} = \underline{0},$$

since $\underline{A}^{-1}\underline{A} = \underline{I}$ and $\underline{A}^{-1}\underline{0} = \underline{0}$. Thus, $\underline{x} = \underline{y}$.

Let $a \neq 0$ be a scalar, and select the special $n \times n$ nonsingular matrix $\underline{A} = a \cdot \underline{I}$, where \underline{I} is the $n \times n$ identity matrix. In this special case,

$$A(\underline{x}) = \underline{A} \cdot \underline{x} = a \cdot \underline{I}\underline{x} = a \cdot \underline{x},$$

that is, the mapping multiplies each vector by scalar a .

◆

Example 5.3. Select again $V = R^n$ and let V' be the set of real numbers. Let the real numbers a_1, \dots, a_n be given. If $\underline{x} \in V$ is an arbitrary vector, then define

$$A(\underline{x}) = a_1x_1 + \dots + a_nx_n$$

where x_1, \dots, x_n are the elements of \underline{x} . This mapping is a special case of the one examined in the previous example with $m = 1$. Notice that $A(\underline{x})$ is a real number. In the literature such linear mappings are called *linear functionals*.

◆

Example 5.4. Let V be the set of all single variable real polynomials of degree not exceeding n . In Example 2.3 we verified that V is a vector space. If $p \in V$, then define mapping A as taking the derivative of p :

$$A(p) = \frac{d}{dt} p.$$

Notice first that $A(p)$ is a polynomial of degree not exceeding $n-1$. Therefore define V' as the set of all real polynomials of degree at most $n-1$. Then A is a linear mapping of V onto V' . The linearity of the mapping follows from the elementary properties of differentiation:

$$A(ap + bq) = \frac{d}{dt}(ap + bq) = a \frac{d}{dt} p + b \frac{d}{dt} q$$

for all $p, q \in V$ and scalars a and b . Let

$$r(t) = c_0 + c_1 t + \dots + c_{n-1} t^{n-1} \in V'$$

be arbitrary. Define

$$p(t) = C + c_0 t + \frac{c_1}{2} t^2 + \frac{c_2}{3} t^3 + \dots + \frac{c_{n-1}}{n} t^n$$

with arbitrary constant term C . Simple differentiation shows that

$$A(p) = \frac{d}{dt} p = r.$$

Since C is not unique, the mapping is not one-to-one.



Example 5.5. Let V be the set $C[a, b]$ of all continuous real functions defined on a finite, closed interval $[a, b]$. Define now mapping A as integrating the functions on interval $[a, b]$. That is, for all $f \in C[a, b]$,

$$A(f) = \int_a^b f(t) dt.$$

Since $A(f)$ is a real number, select $V' = R$. Mapping A is a linear mapping of $C[a, b]$ onto R . The linearity follows immediately from the linearity of the integral, since for all $f, g \in C[a, b]$ and real constants c and d ,

$$A(cf + dg) = \int_a^b (cf(t) + dg(t)) dt = c \int_a^b f(t) dt + d \int_a^b g(t) dt = cA(f) + dA(g).$$

The fact that the range of this mapping is the entire real line follows from the observation, that if for any arbitrary real number c we select

$$p(t) = \frac{c}{b-a} \text{ for } t \in [a, b], \text{ then}$$

$$A(p) = \int_a^b \frac{c}{b-a} dt = c.$$

◆

We will next give a simple sufficient and necessary condition that a linear mapping is one-to-one. The condition will be based on the following concept.

Definition 5.3. Let $A:V \mapsto V'$ be a linear mapping. The *nullspace* of A is defined as

$$N(A) = \{\underline{x} \in V \mid A(\underline{x}) = \underline{0}'\}.$$

First we show that $N(A)$ is always a subspace in V .

Theorem 5.4. Let $A:V \mapsto V'$ be a linear mapping. Then $N(A)$ is a subspace of V .

Proof. We will show that $N(A)$ satisfies the condition of Theorem 2.2. Let $\underline{x}, \underline{y} \in N(A)$ and a, b be scalars. Then

$$A(a\underline{x} + b\underline{y}) = aA(\underline{x}) + bA(\underline{y}) = a \cdot \underline{0}' + b \cdot \underline{0}' = \underline{0}',$$

therefore $a\underline{x} + b\underline{y} \in N(A)$. Hence $N(A)$ is a subspace.

♣

We are now ready to present the simple condition that guarantees that a linear mapping is one-to-one.

Theorem 5.5. A linear mapping A is one-to-one if and only if $N(A) = \{\underline{0}\}$.

Proof. Assume first that $N(A)$ consists of only the zero element. Assume furthermore that with some \underline{x} and $\underline{y} \in V$, $A(\underline{x}) = A(\underline{y})$. Then

$$\underline{0}' = A(\underline{x}) - A(\underline{y}) = 1 \cdot A(\underline{x}) + (-1) \cdot A(\underline{y}) = A(1 \cdot \underline{x} + (-1) \cdot \underline{y}) = A(\underline{x} - \underline{y}),$$

therefore $\underline{x} - \underline{y} \in N(A)$. Consequently $\underline{x} - \underline{y} = \underline{0}$, that is, $\underline{x} = \underline{y}$ showing that mapping A is one-to-one.

Assume next that A is one-to-one. Let $\underline{z} \in N(A)$ be arbitrary, then $A(\underline{z}) = \underline{0}'$. Since $A(\underline{0}) = \underline{0}'$ (from the proof of Theorem 5.2) and A is one-to-one, $\underline{z} = \underline{0}$. Hence $N(A) = \{\underline{0}\}$.

Assume next that V is a finitely generated vector space. If a linear mapping $A: V \mapsto V'$ is one-to-one, then V and $R(A)$ are isomorphic. If $\dim(V) = n$, then $\dim(R(A)) = n$. Since $N(A)$ consists of only the zero element, $\dim(N(A)) = 0$. We may therefore notice that in this case, $\dim(V) = \dim(R(A)) + \dim(N(A))$. As we will next show, this relation always holds regardless of the fact that the mapping is one-to-one or not.

Theorem 5.6. Let $A: V \mapsto V'$ be a linear mapping, and assume that V is finitely generated. Then

$$\dim(V) = \dim(R(A)) + \dim(N(A)).$$

Proof. Let $\{\underline{x}_1, \dots, \underline{x}_k\}$ be a basis in $N(A)$ and assume that $\{A(\underline{y}_1), \dots, A(\underline{y}_l)\}$ is a basis in $R(A)$. In order to prove the assertion, it is sufficient to show that system $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l\}$ is a basis of V . To do so, we have to prove two things. First we have to verify that the system consists of linearly independent elements, and second, we have to show that the system generates V .

To prove linear independence assume that

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k + b_1 \underline{y}_1 + \dots + b_l \underline{y}_l = \underline{0} \quad (5.5)$$

with some scalars a_i and b_j ($i = 1, 2, \dots, k; j = 1, 2, \dots, l$). Apply mapping A to both sides of the equality. The linearity of A implies that

$$a_1 A(\underline{x}_1) + \dots + a_k A(\underline{x}_k) + b_1 A(\underline{y}_1) + \dots + b_l A(\underline{y}_l) = \underline{0}',$$

where $\underline{0}'$ is the zero element of V' . Since \underline{x}_i ($i = 1, 2, \dots, k$) is in the null-space of A , $A(\underline{x}_i) = \underline{0}'$ for all i . Therefore

$$b_1 A(\underline{y}_1) + \dots + b_l A(\underline{y}_l) = \underline{0}'.$$

Notice that $\{ A(\underline{y}_1), \dots, A(\underline{y}_l) \}$ is a basis of $R(A)$, and the linear combination of the basis elements is zero only if all coefficients are equal to zero: $b_1 = \dots = b_l = 0$. Then from (5.5) we have

$$a_1 \underline{x}_1 + \dots + a_k \underline{x}_k = \underline{0},$$

and by using the fact that $\{\underline{x}_1, \dots, \underline{x}_k\}$ is a basis in $N(A)$, we conclude that $a_1 = \dots = a_k = 0$. (Since a linear combination of basis elements is zero only if all coefficients equal zero.) Hence, all coefficients in (5.5) are equal to zero implying that vectors $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l$ are linearly independent.

We have to show next that this system generates V . Let $\underline{v} \in V$ be an arbitrary element. Since $A(\underline{v}) \in R(A)$ and $\{ A(\underline{y}_1), \dots, A(\underline{y}_l) \}$ is a basis of $R(A)$,

$$A(\underline{v}) = c_1 A(\underline{y}_1) + \dots + c_l A(\underline{y}_l)$$

with some scalars c_1, \dots, c_l . The linearity of A implies that

$$A(\underline{v} - (c_1 \underline{y}_1 + \dots + c_l \underline{y}_l)) = \underline{0}',$$

from which we see that

$$\underline{v} - (c_1 \underline{y}_1 + \dots + c_l \underline{y}_l) \in N(A).$$

Since $\{\underline{x}_1, \dots, \underline{x}_k\}$ is a basis of $N(A)$, this vector can be expressed as a linear combination of the basis elements:

$$\underline{v} - (c_1 \underline{y}_1 + \dots + c_l \underline{y}_l) = d_1 \underline{x}_1 + \dots + d_k \underline{x}_k$$

with some scalars d_1, \dots, d_k . That is,

$$\underline{v} = d_1 \underline{x}_1 + \dots + d_k \underline{x}_k + c_1 \underline{y}_1 + \dots + c_l \underline{y}_l,$$

which shows that \underline{v} is a linear combination of elements $\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l$. Therefore system $\{\underline{x}_1, \dots, \underline{x}_k, \underline{y}_1, \dots, \underline{y}_l\}$ generates V . This fact and the linear independence of these vectors imply that the system is a basis in V .

✱

Corollary. Assume that V is finitely generated. A linear mapping $A: V \mapsto V'$ is one-to-one if and only if V and $R(A)$ are isomorphic, which holds if and only if $\dim(V) = \dim R(A)$. From the assertion of the theorem we know that this relation holds if and only if $\dim(N(A)) = 0$ which is equivalent to the condition that $N(A)$ consists of only the zero vector. This observation shows that Theorem 5.5 is implied by Theorem 5.6.

Let $A:V \mapsto V'$ be a linear mapping. If A is one-to-one, then for all $\underline{z} \in R(A)$, there is a unique $\underline{x} \in V$ such that $A(\underline{x}) = \underline{z}$. If A is not one-to-one, then such an element \underline{x} is not necessarily unique. For all $\underline{z} \in R(A)$ denote

$$C(\underline{z}) = \{\underline{x} \in V \mid A(\underline{x}) = \underline{z}\}.$$

Set $C(\underline{z})$ contains all elements $\underline{x} \in V$ which have the same image \underline{z} in V' with respect to mapping A . Notice first that if \underline{x}_0 is an arbitrary element of $C(\underline{z})$, then

$$C(\underline{z}) = \{\underline{x}_0 + \underline{y} \mid \underline{y} \in N(A)\}, \quad (5.6)$$

that is, $C(\underline{z})$ is a linear manifold with directing space $N(A)$ (which follows from Definition 3.5). To show that relation (5.6) holds assume first that \underline{x} and \underline{x}_0 are in $C(\underline{z})$. Then $A(\underline{x}) = A(\underline{x}_0) = \underline{z}$, which implies that

$$\underline{0}' = A(\underline{x}) - A(\underline{x}_0) = A(\underline{x} - \underline{x}_0),$$

and so, $\underline{x} - \underline{x}_0 \in N(A)$. Assume next that $\underline{y} \in N(A)$ is arbitrary, then $A(\underline{x}_0 + \underline{y}) = A(\underline{x}_0) + A(\underline{y}) = \underline{z} + \underline{0}' = \underline{z}$, that is, $\underline{x}_0 + \underline{y} \in C(\underline{z})$.

From the definition of set $C(\underline{z})$ it is also obvious that for $\underline{z} \neq \underline{z}'$,

$$C(\underline{z}) \cap C(\underline{z}') = \{\emptyset\},$$

where $\{\emptyset\}$ denotes the empty set, since for any common element \underline{x} of $C(\underline{z})$ and $C(\underline{z}')$, the image $A(\underline{x})$ cannot be \underline{z} and \underline{z}' at the same time. Furthermore,

$$\bigcup_{\underline{z} \in R(A)} C(\underline{z}) = V,$$

since mapping A is defined for all elements of V . The last two observations can be summarized by saying that V is the union of the disjoint sets $C(\underline{z})$ for all $\underline{z} \in R(A)$. For simple reference, $C(\underline{z})$ will be called the *equivalence class associated to \underline{z}* .

A linear structure will be next defined on the set of the equivalence classes generated by a linear mapping A . The sum of classes $C(\underline{z})$ and $C(\underline{z}')$ is defined as the class $C(\underline{z} + \underline{z}')$ associated to $\underline{z} + \underline{z}'$, and for any scalar a , the a -multiple of $C(\underline{z})$ is defined as the class $C(a\underline{z})$ associated to $a\underline{z}$. We will prove that these two operations satisfy all conditions of Definition 2.1 implying that the equivalence classes form a vector space. Conditions (i) and (ii) follow from the commutativity and associativity of the addition defined in V . If the null element among the equivalence classes is defined as $C(\underline{0}')$, and $-C(\underline{z})$ is defined as $C(-\underline{z})$, then properties (iii) and (iv) are obviously satisfied. Relations (v), (vi), and (vii) follow from the fact that vector space V' satisfies the same conditions, and (viii) is implied by the simple fact that for all $\underline{z} \in R(A)$,

$$1 \cdot C(\underline{z}) = C(1 \cdot \underline{z}) = C(\underline{z}).$$

This vector space of the equivalence classes is called the *factor space of A* and denoted by V/A . Notice that C can also be considered as a mapping which maps $R(A)$ onto V/A . From the above observations it follows that C is linear and one-to-one, therefore is an isomorphism, and vector spaces $R(A)$ and V/A are isomorphic.

5.4 The Vector Space of Linear Mappings

Let V and V' be (not necessarily different) vector spaces. Assume that either both are real or both are complex. Let $L(V, V')$ denote the set of all linear mappings from V into V' . If $A, B \in L(V, V')$, then we can define the sum of mappings A and B as the mapping assigning the sum $A(\underline{x}) + B(\underline{x})$ to all $\underline{x} \in V$. That is, $\underline{A} + \underline{B}$ is defined as

$$(A + B)(\underline{x}) = A(\underline{x}) + B(\underline{x}) \quad (5.7)$$

for all $\underline{x} \in V$. We can similarly define the scalar-multiple of a linear mapping. The c -multiple of A is the mapping that assigns $cA(\underline{x})$ to every $\underline{x} \in V$, that is,

$$(cA)(\underline{x}) = cA(\underline{x}). \quad (5.8)$$

Theorem 5.7. Mappings $\underline{A} + \underline{B}$ and aA are linear.

Proof. We will apply Theorem 5.2. Let a and b be scalars and assume that \underline{x} and \underline{y} are arbitrary elements of V . Then the linearity of A and B implies that

$$\begin{aligned} (A + B)(a\underline{x} + b\underline{y}) &= A(a\underline{x} + b\underline{y}) + B(a\underline{x} + b\underline{y}) \\ &= aA(\underline{x}) + bA(\underline{y}) + aB(\underline{x}) + bB(\underline{y}) \\ &= a(A(\underline{x}) + B(\underline{x})) + b(A(\underline{y}) + B(\underline{y})) = a(A + B)(\underline{x}) + b(A + B)(\underline{y}), \end{aligned}$$

and

$$\begin{aligned} (cA)(a\underline{x} + b\underline{y}) &= c \cdot A(a\underline{x} + b\underline{y}) = c(aA(\underline{x}) + bA(\underline{y})) \\ &= a(cA(\underline{x})) + b(cA(\underline{y})) = a(cA)(\underline{x}) + b(cA)(\underline{y}) \end{aligned}$$

showing that both $\underline{A} + \underline{B}$ and cA are linear mappings.

✦

Next we will show that $L(V, V')$ is a vector space if addition and multiplication by scalars are defined as above. We will verify that all conditions of Definition 2.1 are satisfied:

(i) Since V' is a vector space, for all $\underline{x} \in V$,

$$(\underline{A} + \underline{B})(\underline{x}) = \underline{A}(\underline{x}) + \underline{B}(\underline{x}) = \underline{B}(\underline{x}) + \underline{A}(\underline{x}) = (\underline{B} + \underline{A})(\underline{x});$$

(ii) Similarly, for all $\underline{x} \in V$,

$$\begin{aligned} ((\underline{A} + \underline{B}) + \underline{C})(\underline{x}) &= (\underline{A} + \underline{B})(\underline{x}) + \underline{C}(\underline{x}) = (\underline{A}(\underline{x}) + \underline{B}(\underline{x})) + \underline{C}(\underline{x}) \\ &= \underline{A}(\underline{x}) + (\underline{B}(\underline{x}) + \underline{C}(\underline{x})) = \underline{A}(\underline{x}) + (\underline{B} + \underline{C})(\underline{x}) = (\underline{A} + (\underline{B} + \underline{C}))(\underline{x}); \end{aligned}$$

(iii) The null element \underline{O} in $L(V, V')$ is defined as $\underline{O}(\underline{x}) = \underline{0}'$ for all $\underline{x} \in V$. Then

$$(\underline{A} + \underline{O})(\underline{x}) = \underline{A}(\underline{x}) + \underline{O}(\underline{x}) = \underline{A}(\underline{x}) + \underline{0}' = \underline{A}(\underline{x});$$

(iv) Define $-\underline{A}$ as $(-1) \cdot \underline{A}$, then

$$\begin{aligned} (\underline{A} + (-\underline{A}))(\underline{x}) &= \underline{A}(\underline{x}) + (-\underline{A})(\underline{x}) = \underline{A}(\underline{x}) + (-1)\underline{A}(\underline{x}) \\ &= 1 \cdot \underline{A}(\underline{x}) + (-1) \cdot \underline{A}(\underline{x}) = (1 + (-1))\underline{A}(\underline{x}) = 0 \cdot \underline{A}(\underline{x}) = \underline{0}' = \underline{O}(\underline{x}) \end{aligned}$$

for all $\underline{x} \in V$;

Conditions (v) - (vii) follow from the observations that for all $\underline{x} \in V$,

$$\begin{aligned}
 (a(A+B))(\underline{x}) &= a(A+B)(\underline{x}) = a(A(\underline{x}) + B(\underline{x})) \\
 &= aA(\underline{x}) + aB(\underline{x}) = (aA)(\underline{x}) + (aB)(\underline{x}) = (aA + aB)(\underline{x}), \\
 ((a+b)A)(\underline{x}) &= (a+b)A(\underline{x}) = aA(\underline{x}) + bA(\underline{x}) \\
 &= (aA)(\underline{x}) + (bA)(\underline{x}) = (aA + bA)(\underline{x}),
 \end{aligned}$$

and

$$((ab)A)(\underline{x}) = (ab)A(\underline{x}) = a(bA(\underline{x})) = a(bA)(\underline{x}) = (a(bA))(\underline{x}).$$

And finally, notice that property (viii) is obvious, since for all $\underline{x} \in V$,

$$(1 \cdot A)(\underline{x}) = 1 \cdot A(\underline{x}) = A(\underline{x}).$$

Consider next the special case when V' is the set of real numbers if V is a real vector space, or V' is the set of complex numbers if V is a complex vector space. Then the vector space $L(V, V')$ is called the *dual space of V* , and is denoted by V^* .

5.5 Multiplication of Linear Mappings, and Inverses

Let V , V' , and V'' be vector spaces. It is assumed that either all are real or all are complex. Let $B \in L(V, V')$ and $A \in L(V', V'')$ be two linear mappings. The product of A and B is denoted by $A \cdot B$ and is defined as

$$(A \cdot B)(\underline{x}) = A(B(\underline{x}))$$

for all $\underline{x} \in V$. That is, mapping B maps \underline{x} into $B(\underline{x})$, and then this image is mapped by A into $A(B(\underline{x}))$. This definition is the same as the composition of functions A and B known from calculus, where the composition was denoted by $A \circ B$. In this book we will use the notation $A \cdot B$ to emphasize the algebraic structure of the set of linear mappings. Mapping $A \cdot B$ is defined on the entire space V and its range is a subset of V'' . Furthermore it is a linear mapping, since for all elements $\underline{x}, \underline{y} \in V$ and scalars a, b ,

$$\begin{aligned}(A \cdot B)(a\underline{x} + b\underline{y}) &= A(B(a\underline{x} + b\underline{y})) = A(a \cdot B(\underline{x}) + b \cdot B(\underline{y})) \\ &= a \cdot A(B(\underline{x})) + b \cdot A(B(\underline{y})) = a \cdot (A \cdot B)(\underline{x}) + b \cdot (A \cdot B)(\underline{y}).\end{aligned}$$

Then Theorem 5.2 implies that $A \cdot B \in L(V, V'')$.

The multiplication of linear mappings is not commutative in general, since if spaces V and V'' are different, then product $B \cdot A$ cannot be defined. Notice that A maps V' into V'' , and B maps V into V' , therefore mapping B , in general, cannot be applied to the element $A(\underline{z}) \in V''$. If $V = V''$, then both $A \cdot B$ and $B \cdot A$ are defined, but they are not necessarily equal. Such an example is given next.

Example 5.6. Let $V = V' = V'' = R^2$, and defined mappings A and B as

$$A(\underline{x}) = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \underline{x} \quad \text{and} \quad B(\underline{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \underline{x}.$$

In Example 5.2 we have seen that multiplication by a matrix is a linear mapping. If x_1 and x_2 denote the elements of \underline{x} , then

$$\begin{aligned}(A \cdot B)\underline{x} &= A(B(\underline{x})) = A\left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) \\ &= A\begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + 2x_2 \\ -2x_1 - 2x_2 \end{pmatrix},\end{aligned}$$

and

$$\begin{aligned}(B \cdot A)\underline{x} &= B(A(\underline{x})) = B\left(\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) \\ &= B\begin{pmatrix} x_1 + x_2 \\ -x_1 - x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 + x_2 \\ -x_1 - x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.\end{aligned}$$

Notice that $(A \cdot B)(\underline{x}) = (B \cdot A)(\underline{x})$ if and only if $x_1 + x_2 = 0$.

We will next show that the multiplication of mappings is associative and distributive. Let $C \in L(V, V')$, $B \in L(V', V'')$, $A \in L(V'', V''')$, where V''' is a vector space of the same type as V , V' , and V'' . Then for all $\underline{x} \in V$,

$$\begin{aligned}((A \cdot B) \cdot C)(\underline{x}) &= (A \cdot B)(C(\underline{x})) = A(B(C(\underline{x}))) \\ &= A((B \cdot C)(\underline{x})) = (A \cdot (B \cdot C))(\underline{x}),\end{aligned}$$

and so,

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Assume next that $A, B \in L(V', V'')$ and $C \in L(V, V')$. Then

$$\begin{aligned}((A + B)C)(\underline{x}) &= (A + B)(C(\underline{x})) = A(C(\underline{x})) + B(C(\underline{x})) \\ &= (AC)(\underline{x}) + (BC)(\underline{x}) = (AC + BC)(\underline{x}),\end{aligned}$$

that is,

$$(A + B)C = AC + BC.$$

If $A \in L(V', V'')$ and $B, C \in L(V, V')$, then similarly,

$$\begin{aligned} (A(B + C))(\underline{x}) &= A((B + C)(\underline{x})) = A(B(\underline{x}) + C(\underline{x})) \\ &= A(B(\underline{x})) + A(C(\underline{x})) = (AB)(\underline{x}) + (AC)(\underline{x}) = (AB + AC)(\underline{x}), \end{aligned}$$

which shows that

$$A(B + C) = AB + AC.$$

Consider next the special case when $V = V' = V''$. Instead of the general notation $L(V, V')$ we may now use the simplified notation $L(V)$ for the set of all linear mappings from V into itself. Such mappings are called *linear transformations* on V . If both A and B are linear transformations in $L(V)$, then $A \cdot B$ is also in $L(V)$. This simple observation gives the possibility to define polynomials of linear transformations. If $A \in L(V)$, then the square A^2 of A is defined as $A^2 = A \cdot A$, and in general, A^n is defined by the recursive relation $A^n = A \cdot A^{n-1}$. We also define $A^0 = I$ where I is the identity transformation for which $I(\underline{x}) = \underline{x}$ for all $\underline{x} \in V$. Let

$$p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n \tag{5.9}$$

be next a single variable polynomial. If V is a real vector space, then we assume that the polynomial coefficients are all real, and if V is complex, then all coefficients are assumed to be complex. Then $p(A)$ is defined as

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \dots + a_n A^n.$$

Powers and polynomials of linear transformations are illustrated in the following two examples.

Example 5.7. Assume that $V = R^n$ and \underline{A} is an $n \times n$ real matrix. Define mapping A as $A(\underline{x}) = \underline{A} \cdot \underline{x}$ for all n -element real vectors \underline{x} . Then

$$A^2(\underline{x}) = A(A(\underline{x})) = A(\underline{A} \cdot \underline{x}) = \underline{A} \cdot (\underline{A} \cdot \underline{x}) = \underline{A}^2 \cdot \underline{x},$$

where \underline{A}^2 is the square of matrix \underline{A} . By using finite induction we can prove that in general,

$$A^n(\underline{x}) = \underline{A}^n \cdot \underline{x}. \quad (5.10)$$

For $n = 1$ this relation is true. Assume that it holds for a positive integer n . Then it is true for $n + 1$, since

$$A^{n+1}(\underline{x}) = A(A^n(\underline{x})) = A(\underline{A}^n \cdot \underline{x}) = \underline{A} \cdot (\underline{A}^n \cdot \underline{x}) = \underline{A}^{n+1} \cdot \underline{x}.$$

Assume that a real polynomial is given by equation (5.9). Then

$$\begin{aligned} p(A)(\underline{x}) &= (a_0 I + a_1 A + a_2 A^2 + \dots + a_n A^n)(\underline{x}) \\ &= a_0 I(\underline{x}) + a_1 A(\underline{x}) + a_2 A^2(\underline{x}) + \dots + a_n A^n(\underline{x}) \\ &= a_0 \underline{I} \cdot \underline{x} + a_1 \underline{A} \cdot \underline{x} + a_2 \underline{A}^2 \cdot \underline{x} + \dots + a_n \underline{A}^n \cdot \underline{x} \\ &= (a_0 \underline{I} + a_1 \underline{A} + a_2 \underline{A}^2 + \dots + a_n \underline{A}^n) \cdot \underline{x}. \end{aligned}$$

Example 5.8. Select V as the set of all continuously differentiable real functions on an interval $[a, b]$. For all $f \in V$, define $A(f)$ as the derivative of f :

$$A(f) = \frac{d}{dt} f.$$

The square of mapping A cannot be defined on the entire V , since the selection of V implies that although $\frac{d}{dt} f$ is continuous, not necessarily differentiable. That is, mapping A cannot be necessarily applied again to the derivative function $\frac{d}{dt} f$.

Select V next as the set of all single variable real polynomials, then $A(f) = \frac{d}{dt} f$ is also a real polynomial for all $f \in V$. Therefore transformation A can be applied again, and

$$A^2(f) = \frac{d^2}{dt^2} f,$$

and in general

$$A^n(f) = \frac{d^n}{dt^n} f.$$

Before defining inverses of linear transformations we recall that the identity transformation I on an arbitrary vector space V was defined as

$$I(\underline{x}) = \underline{x}$$

for all $\underline{x} \in V$. If $A \in L(V)$ is any linear transformation, then from this definition it is clear that

$$A \cdot I = A \quad \text{and} \quad I \cdot A = A.$$

Let $A \in L(V)$ be a linear transformation on a vector space V . The inverse of A is defined as follows.

Definition 5.4. Linear transformation $A \in L(V)$ is called *invertible* if $R(A) = V$ and the transformation is one-to-one. The inverse A^{-1} of A is defined as for all $\underline{x} \in V$,

$$A^{-1}(\underline{x}) = \{\underline{y} \in V \mid \underline{x} = A(\underline{y})\}. \quad (5.11)$$

Notice first that $A^{-1}(\underline{x})$ has exactly one element, since A is one-to-one. Therefore A^{-1} is a mapping of V into itself. Using Theorem 5.2 we can easily prove that A^{-1} is also a linear transformation. Let a, b be two scalars and $\underline{x}, \underline{y} \in V$. Then

$$A^{-1}(a\underline{x} + b\underline{y}) = aA^{-1}(\underline{x}) + bA^{-1}(\underline{y}),$$

since

$$A(aA^{-1}(\underline{x}) + bA^{-1}(\underline{y})) = a \cdot A(A^{-1}(\underline{x})) + b \cdot A(A^{-1}(\underline{y})) = a\underline{x} + b\underline{y}.$$

Here we used the fact that for all $\underline{x} \in V$,

$$A(A^{-1}(\underline{x})) = \underline{x} \quad (5.12)$$

which follows from relation (5.11) by selecting $\underline{y} = A^{-1}(\underline{x})$.

Inverse transformations have the following properties.

Theorem 5.8.

a) Let $A \in L(V)$ be invertible. Then

$$(i) (A^{-1})^{-1} = A;$$

$$(ii) A \cdot A^{-1} = A^{-1} \cdot A = I;$$

b) If $A, B, C \in L(V)$, then

$$(iii) A \cdot B = C \cdot A = I \text{ implies that } A \text{ is invertible and } B = C = A^{-1};$$

c) If $A, B \in L(V)$ are invertible, then

$$(iv) A \cdot B \text{ is invertible and } (A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

Proof. (i) Transformation A^{-1} is invertible, since it satisfies the conditions of Definition 5.4. First, $R(A^{-1}) = V$, since A is defined on the entire space V . Next we show that A^{-1} is one-to-one. Assume that $A^{-1}(\underline{x}) = A^{-1}(\underline{y})$. Let \underline{z} denote this common element, then from (5.11) we conclude that

$$\underline{x} = A(\underline{z}) \quad \text{and} \quad \underline{y} = A(\underline{z}),$$

that is, $\underline{x} = \underline{y}$. Therefore $(A^{-1})^{-1}$ exists. Let $\underline{u} \in V$ be arbitrary, and denote

$$\underline{v} = (A^{-1})^{-1}(\underline{u}).$$

Denote A^{-1} by B , then

$$\underline{v} = B^{-1}(\underline{u}).$$

Applying B on both sides of this equality and using relation (5.12) we have

$$B(\underline{v}) = B(B^{-1}(\underline{u})) = \underline{u},$$

that is

$$A^{-1}(\underline{v}) = \underline{u}.$$

Apply next A on both sides and use relation (5.11) again to see that

$$\underline{v} = A(A^{-1}(\underline{v})) = A(\underline{u}),$$

that is,

$$(A^{-1})^{-1}(\underline{u}) = A(\underline{u}),$$

which is equivalent to the assertion.

(ii) The first part, $A \cdot A^{-1} = I$, is equivalent to equality (5.12), and the second part is a consequence of the first part by replacing A by A^{-1} and using property (i).

(iii) First we prove that A is invertible by showing that A is one-to-one and $R(A) = V$. Assume first that $A(\underline{x}) = A(\underline{y})$. Then

$$\underline{x} = I(\underline{x}) = (C \cdot A)\underline{x} = C(A(\underline{x})) = C(A(\underline{y})) = (C \cdot A)(\underline{y}) = I(\underline{y}) = \underline{y},$$

that is, A is one-to-one. Assume now that $\underline{x} \in V$ is an arbitrary element.

Then $B(\underline{x}) \in V$, and

$$A(B(\underline{x})) = (A \cdot B)(\underline{x}) = I(\underline{x}) = \underline{x},$$

therefore $\underline{x} \in R(A)$, and so $R(A) = V$.

We prove next that $B = C$. For all $\underline{x} \in V$,

$$\begin{aligned} C(\underline{x}) &= C(I(\underline{x})) = C((A \cdot B)(\underline{x})) = (C \cdot (A \cdot B))(\underline{x}) = ((C \cdot A) \cdot B)(\underline{x}) \\ &= (I \cdot B)(\underline{x}) = I(B(\underline{x})) = B(\underline{x}). \end{aligned}$$

And finally, we verify that $B = C = A^{-1}$. For all $\underline{x} \in V$, similarly to the previous calculation we have

$$\begin{aligned} C(\underline{x}) &= C(I(\underline{x})) = C((A \cdot A^{-1})(\underline{x})) = (C \cdot (A \cdot A^{-1}))(\underline{x}) = ((C \cdot A) \cdot A^{-1})(\underline{x}) \\ &= (I \cdot A^{-1})(\underline{x}) = I(A^{-1}(\underline{x})) = A^{-1}(\underline{x}). \end{aligned}$$

(iv) Simple calculation shows that

$$\begin{aligned} (B^{-1} \cdot A^{-1}) \cdot (A \cdot B) &= B^{-1} (A^{-1} \cdot (A \cdot B)) \\ &= B^{-1} ((A^{-1} \cdot A) \cdot B) = B^{-1} \cdot (I \cdot B) = B^{-1} \cdot B = I, \end{aligned}$$

similarly,

$$\begin{aligned} (A \cdot B) \cdot (B^{-1} \cdot A^{-1}) &= A \cdot (B \cdot (B^{-1} \cdot A^{-1})) \\ &= A \cdot ((B \cdot B^{-1}) \cdot A^{-1}) = A \cdot (I \cdot A^{-1}) = A \cdot A^{-1} = I. \end{aligned}$$

And finally, property (iii) implies the assertion. ♣

In the case of finitely generated subspaces property (iii) can be significantly simplified.

Theorem 5.9. Let V be a finitely generated vector space and $A \in L(V)$. Then the following are equivalent to each other:

- (a) Transformation A is invertible;
- (b) There is a $B \in L(V)$ such that $A \cdot B = I$;

(c) There is a $C \in L(V)$ such that $C \cdot A = I$.

Proof. Notice first that (a) implies (b) and (c), since in the case of invertible A , the selections $B = A^{-1}$ and $C = A^{-1}$ are suitable. Next we show that (b) implies (a) by verifying that condition (b) implies that $R(A) = V$ and A is one-to-one. Let $\underline{v} \in V$ be arbitrary. Then $B(\underline{v}) \in V$, and

$$A(B(\underline{v})) = (A \cdot B)(\underline{v}) = I(\underline{v}) = \underline{v}$$

which shows that $\underline{v} \in R(A)$. That is, $R(A) = V$. To prove that A is one-to-one assume that $A(\underline{x}) = A(\underline{y})$ with some $\underline{x}, \underline{y} \in V, \underline{x} \neq \underline{y}$. Then $A(\underline{x} - \underline{y}) = \underline{0}$, therefore $N(A)$ has at least one nonzero element, and so $\dim(N(A)) \geq 1$. Then Theorem 5.6 implies that $\dim(R(A)) < n$, where n is the dimension of V . Since $(A \cdot B)(\underline{x}) = A(B(\underline{x})) \in R(A)$, and $\dim(R(A \cdot B)) \leq \dim(R(A))$ therefore $\dim(R(A \cdot B)) < n$. Since $R(I) = V$ and $\dim(R(I)) = n$, $A \cdot B \neq I$, and this contradiction implies that A^{-1} exists, and relations $\underline{x} = A(B(\underline{x}))$ (for all $\underline{x} \in V$) and (5.11) imply that $B(\underline{x}) = A^{-1}(\underline{x})$ for all $\underline{x} \in V$, that is, $B = A^{-1}$.

The fact that condition (c) implies (a) can be proved in a similar way, the details are left as an exercise.

Remark. The assertion of the theorem does not hold if V is infinite dimensional, as it is illustrated in the following example. ♣

Example 5.9. Let V be the set of all single variable real polynomials. Let $p(t) = a_0 + a_1t + a_2t^2 + \dots + a_nt^n \in V$. Define

$$A(p)(t) = a_1 + 2a_2t + \dots + na_nt^{n-1} (= p'(t))$$

and

$$B(p)(t) = a_0 t + \frac{a_1 t^2}{2} + \frac{a_2 t^3}{3} + \dots + \frac{a_n t^{n+1}}{n+1} \left(= \int_0^t p(\tau) d\tau \right).$$

It is easy to see that

$$\begin{aligned} (A \cdot B)(p)(t) &= A(B(p))(t) \\ &= \frac{d}{dt} \left(a_0 t + \frac{a_1 t^2}{2} + \frac{a_2 t^3}{3} + \dots + \frac{a_n t^{n+1}}{n+1} \right) = p(t), \end{aligned}$$

that is, $A \cdot B = I$, however A^{-1} does not exist, since A is not one-to-one. This fact follows from the observation that for an arbitrary real number a ,

$$\frac{d}{dt}(p(t) + a) = \frac{d}{dt} p(t).$$

5.6 Matrix Representations of Linear Mappings

In this section we assume that V and V' are finitely generated vector spaces, and either both are real or both are complex. Suppose that $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ and $B' = \{\underline{v}_1, \dots, \underline{v}_m\}$ are bases in V and V' , respectively. Let $A \in L(V, V')$ be a linear mapping of V into V' . If $\underline{x} \in V$ is an arbitrary element, then

$$\underline{x} = x_1 \underline{u}_1 + \dots + x_n \underline{u}_n,$$

where the scalars x_1, \dots, x_n are the coordinates of \underline{x} in basis B . Then the linearity of mapping A implies that

$$A(\underline{x}) = x_1 A(\underline{u}_1) + \dots + x_n A(\underline{u}_n). \quad (5.13)$$

Notice that for all $i = 1, 2, \dots, n$, $A(\underline{u}_i) \in V'$, therefore this vector can be expressed as the linear combination of the basis elements of V' :

$$A(\underline{u}_i) = a_{i1} \underline{v}_1 + \dots + a_{mi} \underline{v}_m \quad (5.14)$$

with some scalars a_{ij}, \dots, a_{mi} . Substituting this equation into (5.13) we have

$$\begin{aligned} A(\underline{x}) &= x_1 (a_{11} \underline{v}_1 + \dots + a_{m1} \underline{v}_m) + \dots + x_n (a_{1n} \underline{v}_1 + \dots + a_{mn} \underline{v}_m) \\ &= (a_{11} x_1 + \dots + a_{1n} x_n) \underline{v}_1 + \dots + (a_{m1} x_1 + \dots + a_{mn} x_n) \underline{v}_m. \end{aligned}$$

This equation shows that the coordinates of the image element $A(\underline{x})$ in basis B' are

$$a_{11} x_1 + \dots + a_{1n} x_n, \dots, a_{m1} x_1 + \dots + a_{mn} x_n,$$

that is, its coordinate vector is the following:

$$A(\underline{x})_{B'} = \begin{pmatrix} a_{11} x_1 + \dots + a_{1n} x_n \\ \dots \\ a_{m1} x_1 + \dots + a_{mn} x_n \end{pmatrix}.$$

Introduce the $m \times n$ matrix

$$A_{B,B'} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

and coordinate vector

$$\underline{x}_B = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$$

to see that

$$A(x)_{B'} = A_{B,B'} \cdot \underline{x}_B. \quad (5.15)$$

This equality and constant matrix $\underline{A}_{B,B'}$ are called the *matrix-representation* of mapping A . In Example 5.2 we have seen that multiplying by an $m \times n$ real matrix is a linear mapping of R^n into R^m . This matrix representation shows that all linear mappings between finitely generated vector spaces can be represented as multiplication by constant matrices. Since the columns of $\underline{A}_{B,B'}$ are the coordinate vectors of the images $A(\underline{u}_1), \dots, A(\underline{u}_n)$ of the basis elements of B , it is easy to see that $\dim(R(A)) = \text{rank}(\underline{A}_{B,B'})$. Equalities (5.13) and (5.14) show that matrix $\underline{A}_{B,B'}$ can be constructed in the following way:

Step 1. Find the images $A(\underline{u}_1), \dots, A(\underline{u}_n)$ of the basis elements;

Step 2. Compute the coordinate vectors of these images in basis B' ;

Step 3. Construct the $m \times n$ matrix the columns of which are the coordinate vectors of $A(\underline{u}_1), \dots, A(\underline{u}_n)$, respectively.

This procedure is illustrated in the following examples.

Example 5.10. Let V and V' be two finitely generated vector spaces and assume that $A(\underline{x}) = \underline{0}'$ for all $\underline{x} \in V$, where $\underline{0}'$ denotes the zero element of V' . Then for all basis elements \underline{u}_i , $A(\underline{u}_i) = \underline{0}'$. Notice that all coordinates of the zero vector $\underline{0}'$ are equal to zero, since

$$\underline{0}' = 0 \cdot \underline{v}_1 + 0 \cdot \underline{v}_2 + \dots + 0 \cdot \underline{v}_m.$$

Therefore all columns of matrix $\underline{A}_{B,B'}$ is the zero vector, consequently $\underline{A}_{B,B'}$ is the $m \times n$ zero matrix.

Example 5.11. Let V be a finitely generated vector space with basis $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ and select $V' = V$ and $B' = B$. Define A as the identity mapping I , then for all basis elements

$$A(\underline{u}_i) = \underline{u}_i$$

with coordinate vector \underline{e}_i , since

$$\underline{u}_i = 0 \cdot \underline{u}_1 + \dots + 0 \cdot \underline{u}_{i-1} + 1 \cdot \underline{u}_i + 0 \cdot \underline{u}_{i+1} + \dots + 0 \cdot \underline{u}_n.$$

Therefore the columns of $\underline{A}_{B,B'}$ are $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$, respectively. That is, $\underline{A}_{B,B'}$ is the $n \times n$ identity matrix.

Example 5.12. Let V be the set of single variable real polynomials of degree at most n , and select $V' = V$. Assume that mapping A is defined

as $A(p) = \frac{d}{dt} p$ for all $p \in V$. From Theorem 2.7 we conclude that vector space V is $n+1$ dimensional with a basis $B = \{1, t, t^2, \dots, t^n\}$, since each polynomial of degree at most n can be uniquely represented as a linear combination of these elements. Select $B' = B$, which means that $\underline{u}_1 = \underline{v}_1 = 1$, $\underline{u}_2 = \underline{v}_2 = t, \dots, \underline{u}_{n+1} = \underline{v}_{n+1} = t^n$. The images of the basis elements are as follows:

$$A(\underline{u}_1) = 0, A(\underline{u}_2) = 1, A(\underline{u}_3) = 2t, \dots, A(\underline{u}_{n+1}) = nt^{n-1}.$$

Using the same basis elements we see that

$$\begin{aligned} A(\underline{u}_1) &= 0 \cdot \underline{u}_1 + 0 \cdot \underline{u}_2 + \dots + 0 \cdot \underline{u}_n + 0 \cdot \underline{u}_{n+1} \\ A(\underline{u}_2) &= 1 \cdot \underline{u}_1 + 0 \cdot \underline{u}_2 + \dots + 0 \cdot \underline{u}_n + 0 \cdot \underline{u}_{n+1} \\ A(\underline{u}_3) &= 0 \cdot \underline{u}_1 + 2 \cdot \underline{u}_2 + \dots + 0 \cdot \underline{u}_n + 0 \cdot \underline{u}_{n+1} \\ &\dots \\ A(\underline{u}_{n+1}) &= 0 \cdot \underline{u}_1 + 0 \cdot \underline{u}_2 + \dots + n \cdot \underline{u}_n + 0 \cdot \underline{u}_{n+1}. \end{aligned}$$

That is, the coordinate vectors of these images are obtained as

$$A(\underline{u}_1)_{B'} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, A(\underline{u}_2)_{B'} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, A(\underline{u}_3)_{B'} = \begin{pmatrix} 0 \\ 2 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, A(\underline{u}_{n+1})_{B'} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ n \\ 0 \end{pmatrix}.$$

And finally, matrix $A_{B',B}$, is constructed as the matrix the columns of which are these coordinate vectors in the same order as above:

$$A_{B,B'} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The elements of this matrix are characterized as

$$a_{ij} = \begin{cases} i & \text{if } 1 \leq i \leq n \text{ and } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

In the previous two sections we have introduced and examined the sums, scalar multiples, products, and inverses of linear mappings. We will now see how their matrix representations are effected by these operations.

Let V and V' be finitely generated vector spaces. As always, it is assumed that either both are real or both are complex. Let $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ and $B' = \{\underline{v}_1, \dots, \underline{v}_m\}$ be bases in V and V' , respectively. Assume that A and C are two linear mappings of V into V' . Use equations (5.13) and (5.14) for both mappings to see that for all $\underline{x} \in V$,

$$A(\underline{x}) = x_1 A(\underline{u}_1) + \dots + x_n A(\underline{u}_n),$$

and

$$C(\underline{x}) = x_1 C(\underline{u}_1) + \dots + x_n C(\underline{u}_n).$$

Therefore

$$(A + C)(\underline{x}) = A(\underline{x}) + C(\underline{x}) = x_1 (A(\underline{u}_1) + C(\underline{u}_1)) + \dots + x_n (A(\underline{u}_n) + C(\underline{u}_n)).$$

From equation (5.2) we know that for all i , the coordinate vector of $A(\underline{u}_i) + C(\underline{u}_i)$ is the sum of the coordinate vectors of $A(\underline{u}_i)$ and $C(\underline{u}_i)$. Therefore the columns in the matrix representation of mapping $A + C$ are the sums of the corresponding columns of the matrix representations of mappings A and C . That is, the matrix representation of $A + C$ can be obtained as the sum of the matrix representations of A and C :

$$(A + C)_{B, B'} = \underline{A}_{B, B'} + \underline{C}_{B, B'}. \quad (5.16)$$

If A is a linear mapping of V into V' and a is a scalar, then

$$\begin{aligned} (aA)(\underline{x}) &= a \cdot A(\underline{x}) = a(x_1 A(\underline{u}_1)) + \dots + a(x_n A(\underline{u}_n)) \\ &= x_1 (a \cdot A(\underline{u}_1)) + \dots + x_n (a \cdot A(\underline{u}_n)). \end{aligned}$$

From equation (5.3) we know that the coordinate vector of a scalar multiple of an element equals the same scalar multiple of the coordinate vector of that element. Therefore each column of the matrix representation of aA can be obtained by multiplying each column of the matrix representation of A by a . That is,

$$(aA)_{B, B'} = a \cdot \underline{A}_{B, B'}. \quad (5.17)$$

In Section 5.4 we have seen that the set $L(V, V')$ of linear mappings of V into V' is a vector space. Notice that the correspondence $A \mapsto \underline{A}_{B, B'}$ is one-to-one, and relations (5.16) and (5.17) imply that it can be considered as a one-to-one linear mapping of $L(V, V')$ onto the set of all $m \times n$ real (or complex) matrices. Hence these vector spaces are isomorphic, and

$$\dim(L(V, V')) = m \cdot n,$$

since in Example 2.19 we have shown that the space of all $m \times n$ real or complex matrices is $m \cdot n$ dimensional.

Assume next that V , V' , and V'' are finitely generated vector spaces, either all are real or all are complex. Let B , B' and B'' be a basis of V , V' , and V'' , respectively. If $C \in L(V, V')$ and $A \in L(V', V'')$, then for all $\underline{x} \in V$,

$$(AC)(\underline{x}) = A(C(\underline{x})).$$

Let \underline{x}_B be the coordinate vector of \underline{x} in basis B , and let $\underline{A}_{B', B''}$ and $\underline{C}_{B, B'}$ be the matrix representations of mappings A and C , respectively. The repeated application of equality (5.15) implies that

$$\begin{aligned} (AC)(\underline{x})_{B''} &= A(C(\underline{x}))_{B''} = \underline{A}_{B', B''} \cdot \underline{C}(\underline{x})_{B'} \\ &= \underline{A}_{B', B''} \cdot \underline{C}_{B, B'} \cdot \underline{x}_B = (\underline{A}_{B', B''} \cdot \underline{C}_{B, B'}) \cdot \underline{x}_B. \end{aligned}$$

Hence we have shown that

$$(\underline{A} \cdot \underline{C})_{B', B''} = \underline{A}_{B', B''} \cdot \underline{C}_{B, B'}, \quad (5.18)$$

that is the matrix representation of a product of two mappings equals the product of their matrix representations.

Consider finally the special case when $V' = V$. Let $A \in L(V)$ be an invertible linear transformation, and let C denote the inverse of A . Then $A \cdot C = I$, and combining relation (5.18) with the matrix representation of the identity transformation derived in Example 5.11 we immediately see that

$$\underline{I} = (I)_{B, B} = (\underline{A} \cdot \underline{C})_{B, B} = \underline{A}_{B, B} \cdot \underline{C}_{B, B}.$$

That is, $\underline{C}_{B,B}$ is the inverse of $\underline{A}_{B,B}$. This observation can be summarized as

$$(\underline{A}^{-1})_{B,B} = \underline{A}_{B,B}^{-1}, \quad (5.19)$$

which can be interpreted by saying that the matrix representation of the inverse of a linear transformation is the inverse of the matrix representation of the original transformation.

Comparing the results of this and the previous sections with the elements of matrix algebra outlined in Chapter 1, an analogy becomes clear. Addition, multiplication by scalars, products, and inverses of linear mappings have the same properties as the corresponding operations defined on real or complex matrices.

As the conclusion of this section, the special class of linear mappings from a finitely generated real (or complex) vector space V into the set of real (or complex) numbers will be briefly examined. As we mentioned earlier in Example 5.3, such linear mappings are called *linear functionals*. The set of all real (or complex) valued linear functionals defined on V is called the *dual space of V* . This notion has already been introduced at the end of Section 5.4. Let n denote the dimension of V . Since V' is one dimensional, the matrix representation of any linear functional is a $1 \times n$ matrix, which is actually a row vector. If A denotes a linear functional, and $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ is a given basis of V , then the matrix representation of mapping A in basis B is given as

$$\underline{A}_{B,B} = (A(\underline{u}_1), \dots, A(\underline{u}_n)).$$

This row vector can be viewed as an element of R^n (or C^n), therefore the dual space of an n -dimensional vector space V is isomorphic to R^n (or C^n).

5.7 Coordinates and Matrix Representation in a New Basis

Let V denote a finitely generated vector space and let $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ be a basis of V . In Section 5.1 we have seen that there is a unique correspondence between an arbitrary element $\underline{x} \in V$ and its coordinate vector \underline{x}_B . In Example 5.1 we have seen that the coordinate vector usually changes if a new basis is selected in V . We will first examine how to obtain the new coordinate vector efficiently.

Assume that $B' = \{\underline{v}_1, \dots, \underline{v}_n\}$ is a new basis in V . From Corollary 1 of Theorem 2.9 we know that B and B' must have the same number of elements. For $j = 1, 2, \dots, n$, let \underline{v}_{jB} denote the coordinate vector of \underline{v}_j in basis B , and define matrix \underline{T} as the $n \times n$ matrix with columns $\underline{v}_{1B}, \underline{v}_{2B}, \dots, \underline{v}_{nB}$. If t_{ij} denotes the (i, j) element of matrix \underline{T} , then for all j ,

$$\underline{v}_{jB} = \begin{pmatrix} t_{1j} \\ t_{2j} \\ \dots \\ t_{nj} \end{pmatrix},$$

and therefore

$$\underline{v}_j = t_{1j}\underline{u}_1 + \dots + t_{nj}\underline{u}_n. \quad (5.20)$$

Assume that $\underline{x} \in V$ is an arbitrary element. Let $\underline{x}_B = (x_i)$ and $\underline{x}_{B'} = (x'_j)$ denote its coordinate vector in basis B and B' , respectively.

The definition of the coordinates imply that

$$\underline{x} = \sum_{i=1}^n x_i \underline{u}_i = \sum_{j=1}^n x'_j \underline{v}_j,$$

and by combining this equation with (5.20) we have

$$\sum_{i=1}^n x_i \underline{u}_i = \sum_{j=1}^n x_j \underline{v}_j = \sum_{j=1}^n x'_j \left(\sum_{i=1}^n t_{ij} \underline{u}_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n t_{ij} x'_j \right) \underline{u}_i.$$

Since $B = \{\underline{u}_1, \dots, \underline{u}_n\}$ is a basis in B and each element of V can be uniquely represented as the linear combination of the basis elements, we conclude that for $i = 1, 2, \dots, n$,

$$x_i = \sum_{j=1}^n t_{ij} x'_j.$$

Notice that x_i is the i^{th} element of the coordinate vector \underline{x}_B , and $\sum_{j=1}^n t_{ij} x'_j$ is the i^{th} element of the product $\underline{T} \cdot \underline{x}_B$. Therefore we can summarize the above derivation as

$$\underline{x}_B = \underline{T} \cdot \underline{x}_B. \quad (5.21)$$

Here \underline{T} is called the *transformation matrix*. We will next prove that \underline{T} is a nonsingular matrix. The columns of \underline{T} are the coordinate vectors of the new basis elements $\underline{v}_1, \dots, \underline{v}_n$ in basis B . The mapping $\underline{x} \mapsto \underline{x}_B$ is linear and one-to-one, so it is an isomorphism. From the proof of Theorem 5.3 we can conclude that the linear independence of the basis elements $\underline{v}_1, \dots, \underline{v}_n$ implies the same for the coordinate vectors. Therefore the columns of \underline{T} are linearly independent, which implies that \underline{T}^{-1} exists. Multiply both sides of equation (5.21) by \underline{T}^{-1} to obtain an equivalent form:

$$\underline{x}_B = \underline{T}^{-1} \underline{x}_B. \quad (5.22)$$

Relations (5.21) and (5.22) are usually called the *basis transformation equations*, and their use is illustrated in the next example.

Example 5.13. Let $V = R^3$ and $B = \{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$, furthermore

$$\underline{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

as in Example 5.1. Since

$$\underline{x} = 1 \cdot \underline{e}_1 + 1 \cdot \underline{e}_2 + 1 \cdot \underline{e}_3,$$

the coordinate vector of \underline{x} in basis B is the following:

$$\underline{x}_B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Select the new basis

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Then

$$\underline{v}_1 = 1 \cdot \underline{e}_1 + 1 \cdot \underline{e}_2 + 1 \cdot \underline{e}_3,$$

$$\underline{v}_2 = 1 \cdot \underline{e}_1 + 1 \cdot \underline{e}_2 + 0 \cdot \underline{e}_3,$$

$$\underline{v}_3 = 1 \cdot \underline{e}_1 + 0 \cdot \underline{e}_2 + 0 \cdot \underline{e}_3,$$

and so

$$\underline{v}_{1B} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{v}_{2B} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{v}_{3B} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore

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

where we have copied coordinate vectors $\underline{v}_{1B}, \underline{v}_{2B}, \underline{v}_{3B}$ into the columns of \underline{T} . The simple application of the elimination method shows that

$$\underline{T}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Then equality (5.22) implies that

$$\underline{x}_B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

which coincides with the result obtained in Example 5.1. ♦

Consider next a linear transformation $A \in L(V)$ on V . If B is a basis of V , then $\underline{A}_{B,B}$ denotes the matrix representation of A . Assume now that a new basis B' is selected. Using equations (5.21) and (5.22) we can easily find the matrix representation of A in the new basis. Let $\underline{x} \in V$ be an arbitrary element. Then

$$(A(\underline{x}))_B = \underline{A}_{B,B} \cdot \underline{x}_B,$$

and by using relation (5.21) we have

$$\underline{T} \cdot (A(\underline{x}))_{B'} = \underline{A}_{B,B} \cdot \underline{T} \cdot \underline{x}_{B'}.$$

Multiply both sides by \underline{T}^{-1} to get

$$(A(\underline{x}))_{B'} = \underline{T}^{-1} \underline{A}_{B,B} \underline{T} \underline{x}_{B'},$$

which implies that the matrix representation of A in basis B' is the following:

$$\underline{A}_{B',B'} = \underline{T}^{-1} \underline{A}_{B,B} \underline{T}. \quad (5.23)$$

Definition 5.5. The $n \times n$ matrices \underline{A} and \underline{C} are called *similar* if there exists an invertible matrix \underline{X} such that

$$\underline{C} = \underline{X}^{-1} \underline{A} \underline{X}.$$

Relation (5.23) shows that the matrix representations of a linear transformation $A \in L(V)$ are similar matrices. It can also be proved that if matrix \underline{A} is the matrix representation of a mapping $A \in L(V)$ in a given basis B and matrix \underline{C} is similar to \underline{A} , then there is another basis B' of V such that \underline{C} is the matrix representation of A in B' . That is, similar matrices can be viewed as the matrix representations of the same linear transformations in different bases. Therefore the ranks of similar matrices are equal.

The use of equation (5.23) is illustrated in the following example.

Example 5.14. Let $V = R^3$ with the natural basis $B = \{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$. Assume that the matrix representation of a transformation A is given by the 3×3 matrix

$$\underline{A}_{B,B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Select the new basis B' with elements

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

The transformation matrix and its inverse have been determined in the previous example:

$$\underline{T} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \underline{T}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Then the matrix representation of A in the new basis B' is the following:

$$\underline{A}_{B',B'} = \underline{T}^{-1} \underline{A}_{B,B} \underline{T} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}.$$

♦

5.8 Applications

In this section some linear transformations defined on the two-dimensional plane will be first examined, and then special integral-transformations will be introduced.

1. Vectors are identified with directed line segments in the plane. Each vector therefore corresponds to a point in the plane given its position vector from the origin $\underline{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ as shown in Figure 5.1. If r is the length of the vector and α is the angle between the vector and the positive half of the horizontal axis then

$$x_1 = r \cdot \cos \alpha \quad \text{and} \quad x_2 = r \cdot \sin \alpha.$$

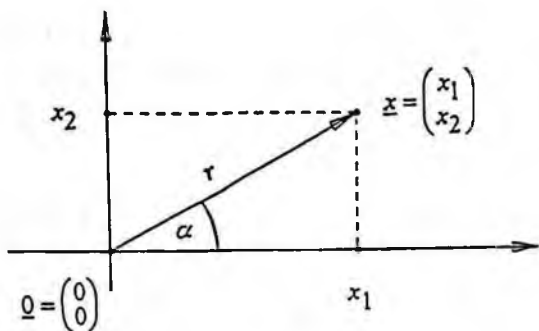


Figure 5.1. Vectors in the plane.

Consider first the transformation A which multiplies a vector by a scalar α . Consider the basis $B = \{\underline{e}_1, \underline{e}_2\}$, where $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Since

$$A(\underline{e}_1) = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \quad \text{and} \quad A(\underline{e}_2) = \begin{pmatrix} 0 \\ \alpha \end{pmatrix},$$

the matrix representation of this transformation is

$$\underline{A}_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} = \alpha \underline{I}.$$

Consider next the transformation B_α which rotates each vector with a given angle α around the origin. As Figure 5.2 shows,

$$B_{\alpha}(\underline{e}_1) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \text{and} \quad B_{\alpha}(\underline{e}_2) = \begin{pmatrix} \cos\left(\alpha + \frac{\pi}{2}\right) \\ \sin\left(\alpha + \frac{\pi}{2}\right) \end{pmatrix} = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}.$$

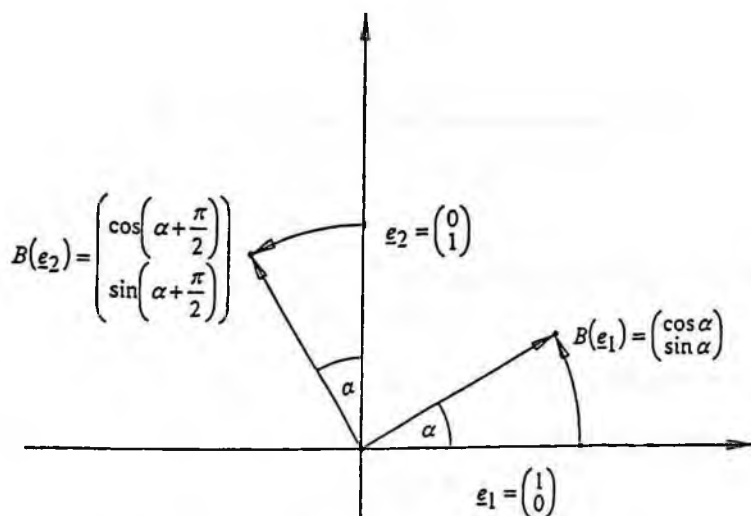


Figure 5.2 Illustration of a rotation.

Therefore the matrix representation of this transformation is

$$\underline{B}_{\alpha} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Consider now the transformation C that gives the mirror image of each vector with respect to the vertical axis. Figure 5.3 shows that

$$C(\underline{e}_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{and} \quad C(\underline{e}_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

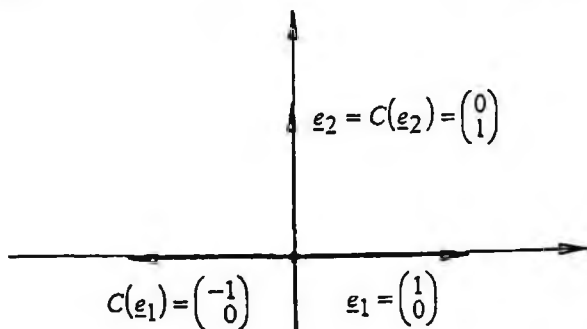


Figure 5.3. Illustration of transformation that gives mirror image with respect to vertical axis.

Therefore the matrix representation is as follows:

$$\underline{C} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Assume that transformation D gives the mirror image of each vector with respect to the horizontal axis. From Figure 5.4 we see that

$$D(\underline{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad D(\underline{e}_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

therefore the matrix representation is given as

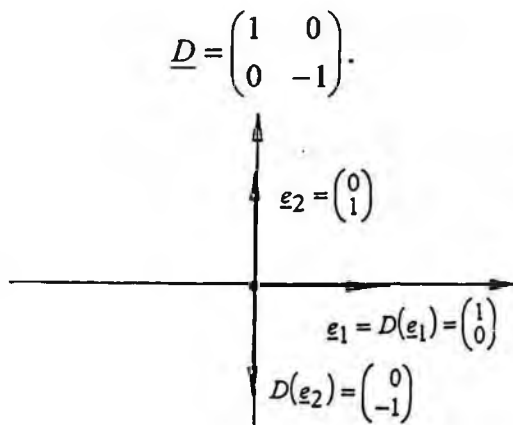


Figure 5.4. Illustration of transformation that gives mirror image with respect to horizontal axis.

The transformation E which gives the mirror image of each vector with respect to the origin is equivalent to a rotation by π . Therefore the matrix representation of E is the following:

$$\underline{E} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to see that these transformations satisfy the following relations:

$$\begin{aligned} \underline{A}_0 &= \underline{O}, \underline{A}_1 = \underline{I}, \\ \underline{B}_0 &= \underline{I}, \underline{B}_\pi = -\underline{I} = \underline{E}, \end{aligned}$$

and

$$\underline{B}_\alpha \underline{B}_\beta = \underline{B}_{\alpha+\beta},$$

since repeated rotations by angles α and β are equivalent to a single rotation by angle $\alpha + \beta$. Notice that this relation can be rewritten as

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}.$$

Compare the corresponding elements on the left and right hand sides to see that

$$\begin{aligned} \cos \alpha \cos \beta - \sin \alpha \sin \beta &= \cos(\alpha + \beta) \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta &= \sin(\alpha + \beta). \end{aligned}$$

This simple derivation provides an elegant new proof of these well-known trigonometric identities.

Notice furthermore that

$$D = -C \quad \text{and} \quad D^2 = C^2 = I.$$

This last identity can be interpreted as

$$D^{-1} = D \quad \text{and} \quad C^{-1} = C.$$

2. Consider next the vector space $V = C[a, b]$ of continuous real functions, and assume that k is a continuous function ($k \in C([a, b] \times [a, b])$) of two variables. Consider the transformation $A: V \mapsto V$ defined as

$$A(x)(t) = \int_a^b k(t, s)x(s)ds \quad (t \in [a, b]) \quad (5.24)$$

for all $x \in C[a, b]$. Here $A(x)$ denotes the image function, the value of which at point t is defined by equation (5.24). Many problems of applied mathematics can be reduced to the solution of the *Fredholm integral equation*:

$$x(t) = \int_a^b k(t, s)x(s)ds + f(t) \quad (t \in [a, b]), \quad (5.25)$$

where $f \in C[a, b]$ is a given function. This equation can be rewritten as

$$x = A(x) + f,$$

or

$$(I - A)(x) = f.$$

If transformation $I - A$ is invertible, then the solution can be obtained in the following form:

$$x = (I - A)^{-1}(f). \quad (5.26)$$

A modified version of equation (5.25) can be written as

$$x(t) = \int_a^t k(t, s)x(s)ds + f(t) \quad (t \in [a, b]) \quad (5.27)$$

which is called the *Volterra integral equation*. Introduce the modified transformation

$$B(x)(t) = \int_a^t k(x, s)x(s)ds,$$

then equation (5.27) can be rewritten as

$$x = B(x) + f,$$

or

$$(I - B)(x) = f.$$

It can be proved that mapping B is invertible, therefore equation (5.27) always has a unique solution in $C[a, b]$. For more detail see, for example Szidarovszky and Yakowitz (1978, Chapter 4).

5.9 Exercises

1. Show that the following transformations are not linear:

$$\text{a) } \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 2 \end{pmatrix}$$

$$\text{b) } \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 \\ x+y \end{pmatrix}$$

$$c) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+1 \\ y+1 \end{pmatrix}$$

$$d) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \sin x \\ y \end{pmatrix}.$$

2. Show that the following transformations are linear:

$$a) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -3x \\ 2y \end{pmatrix}$$

$$b) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ x-y \end{pmatrix}$$

$$c) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x+3y \\ x-y \end{pmatrix}$$

$$d) \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x+2y \\ x-2y \end{pmatrix}.$$

3. Find the matrix representation of the transformations of the previous problem.

4. Denote the transformations

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -3x \\ 2y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ x-y \end{pmatrix}$$

by A and B , respectively.

Find $A+B$, $A-B$, $2A$, $3B$, $2A+3B$, AB , A^2 , A^2B , and BA .

5. Check which of the mappings of Problem 2 are one-to-one.

6. Let V be the set of all single variable real polynomials of degree at most n . If $p(x) = a_0 + a_1x + \dots + a_nx^n$, then define $A(p)$ as

$$A(p)(x) = a_0 + \frac{a_1x}{1!} + \frac{a_2x^2}{2!} + \dots + \frac{a_nx^n}{n!}. \text{ Is } A \text{ a linear}$$

transformation?

7. Find the ranks of the resulting transformations of Problem 4.

8. Prove that if $A \in L(V, V)$, then $A^2 = O$ if and only if $R(A) \subseteq N(A)$.

9. Let $A \in L(V, V)$. Is it possible that $R(A) \cap N(A) \neq \{0\}$?

10. Is there any linear mapping $A \in L(R^3, R^2)$ such that

$$A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{and} \quad A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}?$$

11. Let V be the set of all continuously differentiable functions on $[0, 1]$, and let W be the set of all continuous functions on the same interval $[0, 1]$. Define $A \in L(V, W)$ as for all $f \in V$, $A(f) = f'$. Is A a linear mapping?

12. Let $\underline{a} \in R^n$ be a fixed vector. Define mapping A as $A(\underline{x}) = \underline{x} + \underline{a}$ for all $\underline{x} \in R^n$. Is A a linear mapping?

13. Let $A \in L(R^n, R^{n-1})$ be defined as

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_3 \\ \vdots \\ x_{n-1} - x_n \end{pmatrix}.$$

- Show that A is a linear mapping;
- Give a matrix representation of A ;
- Find $N(A)$ and $R(A)$;
- Find the rank of A .

14. Let A be a mapping of R^3 into itself such that $A(\underline{x})$ is the mirror image of \underline{x} with respect to the origin. Find a matrix representation of A .

15. Find the coordinates of vector \underline{x} in basis $B = \{\underline{x}_1, \underline{x}_2\}$, where

$$\underline{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \underline{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \underline{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

16. Let $V = C[a, b]$, and let $g \in C[a, b]$ be a given function. Define the mapping $A: V \rightarrow R$ as

$$A(f) = \int_a^b f(t)g(t)dt.$$

Is this mapping linear?

17. Find the inverses of the linear transformation given in Problem 2.

18. Repeat Example 5.13 for vectors

$$\underline{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \underline{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

19. Repeat Example 5.14 for matrix

$$\underline{A}_{B,B} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix}$$

and basis elements as given in the previous problem.

20. Show how the matrix of the previous problem changes in the new basis

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \underline{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \underline{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

21. Assume that $k(t,s) = \alpha \neq \frac{1}{b-a}$ for all t and s , where α is a

given constant. Show that equation (5.25) has a unique solution

$$x(t) = f(t) + c$$

with

$$c = \frac{\alpha \int_a^b f(s) ds}{1 - \alpha(b - a)}.$$

22. In addition to the assumptions of the previous problem assume that f is continuously differentiable. Show that the solution of the Volterra equation (5.27) solves the initial-value-problem

$$\dot{x}(t) = \alpha x(t) + \dot{f}(t), \quad x(a) = f(a).$$

23. (Continuation of Problem 22). Show that under the conditions of the previous problem,

$$x(t) = e^{\alpha t} \left(f(a)e^{-\alpha a} + \int_a^t \dot{f}(s)e^{-\alpha s} ds \right).$$

24. Assume that $k(t, s) \equiv K(s)$, where K is a continuous function on $[a, b]$, and f is continuously differentiable. Show that the solution of the Volterra equation (5.27) solves the initial-value-problem

$$\dot{x}(t) = K(t)x(t) + \dot{f}(t), \quad x(a) = f(a).$$

25. (Continuation of Problem 24). Generalize the result of Problem 23. for the more general case of the previous problem.

Chapter 6

Eigenvalues, Invariant Subspaces, Canonical Forms

6.1 Introduction

In this chapter the general theory of the space $L(V)$ of linear transformations on a vector space V will be discussed. It will always be assumed that V is finitely generated, and therefore in a given basis, each $A \in L(V)$ has an $n \times n$ matrix representation with some n . The main objective of this chapter is to introduce methodology that can be used to transform matrices into simple forms. Most of such methods are based on the idea of decomposing space V into special low dimensional subspaces with the property that a given transformation A maps each subspace into itself. Therefore A can also be decomposed into more simple transformations of lower ranks, which will result in matrix representations having special structures.

6.2 Basic Concepts

Let V be a vector space and assume that $A \in L(V)$ is a linear transformation of V into itself. The methodology to be introduced and examined in this chapter will be based on the following concept.

Definition 6.1. A subspace V_1 of V is called an *A-invariant subspace* if for all $\underline{x} \in V_1$, $A(\underline{x}) \in V_1$.

Notice first that V itself is obviously an A -invariant subspace, and the subspace $V_1 = \{\underline{0}\}$ is also A -invariant, since $A(\underline{0}) = \underline{0}$. These subspaces are the trivial A -invariant subspaces. A nontrivial case is shown in the following example.

Example 6.1. Select $V = R^2$, and for all $\underline{x} \in V$, define

$$A(\underline{x}) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \cdot \underline{x}.$$

Consider the subspace generated by vector $\underline{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$:

$$V_1 = \left\{ \underline{x} \mid \underline{x} = \begin{pmatrix} a \\ 2a \end{pmatrix}, a \in R \right\}.$$

We can easily prove that V_1 is A -invariant. Simple calculation shows that for all vectors of V_1 ,

$$A(\underline{x}) = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} a \\ 2a \end{pmatrix} = \begin{pmatrix} a + 2a \\ 2a + 4a \end{pmatrix} = \begin{pmatrix} 3a \\ 6a \end{pmatrix} \in V_1.$$

We start our discussion with the simplest nontrivial A -invariant subspaces. Let V_1 be a one-dimensional subspace of V , then it is generated by a nonzero vector \underline{v} :

$$V_1 = \{ \underline{x} \mid \underline{x} = a\underline{v}, a \text{ is a scalar} \}.$$

This subspace is A -invariant if and only if for all $\underline{x} \in V_1$, $A(\underline{x})$ is a constant multiple of \underline{x} . This condition can be reformulated by saying that

$$A(\underline{x}) = \lambda \underline{x} \quad (6.1)$$

with some scalar λ . The selection $\underline{x} = \underline{0}$ always satisfies this equation with arbitrary value of λ , therefore this trivial solution has no interest to us.

Definition 6.2. Let $A \in L(V)$ be a linear transformation. A scalar λ is called an *eigenvalue* of A , if there is an $\underline{x} \neq \underline{0}$ in V such that equation (6.1) is satisfied. The solutions $\underline{x} \neq \underline{0}$ are called the *eigenvectors* of A . Equation (6.1) is usually called the *eigenvector equation* of mapping A .

Notice that if V is a real vector space, then scalars are also real, therefore only real eigenvalues are considered. If V is complex, then eigenvalues can be both real and complex numbers.

First we point out an important relation between the eigenvalues of a linear transformation and the transformation being one-to-one. From Theorem 5.5 we know that a linear transformations A is one-to-one if and only if its nullspace consists of only the zero element. The eigenvector equation implies that this is the case if and only if $\lambda = 0$ is not an eigenvalue of A . That is, A is one-to-one if and only if all eigenvalues of A are nonzero.

First we will prove the linear independence of eigenvectors associated to distinct eigenvalues. The theorem will play an important role later in this chapter in transforming matrices into diagonal form.

Theorem 6.1. Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of a transformation $A \in V(L)$, and let $\underline{x}_1, \dots, \underline{x}_k$ be associated eigenvectors. Then $\underline{x}_1, \dots, \underline{x}_k$ are linearly independent.

Proof. Assume that a linear combination of these elements is zero,

$$a_1 \underline{x}_1 + a_2 \underline{x}_2 + \dots + a_k \underline{x}_k = \underline{0}, \quad (6.2)$$

where at last one coefficient, say, a_k is nonzero. By renumbering the elements we can always have nonzero coefficient in the last term. Multiply this equation by λ_1 to have

$$a_1 \lambda_1 \underline{x}_1 + a_2 \lambda_1 \underline{x}_2 + \dots + a_k \lambda_1 \underline{x}_k = \underline{0}. \quad (6.3)$$

Apply mapping A to both sides of equality (6.2),

$$a_1 A(\underline{x}_1) + a_2 A(\underline{x}_2) + \dots + a_k A(\underline{x}_k) = \underline{0}.$$

Use the fact that for $i = 1, 2, \dots, k$, $A(\underline{x}_i) = \lambda_i \underline{x}_i$ to rewrite this last relation as

$$a_1 \lambda_1 \underline{x}_1 + a_2 \lambda_2 \underline{x}_2 + \dots + a_k \lambda_k \underline{x}_k = \underline{0}.$$

Subtract this equality from (6.3) to see that

$$a_2 (\lambda_1 - \lambda_2) \underline{x}_2 + \dots + a_k (\lambda_1 - \lambda_k) \underline{x}_k = \underline{0},$$

that is,

$$a_2^{(1)} \underline{x}_2 + \dots + a_k^{(1)} \underline{x}_k = \underline{0}, \quad (6.4)$$

where $a_i^{(1)} = a_i (\lambda_1 - \lambda_i)$ for $i = 2, \dots, k$. Notice that for all i , $a_i^{(1)} \neq 0$ if and only if $a_i \neq 0$, since the eigenvalues are distinct and therefore $\lambda_1 - \lambda_i \neq 0$. The above derivation shows that if a linear combination of $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_k$ is zero, then \underline{x}_1 can be eliminated from the zero linear combination, and all nonzero coefficients remain nonzero. Repeat the same

procedure for equation (6.4) to eliminate \underline{x}_2 , and repeat again for the resulting equation to eliminate \underline{x}_3 , and so on. Finally after eliminating all elements $\underline{x}_1, \dots, \underline{x}_{k-1}$, we have

$$a_k^{(k-1)} \underline{x}_k = \underline{0}.$$

Since $\underline{x}_k \neq \underline{0}$, $a_k^{(k-1)} = 0$, which implies that $a_k = 0$. This conclusion contradicts the initial assumption that a_k is nonzero. Thus the proof is complete.

♣

In the special case when V is finitely generated, with some n , mapping A has an $n \times n$ matrix representation and each element $\underline{x} \in V$ has an n -element coordinate vector in any given basis B . For the sake of simplicity let \underline{A} and \underline{x} denote the matrix representation of A and the coordinate vector of \underline{x} . In this notation we do not indicate basis B in the subscripts of these symbols as we have done earlier in Chapter 5, since the basis is considered now fixed. Then equation (6.1) can be rewritten as

$$\underline{A}\underline{x} = \lambda \underline{x},$$

which is equivalent to the homogeneous equation

$$(\underline{A} - \lambda \underline{I})\underline{x} = \underline{0}, \quad (6.5)$$

where \underline{I} is the $n \times n$ identity matrix and $\underline{0}$ is the n -element zero vector. A scalar λ is an eigenvalue of \underline{A} if and only if this equation has a nonzero solution \underline{x} . Since $\underline{0}$ is always a solution of this equation, nonzero solution exists if and only if multiple solutions exist. From Theorem 4.6 we know that a necessary and sufficient condition for the existence of multiple solutions is given by the condition that the determinant of the coefficient matrix is zero:

$$\det(\underline{A} - \lambda \underline{I}) = 0. \quad (6.6)$$

Equation (6.6) is called the *characteristic equation* of matrix \underline{A} . Notice that $\underline{A} - \lambda \underline{I}$ can be obtained easily from \underline{A} , we have to subtract λ from each of its diagonal elements. That is, equation (6.6) has the equivalent form:

$$\det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2,n-1} & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} - \lambda \end{pmatrix} = 0. \quad (6.7)$$

Expanding this determinant with respect to the first row it is easy to show by mathematical induction that the left-hand side is always a polynomial of degree n and the coefficient of λ^n is $(-1)^n$. This polynomial is called the *characteristic polynomial* of matrix \underline{A} . It is easy to see that the characteristic polynomial does not depend on the basis being selected in V . The matrix representation of \underline{A} in another basis is $\underline{T}^{-1} \underline{A} \underline{T}$, where \underline{T} is a nonsingular matrix. The characteristic polynomial of $\underline{T}^{-1} \underline{A} \underline{T}$ is the following:

$$\begin{aligned} \det(\underline{T}^{-1} \underline{A} \underline{T} - \lambda \underline{I}) &= \det(\underline{T}^{-1} (\underline{A} - \lambda \underline{I}) \underline{T}) \\ &= \det(\underline{T}^{-1}) \cdot \det(\underline{A} - \lambda \underline{I}) \cdot \det(\underline{T}). \end{aligned}$$

From the Corollary of Theorem 4.4 we know that

$$\det(\underline{T}^{-1}) = \frac{1}{\det(\underline{T})},$$

therefore the first and last factors cancel showing that the characteristic polynomial of $T^{-1}AT$ coincides with that of matrix A .

Based on the above discussion, the eigenvalues and eigenvectors of an $n \times n$ real or complex matrix can be obtained by using the following procedure:

Step 1. Expand determinant (6.7) to obtain the characteristic polynomial of matrix A ;

Step 2. Find the (real or complex) roots of the characteristic polynomial to get the eigenvalues;

Step 3. For each eigenvalue, solve the homogeneous system (6.5) to recover the eigenvectors.

This procedure is illustrated in the following example.

Example 6.2. Consider the 3-3 matrix

$$\underline{A} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Equation (6.7) has now the form

$$\det \begin{pmatrix} 1 - \lambda & 4 & 1 \\ 1 & 2 - \lambda & 3 \\ 1 & 3 & 2 - \lambda \end{pmatrix} = 0.$$

Expanding the determinant with respect to its first row, the left-hand side can be written as

$$\begin{aligned}
 & (1-\lambda) \cdot \det \begin{pmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{pmatrix} - 4 \det \begin{pmatrix} 1 & 3 \\ 1 & 2-\lambda \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 1 & 2-\lambda \\ 1 & 3 \end{pmatrix} \\
 &= (1-\lambda) \cdot (4-4\lambda+\lambda^2-9) - 4 \cdot (2-\lambda-3) + 1 \cdot (3-2+\lambda) \\
 &= (1-\lambda)(-5-4\lambda+\lambda^2) + (4\lambda+4) + (\lambda+1) = -\lambda^3 + 5\lambda^2 + 6\lambda.
 \end{aligned}$$

The roots of this polynomial are

$$\lambda_1=0, \lambda_2=-1, \text{ and } \lambda_3=6.$$

For eigenvalue $\lambda_1=0$, equation (6.5) has the form

$$\begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

that is,

$$\begin{aligned}
 x_1 + 4x_2 + x_3 &= 0 \\
 x_1 + 2x_2 + 3x_3 &= 0 \\
 x_1 + 3x_2 + 2x_3 &= 0.
 \end{aligned}$$

The elimination procedure is shown in Table 6.1, where the right-hand side numbers are not presented, since they are always equal to zero.

Table 6.1. Elimination for $\lambda_1 = 0$.

x_1	x_2	x_3
1	4	1
1	2	3
1	3	2
1	4	1
0	-2	2
0	-1	1
1	4	1
0	-2	2
0	0	0

Hence, x_3 is the free variable, and back substitution shows that $x_2 = x_3$ and $x_1 = -4x_2 - x_3 = -5x_3$. The general solution is therefore the following:

$$\underline{x} = \begin{pmatrix} -5x_3 \\ x_3 \\ x_3 \end{pmatrix}.$$

Thus, this vector is the eigenvector associated to $\lambda_1 = 0$ with arbitrary nonzero value of x_3 .

For $\lambda_2 = -1$, equation (6.5) can be written as

$$\begin{pmatrix} 2 & 4 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which is equivalent to the system

$$2x_1 + 4x_2 + x_3 = 0$$

$$x_1 + 3x_2 + 3x_3 = 0$$

$$x_1 + 3x_2 + 3x_3 = 0.$$

Similarly to the first eigenvalue, one may verify that the general solution is

$$\underline{x} = \begin{pmatrix} \frac{9}{2}x_3 \\ -\frac{5}{2}x_3 \\ x_3 \end{pmatrix}.$$

For $\lambda_3 = 6$, equation (6.5) is the following:

$$\begin{pmatrix} -5 & 4 & 1 \\ 1 & -4 & 3 \\ 1 & 3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

and the general solution is:

$$\underline{x} = \begin{pmatrix} x_3 \\ x_3 \\ x_3 \end{pmatrix}.$$

If we wish to determine a particular eigenvector associated to each eigenvalue, then we might select a particular value for x_3 . If $x_3 = 1$ is the choice, then the eigenvectors for $\lambda = 0, -1$, and 6 are

$$\begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{9}{2} \\ 5 \\ -2 \\ 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

respectively. ♦

The procedure discussed above is very useful to find one-dimensional invariant subspaces for any given $n \times n$ real or complex matrix. However to find higher dimensional invariant subspaces, the straightforward extension of this algorithm becomes very complicated, and therefore has no practical value. The case of finding two dimensional invariant subspaces of a 3×3 matrix is illustrated in the following example.

Example 6.3. Consider again matrix

$$\underline{A} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Any two dimensional subspace is characterized as

$$V_1 = \{\lambda \underline{x} + \mu \underline{y} \mid \{\underline{x}, \underline{y}\} \text{ is a basis of } V_1, \text{ and } \lambda, \mu \text{ are scalars}\}.$$

This subspace is \underline{A} -invariant if and only if both \underline{Ax} and \underline{Ay} are in V_1 , since for any vector $\underline{z} \in V_1$,

$$\underline{Az} = \underline{A}(\lambda \underline{x} + \mu \underline{y}) = \lambda \underline{Ax} + \mu \underline{Ay} \in V_1.$$

This observation can be mathematically formulated as the system:

$$\begin{aligned}\underline{Ax} &= \alpha \underline{x} + \beta \underline{y} \\ \underline{Ay} &= \gamma \underline{x} + \delta \underline{y},\end{aligned}$$

where vectors $\underline{x}, \underline{y}$ and scalars $\alpha, \beta, \gamma, \delta$ are the unknowns. These equations can be summarized as the homogeneous system

$$\begin{pmatrix} \underline{A} - \alpha \underline{I} & -\beta \underline{I} \\ -\gamma \underline{I} & \underline{A} - \delta \underline{I} \end{pmatrix} \begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \end{pmatrix}.$$

Similarly to the eigenvector equations the unknown scalars can be obtained as the solutions of the equation

$$\det \begin{pmatrix} \underline{A} - \alpha \underline{I} & -\beta \underline{I} \\ -\gamma \underline{I} & \underline{A} - \delta \underline{I} \end{pmatrix} = 0.$$

Notice that this determinant is $6 \cdot 6$ with four unknown parameters, therefore its expansion becomes complicated. Notice that if a k^{th} dimensional invariant subspace of an $n \times n$ matrix is to be determined, then this equation has k^2 unknown scalars, and the size of the determinant is $kn \times kn$.

◆

Due to the difficulty arising from the increasing size and more complicated structure of the determinant being expanded, a different approach has to be developed to find invariant subspaces. This new methodology is based on the following simple result.

Theorem 6.2. Let $A \in L(V)$, and assume that for a linear transformation $B \in L(V)$, $AB = BA$. Then the nullspace of B is an A -invariant subspace.

Proof. Let $\underline{x} \in N(B)$ be an arbitrary element. Then $A(\underline{x}) \in N(B)$, since

$$B(A(\underline{x})) = (B \cdot A)(\underline{x}) = (A \cdot B)(\underline{x}) = A(B(\underline{x})) = A(\underline{0}) = \underline{0},$$

which proves the assertion. ♣

If we wish to use this theorem in constructing A -invariant subspaces, then we have to find a large class of mappings that commute with A . Let p be a real or complex polynomial depending on whether V is a real or complex vector space. Since A commute with A^k for all $k \geq 0$, A and $p(A)$ commute as well. This simple observation shows that $N(p(A))$ is an A -invariant subspace for all polynomials p . Our general methodology for constructing invariant subspace will be mainly based on this idea. In developing further results we will use the fundamental properties of matrix polynomials, which are the subject of the next section.

6.3 Matrix Polynomials

We start this section with the summary of the most important properties of real and complex polynomials. As in some earlier examples, a single variable single valued function of t having the form

$$p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

is called a *polynomial of degree n* if the coefficients are real or complex numbers, and the leading coefficient, $a_n \neq 0$. If all coefficients are real, then the polynomial is called real, otherwise the polynomial is complex. A real (or complex) number t^* is called a *root of p* if $p(t^*) = 0$. We will use the following facts (see, for example, Herstein, 1964) regarding the roots of an n^{th} degree polynomial p :

1. For any complex polynomial of degree n , there are n (not necessarily distinct) real or complex roots.
2. If all coefficients are real and complex roots exist, they occur in conjugate pairs.
3. If n is odd and all coefficients are real, there is at least one real root.
4. If t^* is a root of p , then necessarily

$$p(t) = (t - t^*)q(t),$$

where q is a polynomial of degree $n - 1$.

The first fact is known as the *fundamental theorem of algebra*, the proof of which is based on the theory of functions of complex variables, therefore it is not included in this book. The second property can be proved by simple substitution. Let t^* be a complex root. If complex conjugate is denoted by overbar, then

$$\begin{aligned} p(\overline{t^*}) &= a_0 + a_1 \overline{t^*} + a_2 \overline{t^{*2}} + \dots + a_n \overline{t^{*n}} \\ &= \overline{a_0 + a_1 t^* + a_2 t^{*2} + \dots + a_n t^{*n}} = \overline{p(t^*)} = \overline{0} = 0, \end{aligned}$$

that is, $\overline{t^*}$ is also a root of p . Fact 3 is a consequence of Fact 2, since the number of complex roots of a real polynomial is always even, but the total number of real and complex roots is odd. Therefore the number of real roots is odd, hence nonzero. The last property is known as the *Factor Theorem*, and it can be proved by using the well-known algebraic identity

$$a^k - b^k = (a-b)(a^{k-1} + a^{k-2}b + \dots + ab^{k-2} + b^{k-1})$$

Simple algebra shows that

$$\begin{aligned} p(t) - p(t^*) &= a_1(t - t^*) + a_2(t^2 - t^{*2}) + \dots + a_n(t^n - t^{*n}) \\ &= a_1(t - t^*) + a_2(t - t^*)(t + t^*) + \dots + a_n(t - t^*)(t^{n-1} + t^{n-2}t^* + \dots + t \cdot t^{*n-2} + t^{*n-1}) \\ &= (t - t^*)q(t) \end{aligned}$$

where

$$q(t) = a_1 + a_2(t + t^*) + \dots + a_n(t^{n-1} + t^{n-2}t^* + \dots + t \cdot t^{*n-2} + t^{*n-1})$$

is a polynomial of degree $n-1$.

Some additional properties of polynomials which will be used later in this chapter are discussed next.

5. If p and q are polynomials of degree n and $m \leq n$, respectively, then there are unique polynomials h and r of degree $n-m$ and at most $m-1$ such that

$$p(t) = q(t)h(t) + r(t). \quad (6.8)$$

Equation (6.8) is called the *Remainder Theorem*, and polynomials h and r are called the *quotient* and the *remainder polynomials*, respectively. Relation (6.8) can be proved in the following way. Denote

$$\begin{aligned} p(t) &= a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0, \\ q(t) &= b_m t^m + b_{m-1} t^{m-1} + \dots + b_1 t + b_0, \end{aligned}$$

and look for polynomials h and r in the following form:

$$h(t) = c_{n-m}t^{n-m} + \dots + c_1t + c_0$$

and

$$r(t) = d_{m-1}t^{m-1} + \dots + d_1t + d_0.$$

Then relation (6.8) is equivalent to equality

$$a_n t^n + \dots + a_1 t + a_0 = (b_m t^m + \dots + b_1 t + b_0) (c_{n-m} t^{n-m} + \dots + c_1 t + c_0) + (d_{m-1} t^{m-1} + \dots + d_1 t + d_0).$$

Comparing the corresponding coefficients of both sides we obtain the system of equations

$$\begin{aligned} a_n &= b_m \cdot c_{n-m} \\ a_{n-1} &= b_m c_{n-m-1} + b_{m-1} c_{n-m} \\ &\vdots \\ a_m &= b_m c_0 + b_{m-1} c_1 + \dots + b_0 c_m, \\ a_{m-1} &= b_{m-1} c_0 + b_{m-2} c_1 + \dots + b_0 c_{m-1} + d_{m-1} \\ &\vdots \\ a_1 &= b_1 c_0 + b_0 c_1 + d_1 \\ a_0 &= b_0 c_0 + d_0, \end{aligned}$$

where for $k > n - m, c_k = 0$. From the first equation the value of c_{n-m} can be determined, since $b_m \neq 0$. From the second equation we have c_{n-m-1} , and from the third equation we obtain c_{n-m-2} . Continuing the process the unknown coefficients are obtained in the order $c_{n-m-2}, c_{n-m-3}, \dots, c_0, d_{m-1}, \dots, d_0$.

Example 6.4. Let $p(t) = t^5 + 2t^4 + t^3 + 2t^2 + t + 1$ and $q(t) = t^2 + t + 1$. Then $n = 5$ and $m = 2$, and equation (6.8) has the form

$$t^5 + 2t^4 + t^3 + 2t^2 + t + 1 = (t^2 + t + 1)(c_3t^3 + c_2t^2 + c_1t + c_0) + (d_1t + d_0),$$

and the comparison of the corresponding coefficients gives the system of equations

$$\begin{aligned} 1 &= c_3 \\ 2 &= c_2 + c_3 \\ 1 &= c_1 + c_2 + c_3 \\ 2 &= c_0 + c_1 + c_2 \\ 1 &= c_0 + c_1 + d_1 \\ 1 &= c_0 + d_0. \end{aligned}$$

The solution is obtained as follows:

$$\begin{aligned} c_3 &= 1 \\ c_2 &= 2 - c_3 = 1 \\ c_1 &= 1 - c_2 - c_3 = -1 \\ c_0 &= 2 - c_1 - c_2 = 2 \\ d_1 &= 1 - c_0 - c_1 = 0 \\ d_0 &= 1 - c_0 = -1, \end{aligned}$$

therefore

$$h(t) = t^3 + t^2 - t + 2 \quad \text{and} \quad r(t) = -1.$$

A real (or complex) polynomial is called *prime* if it cannot be represented as the product of two real (or complex) polynomials of lower degree. From properties 1. and 4. we know that in the set of complex polynomials only the linear polynomials are primes. In the set of real polynomials all linear polynomials are primes, a quadratic polynomial is prime if and only if it has no real root, and no real polynomial of degree three or more is prime. This fact follows from the observation that if the degree of a real polynomial p is at least three, then it can be factored either as $(t-t^*)q(t)$ where t^* is a real root of p , or as $(t-t^*)(t-\bar{t}^*)q(t)$, where t^* and \bar{t}^* are complex conjugate roots. In the second case, the quadratic polynomial

$$(t-t^*)(t-\bar{t}^*) = t^2 - t(t^* + \bar{t}^*) + t^* \bar{t}^*$$

is real, since both $t^* + \bar{t}^*$ and $t^* \bar{t}^*$ are real numbers.

6. Every polynomial can be decomposed as the product of powers of primes. For complex polynomials this decomposition has the form

$$p(t) = a_n (t-t_1^*)^{m_1} (t-t_2^*)^{m_2} \dots (t-t_r^*)^{m_r}, \quad (6.9)$$

where $t_1^*, t_2^*, \dots, t_r^*$ are the distinct roots of p with multiplicities m_1, m_2, \dots, m_r . In the case of real polynomials this decomposition can be given as

$$p(t) = a_n (t-t_1^*)^{m_1} \dots (t-t_l^*)^{m_l} q_{l+1}(t)^{m_{l+1}} \dots q_r(t)^{m_r}, \quad (6.10)$$

where t_1^*, \dots, t_l^* polynomials with unit leading coefficients. For $i = l+1, \dots, r$, let z_i and \bar{z}_i denote the distinct real roots with multiplicities m_1, \dots, m_l , and q_{l+1}, \dots, q_r are prime quadratic roots of q_i , then m_i is the common multiplicity of these complex roots. It can be proved that

decompositions (6.9) and (6.10) are unique up to the order of the prime factors. This property is called the *Primary Decomposition Theorem*.

The *greatest common divisor* of polynomials is next introduced. Let p_1, \dots, p_k be real (or complex) polynomials, and assume that for $i = 1, 2, \dots, k$, p_i factored into powers of distinct primes:

$$p_i(t) = a_n^{(i)} q_1(t)^{m_1^{(i)}} \dots q_r(t)^{m_r^{(i)}} q_{r+1}(t)^{m_{r+1}^{(i)}} \dots q_k(t)^{m_k^{(i)}}.$$

Without losing generality we may assume that the prime factors $q_1(t), \dots, q_r(t)$ are common in these polynomials, and there is no other common prime factor. For all $j = 1, 2, \dots, r$, define

$$m_j = \min\{m_j^{(1)}, \dots, m_j^{(k)}\}$$

Then the greatest common divisor is defined as the polynomial

$$q_1(t)^{m_1} \dots q_r(t)^{m_r}.$$

That is, the greatest common divisor is the product of the powers of the distinct common prime factors, and the power of each common prime factor is the smallest value among the powers of this factor in the different polynomials. This definition implies that any common factor of given polynomials is a divisor of their greatest common factor.

We say that polynomials p_1, \dots, p_k are *relative prime*, if they have no common divisor of degree at least one. Relative prime polynomials satisfy the following condition.

7. Assume that polynomials p_1, \dots, p_k are relative prime. Then there exist polynomials q_1, q_2, \dots, q_k such that

$$p_1(t)q_1(t) + p_2(t)q_2(t) + \dots + p_k(t)q_k(t) = 1$$

for all t .

This result can be proved in the following way. Consider the set of polynomials

$$P = \left\{ \sum_{i=1}^k p_i q_i \mid q_i \text{ is a polynomial for all } i \right\}.$$

It is obvious that $p_1, p_2, \dots, p_k \in P$, all linear combinations of p_1, \dots, p_k belong to P , and if $p \in P$, then for any polynomial q , $pq \in P$. Since the degrees of the nonzero elements of P are nonnegative integers, there is a nonzero polynomial $r \in P$ with smallest degree. Since $p \in P$, the degree of r is not larger than the degree of p . First we prove that r is a divisor of all p_i ($i=1, 2, \dots, k$). Using the remainder theorem we see that there are polynomials q_i and r_i such that

$$p_i = r \cdot q_i + r_i,$$

where either $r_i = 0$ or the degree of r_i is smaller than the degree of r . In the first case r is a divisor of p_i , and in the second case

$$r_i = p_i - r \cdot q_i \in P,$$

which contradicts the selection of d (since the degree of r_i is smaller than the degree of r , r must not be the element of P with smallest degree). Therefore only $r_i = 0$ is possible. Since polynomials p_1, \dots, p_k are relatively prime, r has to be a nonzero constant. Hence

$$r = p_1 q_1 + \dots + p_k q_k,$$

and by selecting $q_i^*(t) = \frac{1}{r} q_i(t)$,

$$1 = p_1(t)q_1^*(t) + \dots + p_k(t)q_k^*(t),$$

which proves the assertion.

The main properties of matrix polynomials are discussed next. Assume that V is finitely generated and $A \in L(V)$ is a linear transformation on V . Let \underline{A} be an $n \times n$ matrix representation of A . In Example 2.19 we have proved that the dimension of the vector space of $n \times n$ matrices is n^2 . Consider now matrices

$$\underline{I}, \underline{A}, \underline{A}^2, \dots, \underline{A}^{n^2}.$$

Since we have $n^2 + 1$ matrices in an n^2 -dimensional vector space, they are linearly dependent. Therefore

$$a_{n^2} \underline{A}^{n^2} + \dots + a_2 \underline{A}^2 + a_1 \underline{A} + a_0 \underline{I} = \underline{O}$$

with some scalars a_0, a_1, \dots, a_{n^2} . That is, there is a polynomial p of degree at most n^2 such that

$$p(\underline{A}) = \underline{O}.$$

Consider now the set of all nonzero polynomials p such that $p(\underline{A}) = \underline{O}$. Since their degrees are nonnegative integers, there is a polynomial of least degree with this property.

Definition 6.3. A nonzero polynomial r of least degree such that $r(\underline{A}) = \underline{O}$ is called a *minimal polynomial* of \underline{A} .

Minimal polynomials satisfy the following properties.

Theorem 6.3. Let r be a minimal polynomial of an $n \times n$ matrix \underline{A} . Then

- a) It is unique up to a constant multiplier;
- b) If for a nonzero polynomial $p, p(\underline{A}) = \underline{0}$, then r is a factor of p .

Proof. a) Assume that r_1 and r_2 are both minimal polynomials of \underline{A} and their common degree is m . The remainder theorem implies that there is a scalar a and a polynomial r^* of degree at most $m-1$ such that

$$r_1 = a \cdot r_2 + r^*.$$

Then

$$r^*(\underline{A}) = r_1(\underline{A}) - ar_2(\underline{A}) = \underline{0} - a \cdot \underline{0} = \underline{0},$$

which is impossible for $r^* \neq 0$ because of the minimality of the degree of r_1 and r_2 . Thus $r^* = 0$, and r_1 is a constant multiple of r_2 .

b) The remainder theorem implies that

$$p = r \cdot q + r^*$$

where q is a polynomial and the degree of r^* is at most $m-1$. Then

$$r^*(\underline{A}) = p(\underline{A}) - r(\underline{A})q(\underline{A}) = \underline{0} - \underline{0} \cdot q(\underline{A}) = \underline{0}.$$

Similarly to the previous part of the proof we see that this is possible only if r^* is the zero polynomial, that is, r is a divisor of p .

✦

Example 6.5. We will now find a minimal polynomial of the 2×2 matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Constant nonzero polynomials have the general form $p_0(t) = a_0$ with some scalar $a_0 \neq 0$. Then $p_0(\underline{A}) = a_0 \underline{I} \neq \underline{O}$. Linear polynomials have the general form $p_1(t) = a_1 t + a_0$. Then

$$p_1(\underline{A}) = a_1 \underline{A} + a_0 \underline{I} = a_1 \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} + a_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is zero if and only if

$$a_1 + a_0 = a_1 = 2a_1 = 2a_1 + a_0 = 0.$$

Since these equations are satisfied only if $a_1 = a_0 = 0$, the minimal polynomials are at least quadratic. The general form of a quadratic polynomial is $p_2(t) = a_2 t^2 + a_1 t + a_0$, so $p_2(\underline{A}) = \underline{O}$ if and only if with some scalars a_2, a_1, a_0 ,

$$a_2 \begin{pmatrix} 3 & 3 \\ 6 & 6 \end{pmatrix} + a_1 \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} + a_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This matrix equation is equivalent to relations

$$3a_2 + a_1 + a_0 = 0$$

$$3a_2 + a_1 = 0$$

$$6a_2 + 2a_1 = 0$$

$$6a_2 + 2a_1 + a_0 = 0.$$

The general solution of these equations is the following: $a_0 = 0, a_1 = -3a_2$, where a_2 is arbitrary. Hence (by the selection of $a_2 = 1$),

$$p_2(t) = t^2 - 3t$$

is a minimal polynomial of \underline{A} .

◆

6.4 The Construction of Invariant Subspaces

We start this section with a simple result, which will play an important role in the construction of invariant subspaces and special matrix forms. Let V be a vector space and assume that for $i = 1, 2, \dots, k$, transformations $B_i \in L(V)$ satisfy the following properties:

a) $B_1 + \dots + B_k = I$

b) $B_i B_j = O$ for $i \neq j$, where O is the zero mapping such that $O(\underline{x}) = \underline{0}$ for all $\underline{x} \in V$.

We will first show that for all i , $B_i^2 = B_i$. Such linear transformations are called *idempotent*. Properties (a) and (b) imply that

$$B_i = B_i \cdot I = B_i \cdot (B_1 + B_2 + \dots + B_k) = B_i^2 + \sum_{j \neq i} B_i B_j = B_i^2,$$

since all terms in the summation are zeros. Define V_i as the range of B_i :

$$V_i = R(B_i) = \{ \underline{x} \mid \underline{x} = B_i(\underline{z}) \text{ with some } \underline{z} \in V \}.$$

Theorem 6.4. Subspace V is the direct sum of V_1, V_2, \dots, V_k :

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_k.$$

Proof. Using Theorem 2.16 we have to prove first that each $\underline{v} \in V$ can be written as the sum $\underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_k$, where for $i = 1, 2, \dots, k, \underline{v}_k \in V_k$. And second, we have to show that if for some elements $\underline{v}_i \in V_i (i = 1, 2, \dots, k), \underline{0} = \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_k$, then necessarily $\underline{v}_1 = \underline{v}_2 = \dots = \underline{v}_k = \underline{0}$.

Let $\underline{v} \in V$ be arbitrary. Then

$$\underline{v} = I(\underline{v}) = (B_1 + B_2 + \dots + B_k)(\underline{v}) = B_1(\underline{v}) + B_2(\underline{v}) + \dots + B_k(\underline{v}).$$

Notice that $B_i(\underline{v}) \in V_i$ for all i , therefore the first statement is verified.

Assume next that for some $\underline{v}_i \in V_i (i = 1, 2, \dots, k)$,

$$\underline{0} = \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_k.$$

The definition of subspaces V_i implies that for all i , there is a $\underline{w}_i \in V$ such that $\underline{v}_i = B_i(\underline{w}_i)$. From this equation we see that for all i ,

$$\underline{0} = B_i(\underline{0}) = B_i(\underline{v}_1) + B_i(\underline{v}_2) + \dots + B_i(\underline{v}_k).$$

If $j \neq i$, then

$$B_i(\underline{v}_j) = B_i(B_j(\underline{w}_j)) = (B_i B_j)(\underline{w}_j) = \underline{0}$$

as the consequence of assumption (b), therefore $B_i(\underline{v}_j) = \underline{0}$, which implies that for all i ,

$$\underline{v}_i = B_i(\underline{w}_i) = B_i^2(\underline{w}_i) = B_i(B_i(\underline{w}_i)) = B_i(\underline{v}_i) = \underline{0},$$

which completes the proof. +

Let now $A \in L(V)$ be a linear transformation and assume that V is finitely generated. Assume furthermore that the minimal polynomial r of A is factored into powers of distinct primes:

$$r(t) = a p_1(t)^{m_1} p_2(t)^{m_2} \dots p_l(t)^{m_l},$$

where the leading coefficient is a . For $i = 1, 2, \dots, l$, define

$$r_i(t) = \frac{r(t)}{p_i(t)^{m_i}}. \quad (6.11)$$

These polynomials are relative prime, and therefore Property 7 discussed in the previous section implies that there are polynomials $q_i(t)$ such that

$$1 = r_1(t)q_1(t) + r_2(t)q_2(t) + \dots + r_l(t)q_l(t). \quad (6.12)$$

Introduce notation

$$f_i(t) = r_i(t)q_i(t) \quad (i = 1, 2, \dots, l),$$

and consider transformations $B_i = f_i(A)$. We will next show that these mappings satisfy all conditions of Theorem 6.4.

Theorem 6.5.

- (i) $I = B_1 + B_2 + \dots + B_l$;
- (ii) $B_i B_j = O$ for $j \neq i$;
- (iii) define $V_i = R(B_i)$ for all i , then

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_i;$$

$$(iv) \text{ for all } i, V_i = N\left(p_i(A)^{m_i}\right).$$

Proof. The proof consists of several steps.

(i) From equation (6.12) and the definition of $f_i(t)$ we have

$$1 = f_1(t) + f_2(t) + \dots + f_i(t).$$

Substitution of A into this equation shows that

$$I = f_1(A) + f_2(A) + \dots + f_i(A),$$

and since $B_i = f_i(A)$, the assertion is valid.

(ii) If $i \neq j$, then

$$B_i B_j = f_i(A) f_j(A) = r_i(A) q_i(A) r_j(A) q_j(A) = 0,$$

since $r_i(t) r_j(t)$ is a multiple of $r(t)$, and $r(A) = 0$.

(iii) Properties (i), (ii) and Theorem 6.4 imply that V is the direct sum of V_1, V_2, \dots, V_i .

(iv) First we show that $V_i \subseteq N\left(p_i(A)^{m_i}\right)$. Let $\underline{v} \in V_i$ be an arbitrary element. Then $\underline{v} = B_i(\underline{w})$ with some $\underline{w} \in V$, and the definition of f_i implies that

$$\begin{aligned} p_i(A)^{m_i}(\underline{v}) &= p_i(A)^{m_i}(B_i(\underline{w})) = p_i(A)^{m_i} f_i(A)(\underline{w}) \\ &= p_i(A)^{m_i} r_i(A) q_i(A)(\underline{w}) = r(A) q_i(A)(\underline{w}) = \underline{0}, \end{aligned}$$

since $r(A) = 0$. Therefore $\underline{v} \in N\left(p_i(A)^{m_i}\right)$.

Next we prove that $V_i \supseteq N(p_i(A)^{m_i})$. Assume that $\underline{v} \in N(p_i(A)^{m_i})$. From property (iii) we know that \underline{v} can be expressed as

$$\underline{v} = \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_l \quad (6.13)$$

where $\underline{v}_i \in V_i$ for all i . Since $V_i = R(B_i)$, there are elements \underline{w}_i such that $\underline{v}_i = B_i(\underline{w}_i)$, therefore

$$\begin{aligned} \underline{0} &= p_i(A)^{m_i}(\underline{v}) = p_i(A)^{m_i}(B_1(\underline{w}_1) + B_2(\underline{w}_2) + \dots + B_l(\underline{w}_l)) \\ &= (p_i(A)^{m_i} B_1)(\underline{w}_1) + (p_i(A)^{m_i} B_2)(\underline{w}_2) + \dots + (p_i(A)^{m_i} B_l)(\underline{w}_l) \\ &= (p_i(A)^{m_i} f_1(A))(\underline{w}_1) + (p_i(A)^{m_i} f_2(A))(\underline{w}_2) + \dots + (p_i(A)^{m_i} f_l(A))(\underline{w}_l). \end{aligned}$$

Notice that for $j = 1, 2, \dots, l$, the j^{th} term can be written as

$$(f_j(A)p_i(A)^{m_i})(\underline{w}_j) = (B_j p_i(A)^{m_i})(\underline{w}_j) = B_j(p_i(A)^{m_i}(\underline{w}_j))$$

which is in $R(B_j) = V_j$. Hence $\underline{0}$ is decomposed as the sum of l elements, where the j^{th} element belongs to V_j for $j = 1, 2, \dots, l$. Then property (iii) and Theorem 2.16 imply that each term is necessarily zero, that is, for all j ,

$$\underline{0} = B_j(p_i(A)^{m_i}(\underline{w}_j)) = (B_j p_i(A)^{m_i})(\underline{w}_j) \quad (6.14)$$

For all $j \neq i$, $p_i(t)^{m_i}$ and $p_j(t)^{m_j}$ are relative prime, therefore Property 7 of the previous section implies that there exist polynomials $s_i(t)$ and $s_j(t)$ such that

$$1 = s_i(t)p_i(t)^{m_i} + s_j(t)p_j(t)^{m_j}.$$

Substituting A into this equation and applying both sides to the element $B_j(\underline{w}_j)$ we have

$$\begin{aligned} B_j(\underline{w}_j) &= (s_i(A)p_i(A)^{m_i})(B_j(\underline{w}_j)) + (s_j(A)p_j(A)^{m_j})(B_j(\underline{w}_j)) \\ &= s_i(A)((B_j p_i(A)^{m_i})(\underline{w}_j)) + s_j(A)((B_j p_j(A)^{m_j})(\underline{w}_j)) = \underline{0}, \end{aligned}$$

since $B_j = f_j(A)$ commutes with $p_i(A)^{m_i}$ and $p_j(A)^{m_j}$, relation (6.14) holds, and $B_j p_j(A)^{m_j} = q_j(A)r(A) = 0$. And finally, equation (6.13) implies that

$$\begin{aligned} \underline{v} &= \underline{v}_1 + \underline{v}_2 + \dots + \underline{v}_l = B_1(\underline{w}_1) + B_2(\underline{w}_2) + \dots + B_l(\underline{w}_l) \\ &= B_i(\underline{w}_i) \in R(B_i) = V_i, \end{aligned}$$

since all other terms are equal to zero. Hence $\underline{v} \in V_i$, that is, $N(p_i(A)^{m_i}) \subseteq V_i$, and thus, the proof is complete.

♣

6.5 Diagonal and Triangular Forms

In this section we will always assume that V is finitely generated. The main result of the previous section will be first applied to answer the question of when a basis of V can be chosen that consists of eigenvectors of A . In such cases the matrix representation of A in this basis becomes diagonal. This important fact can be proved as follows. Assume that $\dim(V) = n$, and

$B = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ is a basis of V consisting of eigenvectors of A . For all $i = 1, 2, \dots, n$, the eigenvector equation

$$A(\underline{x}_i) = \lambda_i \underline{x}_i$$

is satisfied, where λ_i is the eigenvalue associated to eigenvector \underline{x}_i . That is,

$$A(\underline{x}_i) = 0 \cdot \underline{x}_1 + \dots + 0 \cdot \underline{x}_{i-1} + \lambda_i \underline{x}_i + 0 \cdot \underline{x}_{i+1} + \dots + 0 \cdot \underline{x}_n,$$

and equation (5.14) implies that the matrix representation of transformation A in basis B is diagonal:

$$\underline{A}_{B,B} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Definition 6.4. A linear transformation $A \in L(V)$ is called *diagonalizable* if there is a basis of V consisting of eigenvectors of A .

In the above discussion we have derived that the matrix representation of a diagonalizable transformation A in basis B is diagonal. Assume first that A has n distinct eigenvalues, then Theorem 6.1 implies that the associated eigenvectors form a basis in V , hence, A is diagonalizable. This condition is only sufficient for a transformation being diagonalizable. The case of the identity transformation shows that transformations may be diagonalizable even if they have multiple eigenvalues. For the $n \times n$ identity transformation, $\lambda = 1$ is the only eigenvalue with multiplicity n , and any arbitrary vector \underline{x} is an associated eigenvector, since $I(\underline{x}) = \underline{x} = 1 \cdot \underline{x}$. Therefore any basis of V consists of eigenvectors.

Example 6.6. Consider again the 3×3 matrix

$$\underline{A} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

which was the subject of Example 6.2. We have shown there that the eigenvalues of \underline{A} are $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = 6$ with associated eigenvectors

$$\underline{x}_1 = \begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix}, \quad \underline{x}_2 = \begin{pmatrix} 9 \\ -5 \\ 2 \end{pmatrix}, \quad \text{and} \quad \underline{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Since the eigenvalues are different, the eigenvectors form a basis in R^3 . Introduce matrix

$$\underline{T} = \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix},$$

then in the new basis $\{\underline{x}_1, \underline{x}_2, \underline{x}_3\}$, matrix \underline{A} will have the diagonal form

$$\underline{T}^{-1} \underline{A} \underline{T} = \begin{pmatrix} -\frac{1}{6} & -\frac{7}{42} & \frac{7}{21} \\ 6 & 42 & 21 \\ 0 & -\frac{6}{42} & \frac{3}{21} \\ \frac{1}{6} & \frac{19}{42} & \frac{8}{21} \end{pmatrix} \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

A necessary and sufficient condition is next given for a matrix being diagonalizable.

Theorem 6.6. A linear transformation A is diagonalizable if and only if the minimal polynomial of A can be factored as

$$r(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_k),$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct values.

Proof. Assume first that A is diagonalizable, and let $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ be a basis of V consisting of eigenvectors of A . Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be the distinct eigenvalues of A , which means that any other eigenvalue of A coincides with some λ_i ($i = 1, 2, \dots, k$). Define

$$p(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_k).$$

We will show that p is the minimal polynomial of A . Since \underline{x}_j is an eigenvector associated to an eigenvalue λ_i ($i = 1, 2, \dots, k$), the eigenvector equation implies that

$$\underline{0} = A(\underline{x}_j) - \lambda_i \underline{x}_j = (A - \lambda_i I)(\underline{x}_j)$$

Multiply both sides by polynomial $\prod_{i \neq i} (A - \lambda_i I)$ to see that $p(A)\underline{x}_j = \underline{0}$.

If $\underline{x} \in V$ is arbitrary, then

$$\underline{x} = c_1 \underline{x}_1 + \dots + c_n \underline{x}_n$$

with some scalars c_i ($i = 1, 2, \dots, n$), and

$$p(A)(\underline{x}) = p(A)(c_1 \underline{x}_1 + \dots + c_n \underline{x}_n) = c_1 \cdot p(A)(\underline{x}_1) + \dots + c_n p(A)(\underline{x}_n) = \underline{0}.$$

Hence $p(A) = 0$. Let $r(A)$ denote the minimal polynomial of A . If $p \neq r$, then Theorem 6.3 implies that r is a factor of p , therefore r is obtained from p by deleting at least one prime factor. Assume that factor $t - \lambda_i$ is among the deleted prime factors, then

$$r(t) = (t - \lambda_{i_1}) \dots (t - \lambda_{i_r}),$$

where the values $\lambda_{i_1}, \dots, \lambda_{i_r}$ differ from λ_i . Let \underline{x}_i be an eigenvector associated to λ_i , then

$$\begin{aligned} r(A)(\underline{x}_i) &= (A - \lambda_{i_1} I) \dots (A - \lambda_{i_r} I)(\underline{x}_i) \\ &= (A - \lambda_{i_1} I) \dots (A - \lambda_{i_{r-1}} I)((\lambda_i - \lambda_{i_r}) \underline{x}_i) \\ &= \dots = (\lambda_i - \lambda_{i_1}) \dots (\lambda_i - \lambda_{i_{r-1}})(\lambda_i - \lambda_{i_r}) \underline{x}_i \neq 0. \end{aligned}$$

That is, $r(A)$ is nonzero contradicting the assumption that r is the minimal polynomial of A .

Assume next that the minimal polynomial of A has the form

$$r(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_k)$$

with distinct values of $\lambda_1, \lambda_2, \dots, \lambda_k$. From Theorem 6.5 we know that

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_k,$$

where for $i = 1, 2, \dots, k$,

$$V_i = N((A - \lambda_i I)).$$

Let $\underline{x}_{i1}, \underline{x}_{i2}, \dots, \underline{x}_{in_i}$ be a basis of V_i . First we show that all basis elements are eigenvectors of A . The definition of V_i implies that for all j ,

$$(A - \lambda_i I)(\underline{x}_{ij}) = \underline{0},$$

that is,

$$A(\underline{x}_{ij}) = \lambda_i \underline{x}_{ij},$$

which shows that \underline{x}_{ij} is an eigenvector of A associated to λ_i . Since V is the direct sum of subspaces V_1, V_2, \dots, V_k , the system

$$\{\underline{x}_{11}, \dots, \underline{x}_{1n_1}, \underline{x}_{21}, \dots, \underline{x}_{2n_2}, \dots, \underline{x}_{k1}, \dots, \underline{x}_{kn_k}\}$$

is a basis of V , which completes the proof. ♣

For an $n \times n$ matrix \underline{A} the assertion of the theorem can be restated as follows. If \underline{A} is real, then there exists a nonsingular real matrix \underline{T} such that $\underline{T}^{-1} \underline{A} \underline{T}$ is diagonal if and only if all roots of the minimal polynomial of \underline{A} are real and distinct. If \underline{A} is complex, then there is a nonsingular (maybe complex) matrix \underline{T} such that $\underline{T}^{-1} \underline{A} \underline{T}$ is diagonal if and only if the roots (which maybe real or complex) of the minimal polynomial of \underline{A} are distinct.

The condition of the theorem can fail under two circumstances. One is that the minimal polynomial has multiple roots; the other is that the minimal polynomial cannot be factored into linear factors. For complex polynomials the second case never occurs, and for real polynomials it occurs if and only if the minimal polynomial has at least one complex root. In such cases we may consider the matrix representation as complex and diagonalize the

matrix with complex arithmetic. The first case however cannot be handled so easily. The best way is to find a class of matrices of a special structure that is as close to the set of diagonal matrices as possible and for all $A \in L(V)$, there is a basis such that the matrix representation of A has that special structure.

We start our analysis with examining a special group of linear transformations.

Definition 6.5. Let $A \in L(V)$ be a linear transformation. It is called *nilpotent* if $A^k = O$ for some positive integer k . The smallest such integer k is called the *degree* of A .

Example 6.7. Consider the 3×3 matrix

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

Simple calculation shows that

$$\underline{A}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \underline{A}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let A be a linear transformation the matrix representation of which is \underline{A} in a given basis. Then A is nilpotent, and its degree is 3. ♦

We will first show that nilpotent matrices are singular. Let \underline{A} be a nonsingular matrix. The repeated application of property (iv) of Theorem

5.8 implies that \underline{A}^k is also nonsingular for all $k \geq 2$ and $(\underline{A}^k)^{-1} = (\underline{A}^{-1})^k$. Therefore nilpotent matrices must be singular, since $\underline{0}$ is a singular matrix.

The following property of nilpotent transformations will be useful later in this section.

Theorem 6.7. Let V be finite dimensional and let $A \in L(V)$ be nilpotent. Then there is a basis $B = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ of V such that $A(\underline{x}_1) = \underline{0}$, and for all $i = 2, \dots, n$, $A(\underline{x}_i)$ belongs to the subspace $V(\underline{x}_1, \dots, \underline{x}_{i-1})$ generated by the previous basis elements $\underline{x}_1, \dots, \underline{x}_{i-1}$.

Proof. The theorem is proved by induction. Let n denote the dimension of the vector space V . Notice first that A is singular, therefore there is an $\underline{x}_1 \neq \underline{0}$ such that $A(\underline{x}_1) = \underline{0}$. As the inductive hypothesis assumes that for an $i < n$, we have found linearly independent elements $\underline{x}_1, \dots, \underline{x}_i$ satisfying the conditions of the theorem. The subspace $V(\underline{x}_1, \dots, \underline{x}_i)$ generated by these elements is a nontrivial subspace since it is generated by less than n elements. If $R(A) \subset V(\underline{x}_1, \dots, \underline{x}_i)$, then let \underline{x}_{i+1} be any element that does not belong to $V(\underline{x}_1, \dots, \underline{x}_i)$. If $R(A) \not\subset V(\underline{x}_1, \dots, \underline{x}_i)$, then there exists a positive integer l such that $R(A^l) \not\subset V(\underline{x}_1, \dots, \underline{x}_i)$ but $R(A^{l+1}) \subset V(\underline{x}_1, \dots, \underline{x}_i)$, since with some k , $A^k = \underline{0}$, and $R(A^k) = \{\underline{0}\} \subset V(\underline{x}_1, \dots, \underline{x}_i)$. Select now \underline{x}_{i+1} such that $\underline{x}_{i+1} \in R(A^l)$ but $\underline{x}_{i+1} \notin V(\underline{x}_1, \dots, \underline{x}_i)$. Then vectors $\underline{x}_1, \dots, \underline{x}_i, \underline{x}_{i+1}$ are linearly independent and $A(\underline{x}_{i+1}) \in V(\underline{x}_1, \dots, \underline{x}_i)$. Thus the proof is complete. ♣

The construction of the basis $B = \{\underline{x}_1, \dots, \underline{x}_n\}$ implies that the matrix representation of a nilpotent transformation A has the special form that all diagonal elements as well as all elements under the diagonal are equal to zero. This is a special upper triangular matrix. We will next prove that all linear transformations of finitely generated vector spaces can be represented by an upper triangular matrix in a special basis.

Theorem 6.8. Let $A \in L(V)$ be a linear transformation. Assume that V is finitely generated and the minimal polynomial of A is given as

$$r(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_l)^{m_l}.$$

Then there is a basis B of V such that the matrix representation of A in this basis has the special form

$$\underline{A}_{B,B} = \begin{pmatrix} \underline{A}_1 & & & \mathbf{O} \\ & \underline{A}_2 & & \\ & & \ddots & \\ \mathbf{O} & & & \underline{A}_l \end{pmatrix}, \quad (6.15)$$

where for $i = 1, 2, \dots, l$, \underline{A}_i is an upper triangular $n_i \times n_i$ matrix with some $n_i \geq m_i$, and all diagonal elements of \underline{A}_i are equal to λ_i .

Proof. As in the earlier theorems, define $V_i = N((A - \lambda_i I)^{m_i})$, and let $n_i = \dim(V_i)$. In Theorem 6.5 we have proved that $V = V_1 \oplus V_2 \oplus \dots \oplus V_l$, and from Theorem 6.2 we know that all subspaces V_i are A -invariant. If for $i = 1, 2, \dots, l$, $B_i = \{\underline{x}_{i1}, \dots, \underline{x}_{in_i}\}$ is a basis of V_i , then $B = \{\underline{x}_{11}, \dots, \underline{x}_{1n_1}, \underline{x}_{21}, \dots, \underline{x}_{2n_2}, \dots, \underline{x}_{l1}, \dots, \underline{x}_{ln_l}\}$ is a basis of V and the matrix representation of A in this basis has the block-diagonal form (6.15) where the size of each diagonal matrix \underline{A}_i is $n_i \times n_i$. In order to complete the proof we have to show that each diagonal block \underline{A}_i has a special upper triangular form with diagonal elements λ_i . Define $A_i \in L(V_i)$ as a linear transformation on V_i such that for all $\underline{x} \in V_i$, $A_i(\underline{x}) = (A - \lambda_i I)(\underline{x})$. The definition of subspace V_i implies that $(A - \lambda_i I)^{m_i} = \mathbf{O}$, that is, $A - \lambda_i I$ is

nilpotent. Then Theorem 6.7 implies that there is a basis $B_i = \{\underline{x}_{i1}, \underline{x}_{i2}, \dots, \underline{x}_{in_i}\}$ of V_i such that

$$(A - \lambda_i I)(\underline{x}_{i1}) = \underline{0},$$

and for $j = 2, \dots, n_i$,

$$(A - \lambda_i I)(\underline{x}_{ij}) \in V(\underline{x}_{i1}, \dots, \underline{x}_{i,j-1})$$

These relations can be rewritten as

$$A(\underline{x}_{i1}) = \lambda_i \underline{x}_{i1}$$

and for $j = 2, \dots, n_i$,

$$A(\underline{x}_{ij}) = \lambda_i \underline{x}_{ij} + a_{1j} \underline{x}_{i1} + \dots + a_{j-1,j} \underline{x}_{i,j-1}.$$

Hence the matrix representation of A in subspace V_i has the special form

$$\underline{A}_i = \begin{pmatrix} \lambda_i & a_{12} & a_{13} & \dots & a_{1n} \\ & \lambda_i & a_{23} & \dots & a_{2n} \\ & & \ddots & & \vdots \\ & & & & a_{n-1,n} \\ \mathbf{0} & & & & \lambda_i \end{pmatrix},$$

which completes the proof. ♣

For $n \times n$ matrices this theorem can be restated in the following way. For an arbitrary $n \times n$ real or complex matrix there is a nonsingular matrix \underline{T} such that $\underline{T}^{-1} \underline{A} \underline{T}$ has the special form (6.15). If the minimal polynomial has only real roots, then $\underline{T}, \underline{T}^{-1}$, and all diagonal blocks \underline{A}_i are real. In the

case of complex roots the corresponding diagonal blocks \underline{A}_i , as well as matrices \underline{T}^{-1} and \underline{T} might have complex elements.

Corollary. Let $r(t)$ and $\varphi(t)$ be the minimal polynomial and characteristic polynomial of A , respectively. Then

- a) r is a factor of φ ;
- b) all roots of φ are roots of r as well;
- c) $\varphi(A) = O$.

Proof. From the matrix representation (6.15) we see that the characteristic polynomial of A (which is independent of the selected basis) is the following:

$$\begin{aligned}\varphi(t) &= (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \dots (t - \lambda_l)^{n_l} \\ &= (t - \lambda_1)^{n_1 - m_1} (t - \lambda_2)^{n_2 - m_2} \dots (t - \lambda_l)^{n_l - m_l} r(t) = q(t)r(t),\end{aligned}$$

from which assertions (a) and (b) follow immediately, furthermore

$$\varphi(A) = (A - \lambda_1 I)^{n_1 - m_1} (A - \lambda_2 I)^{n_2 - m_2} \dots (A - \lambda_l I)^{n_l - m_l} r(A) = O$$

since $r(A) = O$.

+

Assertion c) is known as the famous *Cayley-Hamilton* theorem, which can be restated by saying that if we substitute any arbitrary $n \times n$ matrix into its characteristic polynomial we always obtain the zero matrix as the result.

6.6 The Jordan Canonical Form

In the previous section we have proved that any real or complex square matrix can be transformed into a block-diagonal form, where each diagonal block is an upper triangular matrix (that might have complex elements). In this section we will present a further refinement of this result by showing

that the upper triangular blocks can be made very specially: the diagonal elements equal to an eigenvalue of the matrix, all elements just above the diagonal are equal to one, and all other elements equal zero. Our analysis will be based on the following concept.

Definition 6.6. Let $A \in L(V)$, and let \underline{x} be a nonzero element of V . The subspace generated by the system $\{\underline{x}, A(\underline{x}), A^2(\underline{x}), \dots\}$ is called a *cyclic subspace*, and is denoted by $V(\underline{x}, A)$.

We show first that subspace $V(\underline{x}, A)$ is A -invariant for all $\underline{x} \in V$. Let $\underline{z} \in V(\underline{x}, A)$ be an arbitrary element. Then Theorem 2.4 implies that

$$\underline{z} = a_1 A^{i_1}(\underline{x}) + a_2 A^{i_2}(\underline{x}) + \dots + a_l A^{i_l}(\underline{x})$$

with some integers $l \geq 1, 0 \leq i_1 < i_2 < \dots < i_l$, and scalars a_1, a_2, \dots, a_l . Then the linearity of A implies that

$$A(\underline{z}) = a_1 A^{i_1+1}(\underline{x}) + a_2 A^{i_2+1}(\underline{x}) + \dots + a_l A^{i_l+1}(\underline{x}) \in V(\underline{x}, A).$$

Notice next that if A is nilpotent with degree k , then $V(\underline{x}, A)$ is generated by $\underline{x}, A(\underline{x}), A^2(\underline{x}), \dots, A^{k-1}(\underline{x})$. Since $A^{k-1} \neq O$, there exists an $\underline{x} \in V$ such that $A^{k-1}(\underline{x}) \neq 0$. We will first prove that in this case the elements $\underline{x}, A(\underline{x}), A^2(\underline{x}), \dots, A^{k-1}(\underline{x})$ are linearly independent. Assume that a linear combination of these elements is zero;

$$a_0 \underline{x} + a_1 A(\underline{x}) + a_2 A^2(\underline{x}) + \dots + a_{k-1} A^{k-1}(\underline{x}) = \underline{0}.$$

Assume that i is the smallest subscript such that $a_i \neq 0$. Apply A^{k-i-1} on both sides of this equation to have

$$a_i A^{k-1}(\underline{x}) = \underline{0}.$$

Since $A^{k-1}(\underline{x}) \neq \underline{0}$, this is possible only if $a_i = 0$, which contradicts the assumption that $a_i \neq 0$. Hence the linear independence of elements $\underline{x}, A(\underline{x}), A^2(\underline{x}), \dots, A^{k-1}(\underline{x})$ is verified, and therefore they form a basis of the subspace $V(\underline{x}, A)$.

Assume again that A is nilpotent on V with degree k . Let $V_1 = R(A)$, and define the linear transformation A_1 on V_1 such that for all $\underline{x} \in V_1$, $A_1(\underline{x}) = A(\underline{x})$. This transformation is called the *restriction* of A to subspace V_1 . We can easily prove that A_1 is also nilpotent, and its degree is $k-1$. First, let $\underline{x} \in V_1$ be an arbitrary element, then $\underline{x} = A(\underline{z})$ with some $\underline{z} \in V$. Then

$$A_1^{k-1}(\underline{x}) = A_1^{k-1}(A(\underline{z})) = A^{k-1}(A(\underline{z})) = A^k(\underline{z}) = \underline{0},$$

that is, the degree of A_1 is not greater than $k-1$. Notice next that there is an $\underline{x} \in V$ such that $A^{k-1}(\underline{x}) \neq \underline{0}$, then by choosing $\underline{y} = A(\underline{x}) \in V_1$,

$$A_1^{k-2}(\underline{y}) = A_1^{k-2}(A(\underline{x})) = A^{k-2}(A(\underline{x})) = A^{k-1}(\underline{x}) \neq \underline{0}$$

showing that the degree of A_1 is not less than $k-1$. Hence the degree of A_1 equals $k-1$.

The main result of this section that will be used in obtaining the Jordan form can be formulated as follows.

Theorem 6.9. Assume that $A \in L(V)$ is nilpotent with degree k . Assume that $A^{k-1}(\underline{u}) \neq \underline{0}$. Then there exists an A -invariant subspace $V_1 \subset V$ such that

$$V = V(\underline{u}, A) \oplus V_1 \quad (6.16)$$

Proof. Finite induction will be used with respect to the degree of A . Assume first that $k = 1$. Then $A(\underline{x}) = \underline{0}$ for all $\underline{x} \in V$. Select an arbitrary element $\underline{u} \neq \underline{0}$. If $\dim(V) = n$, then there exist elements $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{n-1}$ such that $\{\underline{u}, \underline{u}_1, \underline{u}_2, \dots, \underline{u}_{n-1}\}$ is a basis of V . Consider the subspace $V_1 = V(\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{n-1})$ which is obviously an A -invariant subspace and the fact that $V(\underline{u}, A) = V(\underline{u})$ implies that (6.16) holds. Assume next that the assertion is true for nilpotent transformations with degrees up to $k - 1$, and consider an A with degree k . The subspace $R(A) \subset V$ is A -invariant, and the restriction of A to $R(A)$ is also nilpotent with degree $k - 1$. Let $\underline{u} \in R(A)$ be an element such that $A^{k-2}(\underline{u}) \neq \underline{0}$. Then by the inductive hypotheses there is an A -invariant subspace V_0 such that

$$R(A) = V(A(\underline{v}), A) \oplus V_0,$$

where $\underline{v} \in V$ is selected as $\underline{u} = A(\underline{v})$. Notice that the first term is generated by the elements $A(\underline{v}), A^2(\underline{v}), \dots, A^{k-1}(\underline{v})$, which are linearly independent. Consider the set

$$\bar{V}_0 = \{\underline{v} \in V \mid A(\underline{v}) \in V_0\}.$$

We will next prove that it is an A -invariant subspace. First we show that \bar{V}_0 is a subspace of V . Let a_1, a_2 be two scalars and $\underline{u}_1, \underline{u}_2 \in \bar{V}_0$. Then

$$A(a_1 \underline{u}_1 + a_2 \underline{u}_2) = a_1 A(\underline{u}_1) + a_2 A(\underline{u}_2) \in V_0,$$

that is, $a_1 \underline{u}_1 + a_2 \underline{u}_2 \in \overline{V}_0$. To prove that \overline{V}_0 is A -invariant, consider an arbitrary element $\underline{u} \in \overline{V}_0$. Then $A(\underline{u}) \in V_0$, and since V_0 is A -invariant, $A(A(\underline{u})) \in V_0$ showing that $A(\underline{u}) \in \overline{V}_0$. Hence, \overline{V}_0 is also A -invariant. We will next show that $V(\underline{v}, A) \cup \overline{V}_0$ generates the entire space V . Let $\underline{x} \in V$ be arbitrary. Since $A(\underline{x}) \in R(A)$, we have

$$A(\underline{x}) = \sum_{i=1}^{k-1} a_i A^i(\underline{v}) + \underline{w},$$

where $\underline{w} \in V_0$ and a_1, a_2, \dots, a_{k-1} are scalars. This equation can be rewritten as

$$A\left(\underline{x} - \sum_{i=1}^{k-1} a_i A^{i-1}(\underline{v})\right) = \underline{w},$$

therefore the definition of \overline{V}_0 implies that

$$\underline{x} - \sum_{i=1}^{k-1} a_i A^{i-1}(\underline{v}) \in \overline{V}_0.$$

That is, \underline{x} can be expressed as the sum of an element of $V(\underline{v}, A)$ and an element of \overline{V}_0 .

Unfortunately, in completing the proof, we cannot select $\underline{u} = \underline{v}$ and $V_1 = \overline{V}_0$, since $V(\underline{v}, A)$ and \overline{V}_0 might have nonzero common elements. We can show however that $V(\underline{v}, A) \cap V_0 = \{0\}$. If

$$\underline{x} \in V(\underline{v}, A) \cap V_0,$$

then

$$A(\underline{x}) \in V(A(\underline{v}), A) \cap V_0,$$

therefore $A(\underline{x}) = \underline{0}$. Since $\underline{x} \in V(\underline{v}, A)$, with some scalars c_0, c_1, \dots, c_{k-1} ,

$$\underline{x} = c_0 \underline{v} + c_1 A(\underline{v}) + \dots + c_{k-1} A^{k-1}(\underline{v}).$$

Assume that for an index $i (i < k-1), c_i \neq 0$. Let i denote the smallest such value. Apply A^{k-i-1} to both sides of this equality to have

$$\underline{A}^{k-i-1}(\underline{x}) = c_i \underline{A}^{k-1}(\underline{v}),$$

which is an obvious contradiction, since the left hand side is zero, and the right-hand side is nonzero. Consequently, \underline{x} is a scalar-multiple of $A^{k-1}(\underline{v})$, and since V_0 is A -invariant,

$$\underline{x} \in V(A(\underline{v}), A) \cap V_0 = \{\underline{0}\},$$

hence $\underline{x} = \underline{0}$.

Notice finally, that both $V(\underline{v}, A) \cap \overline{V}_0$ and V_0 are subspaces of \overline{V}_0 , and their intersection is only the zero element. Therefore a basis of \overline{V}_0 can be obtained as

$$\{B_1, B_2, \underline{v}_1, \underline{v}_2, \dots, \underline{v}_l\},$$

where B_1 is a basis of $V(\underline{v}, A) \cap \overline{V}_0$, B_2 is a basis of V_0 , and $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_l$ are the elements which complete the linearly independent system $\{B_1, B_2\}$ into a basis of \overline{V}_0 . Define now

$$V_1 = V(B_2, \underline{v}_1, \dots, \underline{v}_l),$$

then obviously

$$V = V(\underline{v}, A) \oplus V_1.$$

The definition of V_1 implies that $V_0 \subseteq V_1 \subseteq \bar{V}_0$ and for all $\underline{x} \in V_1$, $A(\underline{x}) \in V_0$. We have assumed that V_0 is an A -invariant subspace, therefore for all $\underline{x} \in V_1$, $A(\underline{x}) \in V_1$ proving that V_1 is also an A -invariant subspace. Thus the proof is complete. ♣

Corollary 1. The repeated application of the assertion of the theorem implies that if A is a nilpotent linear transformation with degree m then there exist elements $\underline{v}_0, \underline{v}_1, \dots, \underline{v}_r \in V$ and positive integers $k = k_0 \geq k_1 \geq k_2 \geq \dots \geq k_r$ such that vectors

$$\underline{v}_0, A(\underline{v}_0), \dots, A^{k_0-1}(\underline{v}_0), \underline{v}_1, A(\underline{v}_1), \dots, A^{k_1-1}(\underline{v}_1), \dots, \underline{v}_r, A(\underline{v}_r), \dots, A^{k_r-1}(\underline{v}_r)$$

form a basis of V , furthermore

$$V = V(\underline{v}_0, A) \oplus V(\underline{v}_1, A) \oplus \dots \oplus V(\underline{v}_r, A).$$

Corollary 2. Let now $A \in L(V)$ be an arbitrary linear transformation. Assume that the minimal polynomial of A is given as

$$r(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}.$$

From Theorem 6.5 we know that

$$V = N((A - \lambda_1 I)^{m_1}) \oplus N((A - \lambda_2 I)^{m_2}) \oplus \dots \oplus N((A - \lambda_r I)^{m_r}).$$

The restriction of mapping $(A - \lambda_i I)$ into the subspace $V_i = N((A - \lambda_i I)^{m_i})$ is nilpotent with degree m_i , since on $V_i, (A - \lambda_i I)^{m_i} = O$, and the exponent m_i cannot be lowered. Otherwise, the exponent m_i could also be lowered in the minimal polynomial r . Use the previous Corollary to see that for all $i = 1, 2, \dots, r$, there exist integers $m_i = m_{i0} \geq m_{i1} \geq \dots \geq m_{is_i}$ and elements $\underline{v}_{i0}, \underline{v}_{i1}, \dots, \underline{v}_{is_i}$ that

$$B_i = \{ \underline{v}_{i0}, (A - \lambda_i I)(\underline{v}_{i0}), \dots, (A - \lambda_i I)^{m_{i0}-1}(\underline{v}_{i0}), \dots, \underline{v}_{is_i}, (A - \lambda_i I)(\underline{v}_{is_i}), \dots, (A - \lambda_i I)^{m_{is_i}-1}(\underline{v}_{is_i}) \}$$

is a basis of V_i and

$$V_i = V(\underline{v}_{i0}, (A - \lambda_i I)) \oplus V(\underline{v}_{i1}, (A - \lambda_i I)) \oplus \dots \oplus V(\underline{v}_{is_i}, (A - \lambda_i I))$$

It is easy to show that all terms of the right-hand side are A -invariant subspaces. Consider next the restriction A_i of mapping A into subspace V_i . The matrix representation of A_i in basis B_i has the special block diagonal form

$$A_{i, B_i, B_i} = \begin{pmatrix} A_{i1} & & & \\ & A_{i2} & & \\ & & \ddots & \\ & & & A_{is_i} \end{pmatrix}$$

where for $j = 1, 2, \dots, s_i$, A_{ij} is an $m_{ij} \times m_{ij}$ matrix:

$$A_{ij} = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \ddots & & \\ & & \ddots & \lambda_i & \\ & & & 1 & \lambda_i \end{pmatrix}.$$

All diagonal elements of A_{ij} are equal to λ_i , and just under the diagonal, all elements are equal to one, and all other elements are equal to zero. This representation follows from the identities

$$\begin{aligned} & A((A - \lambda_i I)^j(\underline{v}_{ip})) \\ &= \begin{cases} (A - \lambda_i I)^{j+1}(\underline{v}_{ip}) + \lambda_i(A - \lambda_i I)^j(\underline{v}_{ip}) & \text{if } 0 \leq j < m_{ip} - 1 \\ \lambda_i(A - \lambda_i I)^j(\underline{v}_{ip}) & \text{if } j = m_{ip} - 1. \end{cases} \end{aligned}$$

And finally, $B = \{B_1, B_2, \dots, B_r\}$ gives a basis of V , and the matrix representation of mapping A in basis B is the following:

$$A_{B,B} = \begin{pmatrix} A_{1, B_1, B_1} & & & & \\ & A_{2, B_2, B_2} & & & \\ & & \ddots & & \\ & & & & A_{r, B_r, B_r} \end{pmatrix}.$$

This representation is called the *Jordan canonical-form* of transformation (or matrix) A . The number of diagonal blocks equals $s_1 + s_2 + \dots + s_r$, each block has the same eigenvalue in the diagonal, it has the form of A_{ij} , and each eigenvalue of A can be found in at least one diagonal block.

6.7 Complexification

In examining the structure of linear mappings, in several cases we had to distinguish between real and complex vector spaces, since different results apply to these different cases. In this section we will introduce a simple method which will enable us to apply directly results originally proven for the complex case to real vector and inner product spaces, and to real mappings. The idea is known as *imbedding*, and it constructs a complex vector space or inner-product space, a certain subset of which is isomorphic to the given real vector space and inner-product space. In addition, the restriction of the complex extension of a real mapping to this subset is equivalent to the original real mapping.

Consider now a real vector space V , and denote $\hat{V} = V \times V$. Define addition and multiplication by complex numbers on \hat{V} in the following ways:

$$(\underline{x}, \underline{y}) + (\underline{u}, \underline{v}) = (\underline{x} + \underline{u}, \underline{y} + \underline{v}),$$

and

$$(a + ib)(\underline{x}, \underline{y}) = (a\underline{x} - b\underline{y}, b\underline{x} + a\underline{y}).$$

It is easy to see that with these operations, \hat{V} is a complex vector space with zero element $(0, 0)$, and if multiplication by only real numbers is considered, then V is isomorphic to the set

$$\hat{V}_1 = \{(\underline{x}, 0) \mid \underline{x} \in V\}.$$

Therefore, $\dim(V) = \dim(\hat{V}_1)$, and if $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ is a basis of V then $\{(\underline{x}_1, 0), (\underline{x}_2, 0), \dots, (\underline{x}_n, 0)\}$ forms a basis in \hat{V}_1 . For easy reference we may use the notation $\underline{x} + i\underline{y}$ for $(\underline{x}, \underline{y})$.

If V is a real inner-product space, then a complex inner-product on \hat{V} can be defined as

$$(\underline{x} + i\underline{y}, \underline{u} + i\underline{v}) = ((\underline{x}, \underline{u}) - (\underline{y}, \underline{v})) + i((\underline{x}, \underline{v}) + (\underline{y}, \underline{u})).$$

The reader may easily verify that this inner-product satisfies the properties of complex inner-products. The length of the elements of \hat{V} are defined as

$$|\underline{x} + i\underline{y}| = \sqrt{(\underline{x} + i\underline{y}, \underline{x} + i\underline{y})} = \sqrt{|\underline{x}|^2 + |\underline{y}|^2},$$

where $|\underline{x}|$ and $|\underline{y}|$ are the lengths of \underline{x} and \underline{y} in V .

Consider next a linear transformation $A \in L(V)$. The extension \hat{A} of A to the complex vector space \hat{V} is defined by the equation

$$\hat{A}(\underline{x} + i\underline{y}) = A(\underline{x}) + i A(\underline{y}),$$

and it is easy to see that $\hat{A} \in L(\hat{V})$, and the restriction of \hat{A} into the real vector space is the original transformation A .

Based on this complexification method, real vector spaces, real inner-product spaces and their mappings can be treated as complex spaces and mappings, respectively, hence all general results proved for complex vector spaces can be directly applied for real cases.

6.8 Applications

The diagonal and Jordan canonical forms of square matrices are very useful in many fields of the application of matrix theory. In this section we will

first focus on dynamic economic systems described by difference or differential equations, and then some simple examples will be presented.

1. Consider first the *discrete time-invariant system* with state transition equation

$$\underline{x}(t+1) = \underline{A} \cdot \underline{x}(t) + \underline{b},$$

where $t = 0, 1, 2, \dots$, and \underline{A} is an $n \times n$ real matrix, \underline{b} is an n -element real vector. Let \underline{T} be an $n \times n$ nonsingular matrix and introduce the new state variable

$$\underline{z} = \underline{T}\underline{x},$$

then

$$\begin{aligned} \underline{z}(t+1) &= \underline{T}\underline{x}(t+1) = \underline{T}(\underline{A}\underline{x}(t) + \underline{b}) \\ &= \underline{T}(\underline{A} \cdot \underline{T}^{-1} \underline{z}(t) + \underline{b}) = (\underline{T}\underline{A}\underline{T}^{-1})\underline{z}(t) + \underline{T} \cdot \underline{b} \end{aligned} \quad (6.17)$$

showing that the new state-transition equation has the same form as originally and the new coefficient matrix becomes $\underline{T}\underline{A}\underline{T}^{-1}$. If matrix \underline{A} is diagonalizable, then there exists a nonsingular matrix such that $\underline{T}\underline{A}\underline{T}^{-1}$ is diagonal. In this case system (6.17) becomes the set of n independent difference equations, and in each of them the unknown function is real valued. Assume that

$$\underline{T}\underline{A}\underline{T}^{-1} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad \text{and} \quad \underline{T}\underline{b} = \begin{pmatrix} b'_1 \\ b'_2 \\ \dots \\ b'_n \end{pmatrix},$$

then equations (6.17) reduce to the following system

$$\begin{aligned} z_1(t+1) &= \lambda_1 z_1(t) + b'_1 \\ z_2(t+1) &= \lambda_2 z_2(t) + b'_2 \\ &\dots \\ z_n(t+1) &= \lambda_n z_n(t) + b'_n. \end{aligned}$$

The state transition matrix of the system can also be determined in a simple way if \underline{A} is diagonalizable. Since

$$\underline{A} = \underline{T}^{-1} \underline{D} \underline{T}$$

with

$$\underline{D} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

for all $t \geq 1$,

$$\begin{aligned} \underline{A}' &= (\underline{T}^{-1} \underline{D} \underline{T})(\underline{T}^{-1} \underline{D} \underline{T}) \dots (\underline{T}^{-1} \underline{D} \underline{T}) \\ &= \underline{T}^{-1} \underline{D} (\underline{T} \underline{T}^{-1}) \underline{D} \dots (\underline{T} \underline{T}^{-1}) \underline{D} \underline{T} = \underline{T}^{-1} \underline{D}' \underline{T} \end{aligned} \quad (6.18)$$

$$= \underline{T}^{-1} \begin{pmatrix} \lambda'_1 & & & \\ & \lambda'_2 & & \\ & & \ddots & \\ & & & \lambda'_n \end{pmatrix} \underline{T}.$$

Example 6.8. Consider matrix

$$\underline{A} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

In Example 6.6 we have shown that

$$\underline{A} = \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 6 \end{pmatrix} \begin{pmatrix} -\frac{1}{6} & -\frac{7}{42} & \frac{7}{21} \\ 0 & -\frac{6}{42} & \frac{3}{21} \\ \frac{1}{6} & \frac{19}{42} & \frac{8}{21} \end{pmatrix},$$

therefore equation (6.18) implies that for all $t \geq 1$,

$$\underline{A}^t = \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ (-1)^t \\ 6^t \end{pmatrix} \begin{pmatrix} -\frac{1}{6} & -\frac{7}{42} & \frac{7}{21} \\ 0 & -\frac{6}{42} & \frac{3}{21} \\ \frac{1}{6} & \frac{19}{42} & \frac{8}{21} \end{pmatrix}.$$

If \underline{A} cannot be diagonalized, then there is a nonsingular matrix such that $\underline{T} \underline{A} \underline{T}^{-1}$ has the Jordan canonical form. Then system (6.17) breaks up into small dimensional systems of the form

$$z_{ij}(t+1) = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \ddots & & \\ & & & \lambda_i & \\ & & & 1 & \lambda_i \end{pmatrix} z_{ij}(t) + \begin{pmatrix} b'_{ij1} \\ b'_{ij2} \\ \vdots \\ b'_{ijm_j} \end{pmatrix},$$

where we use the notation of the previous section. Notice that this system can be rewritten as follows:

$$\begin{aligned} z_{ij1}(t+1) &= \lambda_i z_{ij1}(t) + b'_{ij1} \\ z_{ij2}(t+1) &= \lambda_i z_{ij2}(t) + z_{ij1}(t) + b'_{ij2} \\ &\vdots \\ z_{ijm_j}(t+1) &= \lambda_i z_{ijm_j}(t) + z_{ij,m_j-1}(t) + b'_{ijm_j}. \end{aligned}$$

From the first equation we can determine function $z_{ij1}(t)$, and then, $z_{ij2}(t)$ can be obtained from the second equation. After $z_{ij2}(t)$ is determined, function $z_{ij3}(t)$ is obtained from the third equation, and so on. Notice, that at each step a real-valued function has to be computed. In summary, in both cases, the solution of an $n \times n$ system has been reduced to the solution of n single equations, which can be done efficiently.

The state-transition matrix of the system can also be obtained in a simple way. Notice first that if

$$\underline{A}_{ij} = \begin{pmatrix} \lambda_i & & & & & \\ & 1 & \lambda_i & & & \\ & & 1 & \ddots & & \\ & & & & \ddots & \\ & & & & & \lambda_i \\ & & & & & 1 & \lambda_i \end{pmatrix}$$

is a Jordan-block, then $\underline{A}_{ij} = \lambda_i \underline{I} + \underline{E}$, where \underline{I} is the $m_{ij} \times m_{ij}$ identity matrix, and \underline{E} is nilpotent with degree m_{ij} . By applying the binomial theorem (which can be used here, since \underline{I} and \underline{E} commute) we have

$$\begin{aligned} \underline{A}_{ij}^t &= (\lambda_i \underline{I} + \underline{E})^t = \sum_{l=0}^t \binom{t}{l} \lambda_i^l \underline{E}^{t-l} \\ &= \begin{pmatrix} \lambda_i^t & & & & 0 \\ \binom{t}{1} \lambda_i^{t-1} & \lambda_i^t & & & \\ \binom{t}{2} \lambda_i^{t-2} & \binom{t}{1} \lambda_i^{t-1} & \ddots & & \\ \vdots & \vdots & & & \lambda_i^t \\ \binom{t}{m_{ij}-1} \lambda_i^{t-m_{ij}+1} & \binom{t}{m_{ij}-2} \lambda_i^{t-m_{ij}+2} & \dots & \binom{t}{1} \lambda_i^{t-1} & \lambda_i^t \end{pmatrix}. \end{aligned}$$

And finally

$$\underline{A}' = \underline{T}^{-1} \left(\begin{array}{cccc} \underline{A}'_{11} & & & \\ & \ddots & & \\ & & \underline{A}'_{1s_1} & \\ & & & \ddots \\ & & & & \underline{A}'_{r,r} \\ & & & & & \ddots \\ & & & & & & \underline{A}'_{rs_r} \end{array} \right) \underline{T} . \quad (6.19)$$

2. Consider next the *continuous time-invariant system* with state transition equation

$$\dot{\underline{x}}(t) = \underline{A} \cdot \underline{x}(t) + \underline{b},$$

where $t \geq 0$ is the continuous time variable, and \underline{A} is an $n \times n$ real matrix, \underline{b} is an n -element real vector. Let \underline{T} again be a nonsingular matrix, and by introducing the new state variable $\underline{z} = \underline{T}\underline{x}$ we have

$$\begin{aligned} \dot{\underline{z}}(t) &= \underline{T}\dot{\underline{x}}(t) = \underline{T}(\underline{A}\underline{x}(t) + \underline{b}) \\ &= \underline{T}(\underline{A}\underline{T}^{-1}\underline{z}(t) + \underline{b}) = (\underline{T}\underline{A}\underline{T}^{-1})\underline{z}(t) + \underline{T} \cdot \underline{b}. \end{aligned} \quad (6.20)$$

That is, similarly to the discrete case, the state-transition equation for $\underline{z}(t)$ has the same form as for the original state variable $\underline{x}(t)$ with the new coefficient matrix $\underline{T}\underline{A}\underline{T}^{-1}$. If matrix \underline{A} is diagonal, then with an appropriate \underline{T} , $\underline{T}\underline{A}\underline{T}^{-1}$ is diagonal, and the state-transition equation (6.20) reduces to the following:

$$\begin{aligned}\dot{z}_1(t) &= \lambda_1 z_1(t) + b'_1 \\ \dot{z}_2(t) &= \lambda_2 z_2(t) + b'_2 \\ &\dots \\ \dot{z}_n(t) &= \lambda_n z_n(t) + b'_n,\end{aligned}$$

which can be solved much more efficiently than the original n -dimensional problem. This idea is illustrated in the next example.

Example 6.9. We will now solve the system

$$\dot{\underline{x}}(t) = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

From equation (6.20) and the results of Example 6.6 we see that by introducing the new variable

$$\underline{z}(t) = \begin{pmatrix} -\frac{1}{6} & -\frac{7}{42} & \frac{7}{21} \\ 0 & -\frac{6}{42} & \frac{3}{21} \\ \frac{1}{6} & \frac{19}{42} & \frac{8}{21} \end{pmatrix} \underline{x}(t),$$

a diagonal system is obtained:

$$\dot{\underline{z}}(t) = \begin{pmatrix} 0 & & \\ & -1 & \\ & & 6 \end{pmatrix} \underline{z}(t) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

which can be rewritten as

$$\begin{aligned}\dot{z}_1 &= 0 \\ \dot{z}_2 &= -z_2 \\ \dot{z}_3 &= 6z_3 + 1.\end{aligned}$$

It is easy to see that the solutions are

$$\begin{aligned}z_1(t) &= c_1 \\ z_2(t) &= c_2 e^{-t} \\ z_3(t) &= -\frac{1}{6} + c_3 e^{6t},\end{aligned}$$

where c_1, c_2, c_3 are arbitrary constants. The solution $\underline{x}(t)$ can then be obtained as

$$\underline{x}(t) = \underline{T}^{-1} \underline{z}(t) = \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 e^{-t} \\ -\frac{1}{6} + c_3 e^{6t} \end{pmatrix}.$$

The state transition matrix of the system can also be determined efficiently if \underline{A} is diagonalizable. Using equation (6.18) we have

$$e^{At} = \sum_{l=0}^{\infty} \frac{t^l}{l!} A^l = \sum_{l=0}^{\infty} \frac{t^l}{l!} \underline{T}^{-1} \begin{pmatrix} \lambda'_1 & & & \\ & \lambda'_2 & & \\ & & \ddots & \\ & & & \lambda'_n \end{pmatrix} \underline{T}$$

$$= \underline{T}^{-1} \left\{ \sum_{l=0}^{\infty} \frac{t^l}{l!} \begin{pmatrix} \lambda'_1 & & & \\ & \lambda'_2 & & \\ & & \ddots & \\ & & & \lambda'_n \end{pmatrix} \right\} \underline{T}$$

$$= \underline{T}^{-1} \begin{pmatrix} \sum_{l=0}^{\infty} \frac{t^l \lambda'_1}{l!} & & & \\ & \sum_{l=0}^{\infty} \frac{t^l \lambda'_2}{l!} & & \\ & & \ddots & \\ & & & \sum_{l=0}^{\infty} \frac{t^l \lambda'_n}{l!} \end{pmatrix} \underline{T}$$

$$= \underline{T}^{-1} \begin{pmatrix} e^{\lambda'_1 t} & & & \\ & e^{\lambda'_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda'_n t} \end{pmatrix} \underline{T}$$

Example 6.10. Consider again matrix

$$\underline{A} = \begin{pmatrix} 1 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

then similarly to Example 6.8 we have

$$e^{\underline{A}t} = \begin{pmatrix} -5 & 9 & 1 \\ 1 & -5 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ e^{-t} \\ e^{6t} \end{pmatrix} \begin{pmatrix} -\frac{1}{6} & -\frac{7}{42} & \frac{7}{21} \\ 0 & -\frac{6}{42} & \frac{3}{21} \\ \frac{1}{6} & \frac{19}{42} & \frac{8}{21} \end{pmatrix}.$$

If \underline{A} cannot be diagonalized, then there is a nonsingular matrix \underline{T} such that \underline{TAT}^{-1} has the Jordan canonical-form. Then system (6.20) is a set of small dimensional systems

$$\dot{z}_{ij}(t) = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \ddots & & \\ & & & \lambda_i & \\ & & & 1 & \lambda_i \end{pmatrix} z_{ij}(t) + \begin{pmatrix} b'_{ij1} \\ b'_{ij2} \\ \vdots \\ b'_{ijm_{ij}} \end{pmatrix}.$$

This system can be rewritten as follows:

$$\begin{aligned}\dot{z}_{ij_1}(t) &= \lambda_i z_{ij_1}(t) + b'_{ij_1} \\ \dot{z}_{ij_2}(t) &= \lambda_i z_{ij_2}(t) + z_{ij_1}(t) + b'_{ij_2} \\ &\dots \\ \dot{z}_{ij_{m_j}}(t) &= \lambda_i z_{ij_{m_j}}(t) + z_{ij_{m_j-1}}(t) + b'_{ij_{m_j}}.\end{aligned}$$

From the first equation we can determine function $z_{ij_1}(t)$, and then, $z_{ij_2}(t)$ can be obtained from the second equation. After $z_{ij_2}(t)$ is determined, function $z_{ij_3}(t)$ is obtained from the third equation, and so on. At each step a single dimensional linear differential equation has to be solved. The solution of an $n \times n$ system has therefore been reduced to the solution of n single dimensional equations, which can be done more efficiently.

Example 6.11. We will now solve the system

$$\dot{\underline{z}} = \begin{pmatrix} 2 & & \\ 1 & 2 & \\ & 1 & 2 \end{pmatrix} \underline{z} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

where the coefficient matrix has the Jordan block-form. This system can be rewritten as

$$\begin{aligned}\dot{z}_1 &= 2z_1 + 1 \\ \dot{z}_2 &= 2z_2 + z_1 + 1 \\ \dot{z}_3 &= 2z_3 + z_2 + 1.\end{aligned}$$

From the first equations we have

$$z_1(t) = -\frac{1}{2} + c_1 e^{2t}$$

where c_1 is a constant. Substituting this solution into the second equation we obtain a single equation for z_2 :

$$\dot{z}_2 = 2z_2 + c_1 e^{2t} + \frac{1}{2}.$$

The solution of this equation is the following:

$$z_2(t) = (c_1 t + c_2) e^{2t} - \frac{1}{4}.$$

Substituting this solution into the third equation we have

$$\dot{z}_3 = 2z_3 + (c_1 t + c_2) e^{2t} + \frac{3}{4},$$

the solution of which is the following:

$$z_3(t) = \left(\frac{c_1 t^2}{2} + c_2 t + c_3 \right) e^{2t} + \frac{3}{4}.$$

The state transition matrix of the original system can also be obtained efficiently. Assume again that the Jordan canonical form of matrix \underline{A} has the blocks

$$\underline{A}_{ij} = \begin{pmatrix} \lambda_i & & & & \\ 1 & \lambda_i & & & \\ & 1 & \ddots & & \\ & & & \lambda_i & \\ & & & 1 & \lambda_i \end{pmatrix},$$

then similarly to the case of diagonal matrices we can easily show that

$$e^{A_t} = \underline{T}^{-1} \begin{pmatrix} e^{A_{11}t} & & & & \\ & \ddots & & & \\ & & e^{A_{11}t} & & \\ & & & \ddots & \\ & & & & e^{A_{r1}t} & \\ & & & & & \ddots & \\ & & & & & & e^{A_{rr}t} \end{pmatrix} \underline{T}.$$

For each Jordan block, $\underline{A}_{ij} = \lambda_i \underline{I} + \underline{E}$, and therefore

$$\begin{aligned} e^{A_{ij}t} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} (\lambda_i \underline{I} + \underline{E})^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{l=0}^k \binom{k}{l} \lambda_i^l \underline{E}^{k-l} \\ &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{l=0}^k \binom{k}{l} \underline{E}^l \lambda_i^{k-l}. \end{aligned}$$

By using the fact that $\underline{E}^l = \underline{0}$ for $l \geq m_{ij}$ and $\binom{k}{l} = 0$ as $k < l$ we see that

$$\begin{aligned} e^{A_{ij}t} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{l=0}^{m_{ij}-1} \binom{k}{l} \underline{E}^l \lambda_i^{k-l} = \sum_{l=0}^{m_{ij}-1} \underline{E}^l \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} \binom{k}{l} \lambda_i^{k-l} \right) \\ &= \sum_{l=0}^{m_{ij}-1} \underline{E}^l \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^l}{d\lambda^l} \left[\frac{t^k}{k!} \lambda^k \right]_{\lambda=\lambda_i} \\ &= \sum_{l=0}^{m_{ij}-1} \frac{1}{l!} \underline{E}^l \frac{d^l}{d\lambda^l} [e^{\lambda t}]_{\lambda=\lambda_i} = \sum_{l=0}^{m_{ij}-1} \underline{E}^l \frac{t^l}{l!} e^{\lambda_i t} \end{aligned}$$

It is easy to see that the eigenvalues of $\underline{I} - \underline{A}$ are the values $1 - \lambda_i$, which are nonzero, therefore $\underline{I} - \underline{A}$ is nonsingular. Notice that

$$\underline{I} + \underline{A} + \underline{A}^2 + \dots + \underline{A}^N = (\underline{I} - \underline{A})^{-1} (\underline{I} - \underline{A}^{N+1}),$$

and since $\underline{A}^{N+1} \rightarrow \underline{O}$ as $N \rightarrow \infty$,

$$\underline{I} + \underline{A} + \underline{A}^2 + \dots + \underline{A}^N + \dots = (\underline{I} - \underline{A})^{-1}.$$

This identity is often used as an iteration method for inverting matrices. Assume that \underline{B}^{-1} has to be determined. Assume that $\underline{A} = \underline{I} - \underline{B}$ satisfies the condition that all of its eigenvalues are inside the unit circle (that is, $|\lambda_i| < 1$ for all eigenvalues of \underline{A}). Then

$$\underline{I} + \underline{A} + \underline{A}^2 + \dots + \underline{A}^N + \dots = (\underline{I} - (\underline{I} - \underline{B}))^{-1} = \underline{B}^{-1}.$$

The infinite series is the limit of the sequence:

$$\begin{aligned} \underline{A}_0 &= \underline{I} \\ \underline{A}_{k+1} &= \underline{I} + \underline{A}_k \cdot \underline{A} \quad (k \geq 0), \end{aligned}$$

since

$$\begin{aligned} \underline{A}_1 &= \underline{I} + \underline{A} \\ \underline{A}_2 &= \underline{I} + (\underline{I} + \underline{A})\underline{A} = \underline{I} + \underline{A} + \underline{A}^2, \end{aligned}$$

and in general

$$\underline{A}_{k+1} = \underline{I} + (\underline{I} + \underline{A} + \dots + \underline{A}^k)\underline{A} = \underline{I} + \underline{A} + \underline{A}^2 + \dots + \underline{A}^{k+1}.$$

4. In equation (6.19) we have seen a closed-form representation of the matrix \underline{A}^t for all $t \geq 1$. This formula was developed for solving discrete dynamic systems. Matrix powers however should be computed in many other cases. One example of such an application is concerned with the vertex matrix of directed graphs. In application 8 of section 1.5 we have seen that if \underline{A} is the vertex-matrix of a directed graph, then the elements of the matrix \underline{A}^r ($r \geq 1$) give the r -step connections from each vertex to all other vertices and to itself. The fast computation of \underline{A}^r is therefore essential in determining the number of higher-step connections.

5. In section 3.8 we have briefly analysed Markov chains, and have seen that future state probabilities can be directly obtained by equation (3.26), which has the form

$$\underline{x}(t+n) = \underline{P}^n \underline{x}(t).$$

Here the entries of $\underline{x}(t)$ and $\underline{x}(t+n)$ give the state probabilities at time periods t and $t+n$, respectively, furthermore \underline{P} is the transition matrix. For efficient computations we have to determine the matrix powers \underline{P}^n for large values of n . This task can be performed for example, by using the method of the first application of this section (equations (6.18) and (6.19)).

6.9 Exercises

1. For an $n \times n$ square matrix $\underline{A} = (a_{ij})$, define the trace of \underline{A} as $tr(\underline{A}) = a_{11} + a_{22} + \dots + a_{nn}$.

Prove that $tr(\underline{A}) = tr(\underline{T}^{-1} \underline{A} \underline{T})$ for all nonsingular $n \times n$ matrices.

2. Show that

a) $tr(\underline{A}^T) = tr(\underline{A})$;

b) $tr(\underline{A} + \underline{B}) = tr(\underline{A}) + tr(\underline{B})$, if \underline{B} is also $n \times n$;

c) $tr(a\underline{A}) = atr(\underline{A})$, if a is a scalar;

d) $tr(\underline{AB}) = tr(\underline{BA})$, if \underline{B} is also $n \times n$;

e) $tr(\underline{A}) = \lambda_1 + \lambda_2 + \dots + \lambda_n$, if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of \underline{A} .

3. Prove that $tr(\underline{AA}^T) \geq 0$ and $tr(\underline{AA}^T) = 0$ if and only if $\underline{A} = \underline{O}$.

4. Find the eigenvalues of the complex matrix

$$\underline{A} = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix}.$$

5. Find the characteristic equations of the following matrices:

a) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}$;

$$\text{b) } \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix};$$

$$\text{c) } \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}.$$

6. Find the eigenvalues and the associated eigenvectors for the following matrices

$$\text{a) } \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix};$$

$$\text{b) } \begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix}, \text{ where } w > 0 \text{ is a given number;}$$

$$\text{c) } \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix};$$

$$\text{d) } \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

7. Repeat Example 6.1 with matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

8. Illustrate Theorem 6.1 for matrices of Problem 6.
9. Repeat Example 6.3 with matrix

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

10. Find $p(\underline{A})$, where

$$p(t) = t^3 + t^2 + 2t + 1$$

and

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

11. Find the minimal polynomial of the matrix of the previous problem. (Repeat Example 6.5.)
12. Diagonalize the matrices of Problem 6.
13. Characterize all $2 \cdot 2$ real nilpotent matrices of degree 2.
14. Assume that the only eigenvalue of an $n \times n$ real matrix is zero. Prove that the matrix is nilpotent.

15. (Continuation of Problem 14). An $n \times n$ real matrix \underline{A} is called *unipotent*, if $\underline{A} - \underline{I}$ is nilpotent. Prove that \underline{A} is unipotent if and only if 1 is the only eigenvalue of \underline{A} .

16. Illustrate the Cayley-Hamilton Theorem with the matrices of Problem 5.

17. Illustrate the Cayley-Hamilton Theorem with the matrices of Problem 6.

18. Find \underline{A}^t for matrices

a) $\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix};$

b) $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}.$

19. Find $e^{\underline{A}t}$ for matrices

a) $\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix};$

b) $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}.$

20. Find $\underline{x}(t)$ for the diagonal discrete system

$$\underline{x}(t+1) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

21. Find $\underline{x}(t)$ for the diagonal continuous system

$$\dot{\underline{x}}(t) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

22. Find $\underline{x}(t)$ for the triangular discrete system

$$\underline{x}(t+1) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

23. Find $\underline{x}(t)$ for the triangular continuous system

$$\dot{\underline{x}}(t) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \underline{x}(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{x}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

24. Find the inverse of matrix

$$\underline{A} = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

by using the method given in Application 3 of section 6.8.

25. Let $\underline{A} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. Show that

$$e^{\underline{A}} = \begin{pmatrix} e^a \cos b & e^a \sin b \\ -e^a \sin b & e^a \cos b \end{pmatrix}.$$

Chapter 7

Special Matrices

7.1 Introduction

In this chapter the most important classes of special matrices will be examined. In the first section the main properties of diagonal, tridiagonal, and triangular matrices will be discussed. In the second section we will focus on idempotent and nilpotent matrices, and the third section will be devoted to special matrices defined in inner-product spaces.

7.2 Diagonal, Tridiagonal, and Triangular Matrices

From Chapter 1 we know that an $n \times n$ square matrix $\underline{A} = (a_{ij})$ is *diagonal* if $a_{ij} = 0$ for all $i \neq j$. That is, a diagonal matrix has the special form

$$\underline{A} = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

where all element under and above the diagonal are equal to zero. The rank of \underline{A} equals the number of nonzero diagonal elements, since the columns with nonzero diagonal elements are the nonzero-multiples of the

corresponding natural basis elements \underline{e}_k . The characteristic polynomial of \underline{A} is the determinant of the diagonal matrix

$$\underline{A} - \lambda \underline{I} = \begin{pmatrix} a_{11} - \lambda & & & 0 \\ & a_{22} - \lambda & & \\ & & \ddots & \\ 0 & & & a_{nn} - \lambda \end{pmatrix}$$

which is the product of the diagonal elements:

$$\varphi(\lambda) = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda),$$

therefore the eigenvalues are the diagonal elements of \underline{A} . Consider now an eigenvalue λ , and assume that

$$\lambda = a_{i_1 i_1} = a_{i_2 i_2} = \dots = a_{i_r i_r}$$

and

$$\lambda \neq a_{jj} \quad \text{for} \quad j \notin \{i_1, i_2, \dots, i_r\}.$$

That is, λ can be found as the diagonal element in positions i_1, i_2, \dots, i_r . Let $\underline{x} = (x_i)$ be an associated eigenvector. Then the eigenvector equation $\underline{A}\underline{x} = \lambda\underline{x}$ implies that

$$a_{ii}x_i = \lambda x_i \quad (i = 1, 2, \dots, n).$$

If $i \in \{i_1, i_2, \dots, i_r\}$, then this equation is satisfied for all x_i , since $\lambda = a_{ii}$, and if $i \notin \{i_1, i_2, \dots, i_r\}$, then $x_i = 0$. Hence the eigenvectors associated to λ have the special form:

$$x_i = \begin{cases} \text{arbitrary} & \text{if } i \in \{i_1, i_2, \dots, i_r\} \\ 0 & \text{if } i \notin \{i_1, i_2, \dots, i_r\} \end{cases} \quad (7.1)$$

Notice that system $\{\underline{e}_{i_1}, \underline{e}_{i_2}, \dots, \underline{e}_{i_r}\}$ forms a basis of the subspace of the eigenvectors associated to λ .

If $\underline{A} = \underline{O}$, then all diagonal elements are equal to zero, therefore $\lambda = 0$ is the only eigenvalue with multiplicity n , and any arbitrary n -element vector is an eigenvector of \underline{A} . Similarly, if $\underline{A} = \underline{I}$, then all diagonal elements are equal to one, therefore $\lambda = 1$ is the only eigenvalue with multiplicity n , and all n -element vectors are eigenvectors of \underline{A} .

Definition 7.1. A matrix $\underline{A} = (a_{ij})$ is called *tridiagonal*, if $a_{ij} = 0$ for $|i - j| \geq 2$.

This definition implies that \underline{A} is tridiagonal if and only if it has the special form

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & 0 \\ 0 & 0 & 0 & \dots & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n,n-1} & a_{nn} \end{pmatrix}$$

Two specialties of tridiagonal matrices will be discussed here: the solution of linear equations with tridiagonal matrices, and an easy way to compute the characteristic polynomials of such matrices.

Table 7.1 Initial elimination table for a tridiagonal system.

x_1	x_2	x_3	\dots	x_{n-2}	x_{n-1}	x_n	b
a_{11}	a_{12}	0	\dots	0	0	0	b_1
a_{21}	a_{22}	a_{23}	\dots	0	0	0	b_2
0	a_{32}	a_{33}	\dots	0	0	0	b_3
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots
0	0	0	\dots	$a_{n-2,n-2}$	$a_{n-2,n-1}$	0	b_{n-2}
0	0	0	\dots	$a_{n-1,n-2}$	$a_{n-1,n-1}$	$a_{n-1,n}$	b_{n-1}
0	0	0	\dots	0	$a_{n,n-1}$	a_{nn}	b_n

In applying the elimination method for solving linear equations with tridiagonal matrices, the procedure can be largely simplified, if at each step, the next diagonal element is selected as the pivot. To illustrate this point, assume that a system of linear equations has a tridiagonal coefficient matrix. The initial elimination table is shown in Table 7.1. Select the pivot element a_{11} by assuming that it is nonzero. Subtract the a_{21}/a_{11} multiple of the first equation from the second equation, then the second row of the elimination table becomes the following:

$0a_{22}^{(1)} a_{23} \dots 000$	$b_2^{(2)}$
----------------------------------	-------------

with $a_{22}^{(1)} = a_{22} - a_{12} \cdot \frac{a_{21}}{a_{11}}$, and $b_2^{(1)} = b_2 - b_1 \cdot \frac{a_{21}}{a_{11}}$. Notice that by omitting

the first row, the remaining system with only $n-1$ equations for the unknowns x_2, x_3, \dots, x_n remains tridiagonal. Select next the pivot element $a_{22}^{(1)}$, then only the element a_{32} has to be eliminated by subtracting the $a_{32}/a_{22}^{(1)}$ multiple of the new second equation from the third equation. In the new table only the third row changes, and by omitting the second equation, the remaining $n-2$ equations for the $n-2$ remaining unknowns x_3, \dots, x_n is tridiagonal again. At each elimination step only one diagonal element and the right hand side number has to be modified, which makes the entire elimination procedure very inexpensive. After elimination is done, the resulting table is also special as shown in Table 7.2, where all elements under the diagonal equal zero. Notice that back substitution also becomes simple in this case, since

Table 7.2. Final elimination table for tridiagonal system.

x_1	x_2	x_2	\dots	x_{n-2}	x_{n-1}	x_n	b
a_{11}	a_{12}	0	\dots	0	0	0	b_1
	$a_{22}^{(1)}$	a_{23}	\dots	0	0	0	$b_2^{(1)}$
		$a_{33}^{(2)}$	\dots	0	0	0	$b_3^{(2)}$
			\ddots	$a_{n-2, n-2}^{(n-3)}$	$a_{n-2, n-1}$	0	$b_{n-2}^{(n-3)}$
					$a_{n-1, n-1}^{(n-2)}$	$a_{n-1, n}$	$b_{n-1}^{(n-2)}$
						$a_{nn}^{(n-1)}$	$b_n^{(n-1)}$

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}, \quad (7.2)$$

and for $i = n-1, n-2, \dots, 2, 1$,

$$x_i = \frac{b_i^{(i-1)} - a_{i,i+1}x_{i+1}}{a_{ii}^{(i-1)}}. \quad (7.3)$$

The characteristic polynomial φ of a tridiagonal matrix A can be determined by using a special recursion as it is shown next. The definition of characteristic polynomials implies that

$$\varphi_n(\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & 0 & \dots & 0 & 0 & 0 \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} - \lambda & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & & a_{n-2,n-2} - \lambda & a_{n-2,n-1} & 0 \\ 0 & 0 & 0 & & a_{n-1,n-2} & a_{n-1,n-1} - \lambda & a_{n-1,n} \\ 0 & 0 & 0 & & 0 & a_{n,n-1} & a_{nn} - \lambda \end{pmatrix}$$

Expand this determinant with respect to its last column. The cofactor φ_{n-1} of $a_{nn} - \lambda$ has the same structure as the entire determinant but instead of n , its size is $n-1$. That is, it is the characteristic polynomial of the $(n-1) \times (n-1)$ tridiagonal matrix with elements a_{ii} ($i=1,2,\dots,n-1$) and $a_{i,i+1}$ and $a_{i+1,i}$ ($i=1,2,\dots,n-2$). The cofactor of the element $a_{n-1,n}$ is the following:

$$-\det \begin{pmatrix} a_{11} - \lambda & a_{12} & 0 & \dots & 0 & 0 & 0 \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} - \lambda & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2,n-3} & a_{n-2,n-2} - \lambda & a_{n-2,n-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & a_{n,n-1} \end{pmatrix}$$

Expanding this determinant with respect to the last row we get the product $a_{n,n-1}\varphi_{n-2}(\lambda)$, where $\varphi_{n-2}(\lambda)$ has the same structure as $\varphi_n(\lambda)$, its size is $(n-2) \times (n-2)$ with elements a_{ii} ($i=1,2,\dots,n-2$) and $a_{i,j+1}$ and $a_{i+1,j}$ ($i=1,2,\dots,n-3$). Hence we derived a recursion to find φ_n :

$$\begin{aligned} \varphi_0(\lambda) &= 1 \\ \varphi_1(\lambda) &= a_{11} - \lambda \end{aligned}$$

and for $k=2,3,\dots,n$,

$$\varphi_k(\lambda) = (a_{kk} - \lambda)\varphi_{k-1}(\lambda) - a_{k-1,k} \cdot a_{k,k-1}\varphi_{k-2}(\lambda). \quad (7.4)$$

Example 7.1. We will first solve the system of linear equations

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 + 2x_2 + x_3 &= 4 \\ x_2 + 2x_3 + x_4 &= 4 \\ x_3 + 2x_4 &= 3, \end{aligned}$$

where the coefficient matrix is tridiagonal. The elimination process is shown in Table 7.3.

Table 7.3. Elimination of Example 7.1.

	x_1	x_2	x_3	x_4	
	1	1	0	0	2
	1	2	1	0	4
	0	1	2	1	4
	0	0	1	2	3
row 2-pivot row	1	1	0	0	2
	0	1	1	0	2
	0	1	2	1	4
	0	0	1	2	3
row 3-pivot row	1	1	0	0	2
	0	1	1	0	2
	0	0	1	1	2
	0	0	1	2	3
row 4-pivot row	1	1	0	0	2
	0	1	1	0	2
	0	0	1	1	2
	0	0	0	1	1

Back substitution shows that the solution is the following:

$$\begin{aligned}x_4 &= 1, \\x_3 &= 2 - x_4 = 1, \\x_2 &= 2 - x_3 = 1, \\x_1 &= 2 - x_2 = 1.\end{aligned}$$

The characteristic polynomial of the coefficient matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

will be next determined by using recursion (7.4):

$$\varphi_0(\lambda) = 1;$$

$$\varphi_1(\lambda) = 1 - \lambda;$$

$$\varphi_2(\lambda) = (2 - \lambda)\varphi_1(\lambda) - 1 \cdot 1 \cdot \varphi_0(\lambda) = (2 - \lambda)(1 - \lambda) - 1 = \lambda^2 - 3\lambda + 1;$$

$$\begin{aligned}\varphi_3(\lambda) &= (2 - \lambda)\varphi_2(\lambda) - 1 \cdot 1 \cdot \varphi_1(\lambda) = (2 - \lambda)(\lambda^2 - 3\lambda + 1) - (1 - \lambda) \\ &= -\lambda^3 + 5\lambda^2 - 6\lambda + 1;\end{aligned}$$

and finally,

$$\begin{aligned}\varphi_4(\lambda) &= (2 - \lambda)\varphi_3(\lambda) - 1 \cdot 1 \cdot \varphi_2(\lambda) \\ &= (2 - \lambda)(-\lambda^3 + 5\lambda^2 - 6\lambda + 1) - (\lambda^2 - 3\lambda + 1) \\ &= \lambda^4 - 7\lambda^3 + 15\lambda^2 - 10\lambda + 1.\end{aligned}$$

Consider next an $n \times n$ upper triangular matrix

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

where all elements under the diagonal are equal to zero. If a system of linear equations has an upper triangular coefficient matrix, then there is no need for elimination, the solution can be easily obtained by back substitution. This matrix \underline{A} is nonsingular if and only if all diagonal elements differ from zero, since

$$\det(\underline{A}) = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn},$$

and this product is nonzero if and only if all factors are nonzero. The characteristic polynomial of \underline{A} equals

$$\varphi(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda),$$

that is, the eigenvalues of \underline{A} are the diagonal elements.

Lower triangular matrices have similar properties. A lower triangular matrix \underline{A} has the form

$$\underline{A} = \begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \dots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}$$

where all elements above the diagonal equal zero. If the coefficient matrix of a system of equations is lower triangular, then there is no need for elimination. The value of x_1 can be obtained from the first equation, and then, x_2 can be determined from the second equation, and so on. Finally, x_n is obtained from the last equation. This procedure is called the *forward substitution*. This matrix \underline{A} is nonsingular if and only if all diagonal elements differ from zero. The characteristic polynomial of \underline{A} is the following:

$$\varphi(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda),$$

and the eigenvalues are again the diagonal elements of \underline{A} .

7.3 Idempotent and Nilpotent Matrices

In section 6.4 we called a square matrix \underline{A} *idempotent*, if $\underline{A}^2 = \underline{A}$. We will first prove that the only nonsingular idempotent matrix is the identity matrix. Notice first that the definition implies that

$$\underline{A}(\underline{I} - \underline{A}) = \underline{O}. \quad (7.5)$$

If \underline{A} is nonsingular, then \underline{A}^{-1} exists. Multiply both sides of this equation by \underline{A}^{-1} from left. Then we have

$$\underline{I} - \underline{A} = \underline{O},$$

that is,

$$\underline{A} = \underline{I}.$$

Consider next the eigenvalue equation of \underline{A} :

$$\underline{A}\underline{x} = \lambda\underline{x}.$$

Since $\underline{A}^2 = \underline{A}$, any eigenvalue λ and associated eigenvector satisfy relation

$$\lambda\underline{x} = \underline{A}\underline{x} = \underline{A}^2\underline{x} = \underline{A}(\underline{A}\underline{x}) = \underline{A}(\lambda\underline{x}) = \lambda\underline{A}\underline{x} = \lambda^2\underline{x},$$

and the fact that the eigenvector \underline{x} differs from zero implies that $\lambda^2 = \lambda$. Therefore the eigenvalues of \underline{A} are 0 and 1. Equation (7.5) implies that for any column \underline{u} of matrix $\underline{I} - \underline{A}$, $\underline{A}\underline{u} = \underline{0}\underline{u}$ showing that the columns of $\underline{I} - \underline{A}$ are eigenvectors of \underline{A} associated to the zero eigenvalue. Equation (7.5) can be rewritten as $(\underline{A} - \underline{I})\underline{A} = \underline{O}$ which shows that any column \underline{v} of \underline{A} satisfies equation $\underline{A}\underline{v} = 1 \cdot \underline{v}$. That is, the columns of \underline{A} are eigenvectors associated to the eigenvalue $\lambda = 1$. Next we show that there are n linearly independent eigenvectors which implies that \underline{A} is diagonalizable. Notice first that $\underline{A} + (\underline{I} - \underline{A}) = \underline{I}$, and therefore

$$n = \text{rank}(\underline{I}) \leq \text{rank}(\underline{A}) + \text{rank}(\underline{I} - \underline{A}).$$

Let $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_k$ be a basis of the column space of \underline{A} , and let $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_l$ be a basis of the column space of $\underline{I} - \underline{A}$. Theorem 6.1 implies that system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_k, \underline{v}_1, \underline{v}_2, \dots, \underline{v}_l\}$ is linearly independent. Since they are in an n -dimensional space and $k+l \geq n$, $k+l$ must be equal to n . Thus, this set of eigenvectors forms a basis of the n -dimensional vector space.

A square matrix \underline{A} is called *nilpotent* if $\underline{A}^k = \underline{O}$ for some positive integer k . The smallest such integer k is called the *degree* of \underline{A} . In Section 6.5 we have seen that nilpotent matrices are always singular. We can easily show that the only eigenvector of \underline{A} is zero. Consider the eigenvector equation of \underline{A} ,

$$\underline{A}\underline{x} = \lambda\underline{x}.$$

Let l be the smallest positive integer such that $\underline{A}^l \underline{x} = \underline{0}$. Such an l exists, since if l is selected as the degree of \underline{A} , $\underline{A}^l = \underline{O}$ as well as $\underline{A}^l \underline{x} = \underline{0}$. The choice of $l=0$ is not appropriate, since $\underline{A}^0 \underline{x} = \underline{I}\underline{x} = \underline{x} \neq \underline{0}$. Multiply both sides of the eigenvector equation by \underline{A}^{l-1} from left to see that

$$\underline{A}^l \underline{x} = \lambda \underline{A}^{l-1} \underline{x}.$$

Since $\underline{A}^{l-1} \underline{x} \neq \underline{0}$ and $\underline{A}^l \underline{x} = \underline{0}$, necessarily $\lambda = 0$. This observation implies that all diagonal elements of the Jordan form of nilpotent matrices equal zero.

7.4 Matrices in Inner Product Spaces

In this section it is assumed that V is a finitely generated real vector space, and there is a real inner-product defined on V . Let $A \in L(V)$ be a linear transformation on V . From Theorem 2.12 we know that there is an orthonormal basis $B = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ in V , where n is the dimension of V . Assume that

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

is the matrix-representation of A in basis B . Notice first that equality (5.14) implies that for all $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, n$,

$$\begin{aligned} (A(\underline{x}_i), \underline{x}_j) &= (\underline{c}_i, \underline{x}_j) = (a_{1i}\underline{x}_1 + a_{2i}\underline{x}_2 + \dots + a_{ni}\underline{x}_n, \underline{x}_j) \\ &= \sum_{l=1}^n a_{li}(\underline{x}_l, \underline{x}_j) = a_{ji}, \end{aligned}$$

where \underline{c}_i denotes the i^{th} column of matrix \underline{A} . Similarly,

$$\begin{aligned} (\underline{x}_i, A(\underline{x}_j)) &= (\underline{x}_i, \underline{c}_j) = (\underline{x}_i, a_{1j}\underline{x}_1 + a_{2j}\underline{x}_2 + \dots + a_{nj}\underline{x}_n) \\ &= \sum_{l=1}^n a_{lj}(\underline{x}_i, \underline{x}_l) = a_{ij}. \end{aligned}$$

Definition 7.2. The *adjoint* of an $A \in L(V)$ is defined as the linear transformation $A' \in L(V)$ such that for all $\underline{x}, \underline{y} \in V$,

$$(A(\underline{x}), \underline{y}) = (\underline{x}, A'(\underline{y})). \quad (7.6)$$

Our first result guarantees the existence of the unique adjoint for all A , and presents the matrix representation of A' .

Theorem 7.1. For all $A \in L(V)$ there is a unique adjoint, and its matrix representation in basis B is given as the transpose \underline{A}'^T of matrix \underline{A} .

Proof. We have shown above that $(A(\underline{x}_i), \underline{x}_j) = a_{ji}$ and $(\underline{x}_i, A'(\underline{x}_j)) = a'_{ij}$, therefore the elements a'_{ij} of the matrix representation \underline{A}' of the adjoint must satisfy the equation $a'_{ij} = a_{ji}$ for all i and j , hence \underline{A}' , if exists, must be the transpose \underline{A}'^T of \underline{A} . Next we show that this transformation satisfies relation (7.6) for all $\underline{x}, \underline{y} \in V$, therefore it is the adjoint of A . Assume that

$$\underline{x} = \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2 + \dots + \alpha_n \underline{x}_n$$

and

$$\underline{y} = \beta_1 \underline{x}_1 + \beta_2 \underline{x}_2 + \dots + \beta_n \underline{x}_n,$$

then

$$\begin{aligned} (A(\underline{x}), \underline{y}) &= \left(\sum_{i=1}^n \alpha_i A(\underline{x}_i), \sum_{j=1}^n \beta_j \underline{x}_j \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (A(\underline{x}_i), \underline{x}_j) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j a_{ji}, \end{aligned}$$

and similarly,

$$\begin{aligned} (\underline{x}, A'(\underline{y})) &= \left(\sum_{i=1}^n \alpha_i \underline{x}_i, \sum_{j=1}^n \beta_j A'(\underline{x}_j) \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (\underline{x}_i, A'(\underline{x}_j)) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j a'_{ij} \\ &= \sum \sum \alpha_i \beta_j a_{ji} = (A(\underline{x}), \underline{y}). \end{aligned}$$

✱

Remark. If V is a complex vector space and the inner product is complex, then a similar proof shows that $\underline{A}' = \overline{\underline{A}}^T$, where overbar denotes complex conjugate.

From Definition 7.2 it is easy to see that the adjoint of A' is A , furthermore

- (i) $(aA)' = aA'$ for all scalars a and mappings A ;
- (ii) $(A + C)' = A' + C'$ for all $A, C \in L(V)$;
- (iii) $(AC)' = C'A'$ for all $A, C \in L(V)$.

We mention also that adjoint mappings can be defined in the more general case when $A \in L(V, W)$ where V and W are two inner product spaces and either both are real or both are complex. For our purposes it is

sufficient to discuss only the case of $V=W$. The general case is examined in a similar way as the one presented in this book.

In this section three special classes of mappings (or matrices) will be examined. They are first defined.

Definition 7.3. Let $A \in L(V)$ be a linear transformation on V .

- (i) A is called *symmetric* (or *self-adjoint* in the complex case) if $A' = A$;
- (ii) A is called *orthogonal* (or *unitary* in the complex case), if A^{-1} exists and $A^{-1} = A'$;
- (iii) A is called *normal* if $AA' = A'A$.

First we will examine the main properties of symmetric (or self-adjoint) matrices. If the matrix is $n \times n$ and is considered as complex, then from Chapter 6, we know that it always has n real or complex eigenvalues. However, if the matrix is self-adjoint, then the eigenvalue are real.

Theorem 7.2. Let \underline{A} be a self-adjoint matrix. Then

- (a) All eigenvalues of \underline{A} are real;
- (b) Eigenvectors associated to different eigenvalues are orthogonal;
- (c) There is an orthonormal basis in V that consists of eigenvectors of \underline{A} .

Proof. (a) Let λ be an eigenvalue of \underline{A} with associated eigenvector \underline{u} . Then the eigenvector equation $\underline{A}\underline{u} = \lambda\underline{u}$ implies that

$$\lambda(\underline{u}, \underline{u}) = (\lambda\underline{u}, \underline{u}) = (\underline{A}\underline{u}, \underline{u}) = (\underline{u}, \underline{A}\underline{u}) = (\underline{u}, \lambda\underline{u}) = \bar{\lambda}(\underline{u}, \underline{u})$$

since \underline{A} is self-adjoint. Since $\underline{u} \neq \underline{0}$, $(\underline{u}, \underline{u}) \neq 0$ as well and so, $\lambda = \bar{\lambda}$ showing that λ is real.

(b) Let λ_1 and λ_2 be distinct eigenvalues of \underline{A} with associated eigenvectors \underline{u}_1 and \underline{u}_2 . Then

$$\begin{aligned}\lambda_1(\underline{u}_1, \underline{u}_2) &= (\lambda_1 \underline{u}_1, \underline{u}_2) = (\underline{A}\underline{u}_1, \underline{u}_2) = (\underline{u}_1, \underline{A}\underline{u}_2) = (\underline{u}_1, \lambda_2 \underline{u}_2) = \bar{\lambda}_2(\underline{u}_1, \underline{u}_2) \\ &= \lambda_2(\underline{u}_1, \underline{u}_2),\end{aligned}$$

since \underline{A} is self-adjoint and λ_2 is real. Therefore

$$(\lambda_1 - \lambda_2) \cdot (\underline{u}_1, \underline{u}_2) = 0$$

which implies that

$$(\underline{u}_1, \underline{u}_2) = 0.$$

c) We will next show by finite induction that for all $k \leq \dim(V)$, there is as orthonormal system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_k\}$ consisting of eigenvectors of \underline{A} . For $k = 1$ the assertion is obvious, since \underline{A} has at least one eigenvalue with associated eigenvector \underline{v} . Then $\underline{u}_1 = \frac{1}{(\underline{v}, \underline{v})} \underline{v}$ satisfies the assertion. Assume next that there is an orthonormal system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{k-1}\}$ of eigenvectors of \underline{A} . Let V_1 denote the subspace generated by this orthonormal system, and let W be the orthogonal complementary subspace of V_1 . For $i = 1, 2, \dots, k-1$ and arbitrary $\underline{x} \in W$,

$$(\underline{A}\underline{x}, \underline{u}_i) = (\underline{x}, \underline{A}\underline{u}_i) = (\underline{x}, \lambda_i \underline{u}_i) = \lambda_i (\underline{x}, \underline{u}_i) = 0,$$

where λ_i is the real eigenvalue associated to \underline{u}_i . Therefore, $\underline{A}\underline{x}$ is orthogonal to the basis of V_1 , which implies that $\underline{A}\underline{x} \in W$. That is, W is an A -invariant subspace of V . Consider now the restriction of transformation A into subspace W . It is also self-adjoint, and has at least one eigenvalue with associated eigenvector \underline{u}_k of unit length. Then system

$\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{k-1}, \underline{u}_k\}$ is an orthonormal system of eigenvectors. The proof is now complete. ♣

Corollary. If matrix \underline{A} is self-adjoint, then its matrix representation in basis $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n\}$ (where $n = \dim(V)$) is diagonal. The diagonal elements are the real eigenvalues, and the columns of the transformation matrix form an orthonormal system. If matrix \underline{A} has complex elements, then \underline{T} also has complex elements. However if \underline{A} is a real symmetric matrix, then all eigenvectors can be given as real vectors since the eigenvalues are real, and if λ is an eigenvalue then the solutions of the homogeneous linear system $(\underline{A} - \lambda \underline{I})\underline{u} = \underline{0}$ determining the associated eigenvectors have real solutions. Hence the transformation matrix can be selected as a real matrix.

Example 7.2. Consider the 2×2 symmetric matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The characteristic polynomial of \underline{A} is the following:

$$\varphi(\lambda) = \det(\underline{A} - \lambda \underline{I}) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 1 & 1 - \lambda \end{pmatrix} = \lambda^2 - 2\lambda,$$

therefore the eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 2$, both are real numbers. If $\underline{v}_1 = (v_{1i})$ and $\underline{v}_2 = (v_{2i})$ are eigenvectors associated to λ_1 and λ_2 , then they satisfy equations

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We may therefore select

$$\underline{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The orthonormal system of eigenvectors can be obtained by normalizing the eigenvectors to unit lengths. Notice that $|\underline{v}_1| = |\underline{v}_2| = \sqrt{2}$, therefore we may select

$$\underline{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \quad \text{and} \quad \underline{u}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

The transformation matrix \underline{T} is given as the matrix with columns \underline{u}_1 and \underline{u}_2 :

$$\underline{T} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{with inverse} \quad \underline{T}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

The matrix representation of \underline{A} in basis $\{\underline{u}_1, \underline{u}_2\}$ is diagonal:

$$\underline{T}^{-1} \underline{AT} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

◆

Consider next a unitary matrix \underline{A} . Since for all $\underline{x}, \underline{y} \in V$,

$$(\underline{Ax}, \underline{Ay}) = (\underline{x}, \underline{A}' \underline{Ay}) = (\underline{x}, \underline{Iy}) = (\underline{x}, \underline{y}), \quad (7.7)$$

unitary transformations preserve the values of the inner products. If we select $\underline{y} = \underline{x}$, then this equation shows that $|\underline{Ax}| = |\underline{x}|$, that is, unitary transformations preserve the lengths of the elements of V . These two simple observations also imply that the image of an orthonormal system is also an orthonormal system in V . Notice also that the columns of the matrix representation \underline{A} of the unitary transformation A form an orthonormal system, since equation (7.7) implies that for all i and j , $(A(\underline{x}_i), A(\underline{x}_j)) = (\underline{x}_i, \underline{x}_j)$. We may also show that this property holds only for unitary transformations. Assume that $B = \{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ is an orthonormal system in V , and $\{A(\underline{x}_1), A(\underline{x}_2), \dots, A(\underline{x}_n)\}$ is also an orthonormal system. This simple fact follows from the observation that the columns of the matrix representation of A in basis B are $A(\underline{x}_1), A(\underline{x}_2), \dots, A(\underline{x}_n)$ which form an orthonormal system. Further properties of unitary matrices are given in the following theorem.

Theorem 7.3. Let \underline{A} be a unitary matrix. Then

- (a) The absolute values of the eigenvalues of \underline{A} are equal to one;
- (b) There is an orthonormal basis in V that consists of eigenvectors of \underline{A} .

Proof. (a) Let λ and \underline{v} be a pair of an eigenvalue and associated eigenvector of \underline{A} . Then

$$(\underline{v}, \underline{v}) = (\underline{A}\underline{v}, \underline{A}\underline{v}) = (\lambda\underline{v}, \lambda\underline{v}) = \lambda(\underline{v}, \lambda\underline{v}) = \lambda\bar{\lambda}(\underline{v}, \underline{v}).$$

Since $(\underline{v}, \underline{v}) \neq 0$, $\lambda\bar{\lambda} = |\lambda|^2 = 1$.

(b) The same proof can be repeated here as the one used in part (c) of Theorem 7.2. We will verify again that for all $k \leq \dim(V)$ there is an orthonormal system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_k\}$ consisting of eigenvectors of \underline{A} . For $k=1$, the assertion is obvious, since $\underline{v}_1 = \frac{1}{(\underline{v}, \underline{v})}\underline{v}$ is again an appropriate selection, where \underline{v} is an eigenvector of \underline{A} . Assume next that $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{k-1}\}$ is an orthonormal system consisting of eigenvectors of \underline{A} . Let V_1 be the subspace generated by this orthonormal system, and let W be the orthogonal complementary subspace of V_1 . We can show that W is A -invariant, since for all $\underline{x} \in W$,

$$(\underline{A}\underline{x}, \underline{u}_i) = (\underline{x}, \underline{A}'\underline{u}_i) = (\underline{x}, \underline{A}^{-1}\underline{u}_i) = \left(\underline{x}, \frac{1}{\lambda_i}\underline{u}_i \right) = \frac{1}{\lambda_i}(\underline{x}, \underline{u}_i) = 0,$$

since if λ_i is an eigenvalue of \underline{A} with associated eigenvector \underline{u}_i , then $\frac{1}{\lambda_i}$ is an eigenvalue of \underline{A}^{-1} with the same associated eigenvector. Therefore $\underline{A}\underline{x}$ is orthogonal to all basis elements of V_1 , which implies that $\underline{A}\underline{x} \in W$. The restriction of transformation A into W has at least one eigenvector \underline{u}_k of unit length, and system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_k\}$ is orthonormal and all \underline{u}_i ($i=1, 2, \dots, k$) are eigenvectors of \underline{A} . Hence the proof is complete. \clubsuit

Remark. If \underline{A} is a real orthogonal matrix, then it can be transformed into diagonal form, however the diagonal matrix as well as the transformation matrix might have complex elements.

Example 7.3. Consider now the matrix

$$\underline{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

which is obviously orthogonal. The characteristic polynomial of \underline{A} is the following:

$$\varphi(\underline{A}) = \det(\underline{A} - \lambda \underline{I}) = \det \begin{pmatrix} \frac{1}{\sqrt{2}} - \lambda & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{pmatrix} = \lambda^2 - \frac{2}{\sqrt{2}}\lambda + 1,$$

and therefore the eigenvalues are

$$\lambda_1 = \frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}} \quad \text{and} \quad \lambda_2 = \frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}.$$

It is clear that $|\lambda_1| = |\lambda_2| = 1$. The associated eigenvectors $\underline{v}_1 = (v_{1i})$ and $\underline{v}_2 = (v_{2i})$ satisfy the homogeneous equations

$$\begin{pmatrix} -\frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and therefore we may select

$$\underline{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Since $|\underline{v}_1| = |\underline{v}_2| = \sqrt{2}$, we may select $\underline{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}$ and $\underline{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \end{pmatrix}$.

The complex inner product of these normalized eigenvectors is as follows:

$$(\underline{u}_1, \underline{u}_2) = \frac{1}{2} (\underline{v}_1, \underline{v}_2) = \frac{1}{2} \underline{v}_1^T \overline{\underline{v}_2} = \frac{1}{2} (i \cdot (i) + 1 \cdot (1)) = 0,$$

that is $\{\underline{u}_1, \underline{u}_2\}$ is an orthonormal basis.

◆

Theorems 7.2 and 7.3 imply that self-adjoint and unitary matrices can be transformed into diagonal form by unitary transformation matrices. A common generalization of these very important matrix types is obtained by assuming that there is an orthonormal basis B in V such that the matrix representation \underline{A} of mapping A in basis B is diagonal. Then \underline{A}' is also diagonal, and its diagonal elements are the complex conjugates of the corresponding diagonal elements of \underline{A} . Since the product of diagonal matrices is commutative, we conclude that $\underline{A}' \underline{A} = \underline{A} \underline{A}'$, that is, \underline{A} must be a normal matrix. The next theorem shows that the converse of this simple result also holds.

Theorem 7.4. Let \underline{A} be a complex normal matrix. Then there exists an orthonormal basis $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n\}$ of V such that all elements \underline{u}_i are common eigenvectors of \underline{A} and \underline{A}' .

Proof. Similar to the proofs of Theorems 7.2 and 7.3 we will prove that for all $k \leq \dim(V)$, there is an orthonormal system $\{\underline{u}_1, \dots, \underline{u}_k\}$ such that all elements \underline{u}_i are common eigenvectors of \underline{A} and \underline{A}' . We will use finite induction. Assume first that $k=1$. Matrix \underline{A} has an eigenvalue λ with associated eigenvector \underline{u} . First we show that $\underline{A}'\underline{u}$ is also an eigenvector associated to λ :

$$\begin{aligned} (\underline{A} - \lambda I)(\underline{A}'\underline{u}) &= \underline{A}\underline{A}'\underline{u} - \lambda\underline{A}'\underline{u} = \underline{A}'\underline{A}\underline{u} - \lambda\underline{A}'\underline{u} \\ &= \underline{A}'(\underline{A}\underline{u} - \lambda\underline{u}) = 0. \end{aligned}$$

Consider next the subspace generated by the eigenvectors of \underline{A} associated to λ . The restriction of transformation \underline{A}' into this subspace is a linear transformation, which has at least one eigenvector \underline{u}_1 , which is a common eigenvector of \underline{A} and \underline{A}' (since all vectors of this subspace are eigenvectors of \underline{A}). Let μ be the eigenvalue of \underline{A}' associated to \underline{u}_1 . Then

$$(\underline{A}\underline{u}_1, \underline{u}_1) = (\lambda\underline{u}_1, \underline{u}_1) = \lambda(\underline{u}_1, \underline{u}_1)$$

and similarly

$$(\underline{A}\underline{u}_1, \underline{u}_1) = (\underline{u}_1, \underline{A}'\underline{u}_1) = (\underline{u}_1, \mu\underline{u}_1) = \bar{\mu}(\underline{u}_1, \underline{u}_1)$$

Since $(\underline{u}_1, \underline{u}_1) \neq 0$, $\lambda = \bar{\mu}$. Hence the eigenvalues are complex conjugates of each other.

Assume next that we have an orthonormal system $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{k-1}\}$ such that each element \underline{u}_i is a common eigenvector of \underline{A} and \underline{A}' . Let V_1 denote the subspace generated by this system, and let W denote the orthogonal complementary subspace of V_1 . Let $\underline{x} \in W$ be an arbitrary element. Then for $i = 1, 2, \dots, k-1$,

$$(\underline{A}\underline{x}, \underline{u}_i) = (\underline{x}, \underline{A}'\underline{u}_i) = (\underline{x}, \bar{\lambda}_i \underline{u}_i) = \lambda_i (\underline{x}, \underline{u}_i) = 0,$$

and

$$(\underline{A}'\underline{x}, \underline{u}_i) = (\underline{x}, \underline{A}\underline{u}_i) = (\underline{x}, \lambda_i \underline{u}_i) = \bar{\lambda}_i (\underline{x}, \underline{u}_i) = 0.$$

Therefore subspace W is A -invariant as well as A' -invariant. Consider next the restriction of transformations \underline{A} and \underline{A}' into W , which have at least one common eigenvector \underline{u}_k . The case of $k = 1$ discussed above provides the proof of the existence of \underline{u}_k . Then $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_{k-1}, \underline{u}_k\}$ is an orthonormal system. Thus the proof is complete. +

Remark 1. A linear transformation A on a complex vector space V is normal if and only if V has an orthonormal basis consisting of common eigenvectors of \underline{A} and \underline{A}' . If a normal mapping is real, then both the transformation matrix and the diagonal form may contain complex elements similarly to the case of real orthogonal matrices.

Example 7.4. In this example we will find all 2×2 real normal matrices. Assume that

$$\underline{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then

$$\underline{A}\underline{A}^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}$$

and

$$\underline{A}^T \underline{A} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

Matrix \underline{A} is normal if and only if

$$\begin{aligned} a^2 + b^2 &= a^2 + c^2 \\ ac + db &= ab + cd \\ c^2 + d^2 &= b^2 + d^2. \end{aligned}$$

This equations are equivalent to the following two relations:

$$\begin{aligned} b^2 &= c^2 \\ (a - d)(b - c) &= 0. \end{aligned} \tag{7.8}$$

If $b = c$, then \underline{A} is symmetric, and if $b = -c$, then $a = d$. In the second case \underline{A} is the sum of $a\underline{I}$ and a skew symmetric matrix, and it is a constant multiple of an orthogonal matrix.

7.5 Definite Matrices

In this section special self-adjoint matrices will be introduced and examined. Let \underline{A} be the $n \times n$ matrix representation of a self-adjoint mapping in an orthonormal basis of a complex vector space.

Definition 7.4. Introduce the following terminology:

- (i) \underline{A} is called *positive definite* if $(\underline{Ax}, \underline{x}) > 0$ for all $\underline{x} \neq \underline{0}$, $\underline{x} \in C^n$;
- (ii) \underline{A} is called *positive semidefinite* if $(\underline{Ax}, \underline{x}) \geq 0$ for all $\underline{x} \in C^n$;
- (iii) \underline{A} is called *negative definite* if $(\underline{Ax}, \underline{x}) < 0$ for all $\underline{x} \neq \underline{0}$, $\underline{x} \in C^n$;
- (iv) \underline{A} is called *negative semidefinite* if $(\underline{Ax}, \underline{x}) \leq 0$ for all $\underline{x} \in C^n$.

We can show that in the case of a real symmetric matrix \underline{A} it is sufficient to check if the conditions of the above definition hold for only real vectors \underline{x} . If \underline{x} is complex, then it can be written as $\underline{x} = \underline{y} + i\underline{z}$ with real \underline{y} and \underline{z} , and

$$\begin{aligned} (\underline{Ax}, \underline{x}) &= (\underline{A}(\underline{y} + i\underline{z}), \underline{y} + i\underline{z}) = (\underline{Ay}, \underline{y}) + (i\underline{Az}, \underline{y}) + (\underline{Ay}, i\underline{z}) + (i\underline{Az}, i\underline{z}) \\ &= (\underline{Ay}, \underline{y}) + i(\underline{Az}, \underline{y}) + \bar{i}(\underline{Ay}, \underline{z}) + i\bar{i}(\underline{Az}, \underline{z}) = (\underline{Ay}, \underline{y}) + (\underline{Az}, \underline{z}), \end{aligned}$$

since $\bar{i} = -i$ and $(\underline{Az}, \underline{y}) = (\underline{z}, \underline{A}^T \underline{y}) = (\underline{z}, \underline{Ay}) = (\underline{Ay}, \underline{z})$. The assertion follows from the facts that both vectors \underline{y} and \underline{z} are real, and $\underline{x} \neq \underline{0}$ if and only if at least one of vectors \underline{y} and \underline{z} is nonzero.

From the definition it is clear that \underline{A} is negative definite if and only if $-\underline{A}$ is positive definite, and \underline{A} is negative semidefinite if and only if $-\underline{A}$ is positive semidefinite. Therefore we will discuss only the properties of positive definite and positive semidefinite matrices.

Theorem 7.5. a) A self-adjoint matrix \underline{A} is positive definite if and only if all eigenvalues of \underline{A} are positive;

b) A self-adjoint matrix \underline{A} is positive semidefinite if and only if all eigenvalues of \underline{A} are nonnegative.

Proof. a) Assume first that \underline{A} is positive definite. Let λ be an eigenvalue of \underline{A} with associated eigenvector \underline{u} . Then

$$0 < (\underline{A}\underline{u}, \underline{u}) = (\lambda\underline{u}, \underline{u}) = \lambda(\underline{u}, \underline{u}),$$

and since $(\underline{u}, \underline{u}) > 0$, λ has to be positive. Assume next that all eigenvalues of \underline{A} are positive. Then

$$\underline{A} = \underline{T} \underline{D} \underline{T},$$

where \underline{T} is a unitary matrix, and \underline{D} is diagonal with positive diagonal elements $\lambda_1, \lambda_2, \dots, \lambda_n$. Let \underline{x} be an arbitrary vector. Then

$$(\underline{A}\underline{x}, \underline{x}) = (\underline{T}' \underline{D} \underline{T}\underline{x}, \underline{x}) = (\underline{D}\underline{T}\underline{x}, \underline{T}\underline{x}).$$

Introduce the notation $\underline{y} = \underline{T}\underline{x}$, and denote the elements of \underline{y} by y_1, y_2, \dots, y_n . Then

$$(\underline{A}\underline{x}, \underline{x}) = (\underline{D}\underline{y}, \underline{y}) = \sum_{i=1}^n \lambda_i y_i \bar{y}_i = \sum_{i=1}^n \lambda_i |y_i|^2 > 0,$$

since $\underline{y} \neq \underline{0}$, all terms are nonnegative and at least one term is positive.

b) The proof is the same as given for part a) with the slight modification that inequalities $<$ and $>$ have to be replaced by \leq and \geq , respectively.

♣

Corollary. A self-adjoint matrix \underline{A} is negative definite if and only if all eigenvalues of \underline{A} are negative, and is negative semidefinite if and only if all eigenvalues of \underline{A} are nonpositive.

Let now matrix \underline{B} be an $m \times n$ real matrix. First we show that $\underline{A} = \underline{B}^T \underline{B}$ is positive semidefinite. Let \underline{x} be an arbitrary n -element vector. Then

$$(\underline{Ax}, \underline{x}) = (\underline{B}^T \underline{Bx}, \underline{x}) = (\underline{Bx}, \underline{Bx}) \geq 0.$$

Assume in addition that the columns of \underline{B} are linearly independent, then $\underline{x} \neq \underline{0}$ implies that $\underline{Bx} \neq \underline{0}$, therefore

$$(\underline{Ax}, \underline{x}) = (\underline{Bx}, \underline{Bx}) > 0,$$

which implies that \underline{A} is positive definite. We will next prove that if \underline{A} is a real $n \times n$ positive semidefinite matrix then there exists an $n \times n$ real matrix \underline{B} such that $\underline{A} = \underline{B}^T \underline{B}$, and if in addition \underline{A} is positive definite, then \underline{B} is nonsingular. From Theorem 7.2 we know that

$$\underline{A} = \underline{T}^T \underline{DT}$$

where \underline{T} is an orthogonal matrix and \underline{D} is diagonal with diagonal elements $\lambda_1, \lambda_2, \dots, \lambda_n$. Select

$$\underline{B} = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \dots & \\ & & & \sqrt{\lambda_n} \end{pmatrix} \underline{T},$$

then obviously, $\underline{A} = \underline{B}^T \underline{B}$, and if \underline{A} is positive definite, then $\lambda_i > 0$ for all i , consequently both factors of \underline{B} are nonsingular, hence \underline{B} is also nonsingular. We mention here that in the decomposition $\underline{A} = \underline{B}^T \underline{B}$, matrix \underline{B} can be selected as an upper triangular matrix. For more details on this decomposition see for example, Szidarovszky and Yakowitz (1978).

Definition 7.5. Let \underline{A} be an $n \times n$ matrix, and let $1 \leq i_1 < i_2 < \dots < i_r \leq n$ be arbitrary integers. The determinant of the principal submatrix

$$\begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \dots & a_{i_1 i_r} \\ a_{i_2 i_1} & a_{i_2 i_2} & \dots & a_{i_2 i_r} \\ \dots & \dots & \dots & \dots \\ a_{i_r i_1} & a_{i_r i_2} & \dots & a_{i_r i_r} \end{pmatrix} \quad (7.9)$$

is called the (i_1, i_2, \dots, i_r) -principal minor of \underline{A} .

An important property of definite and semidefinite matrices is given next.

Theorem 7.6. a) All principal minors of a real positive definite matrix are positive;

b) All principal minors of a real positive semidefinite matrix are nonnegative.

Proof. a) First we prove that if \underline{A} is positive definite, then $\det(\underline{A}) > 0$. From Theorem 7.2 we know that there exists a real orthogonal matrix \underline{T} such that

$$\underline{A} = \underline{T} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \underline{T}^{-1},$$

where all eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are positive. Then

$$\det(\underline{A}) = \det(\underline{T}) \lambda_1 \lambda_2 \dots \lambda_n \det(\underline{T}^{-1}) = \lambda_1 \lambda_2 \dots \lambda_n > 0,$$

since

$$\det(\underline{T}) \det(\underline{T}^{-1}) = \det(\underline{T} \underline{T}^{-1}) = \det(\underline{I}) = 1.$$

Next we show that if \underline{A} is positive definite, then all principal submatrices of \underline{A} are also positive definite, which implies the assertion. Notice first that for all \underline{x} , $(\underline{A}\underline{x}, \underline{x}) = \underline{x}^T \underline{A}\underline{x}$. Let \underline{A}_1 be the principal submatrix (7.9) and let \underline{z} be an r -element vector. Define the n -element vector as

$$x_j = \begin{cases} z_i & \text{if } j = i_i \\ 0 & \text{otherwise,} \end{cases}$$

then it is easy to see that

$$\underline{z}^T \underline{A}_1 \underline{z} = \underline{x}^T \underline{A} \underline{x} > 0$$

since $\underline{x} \neq \underline{0}$ if and only if $\underline{z} \neq \underline{0}$.

b) The case of positive semidefinite matrices can be proved in the same way as shown above with the slight differences that the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ are nonnegative, and $\underline{x}^T \underline{A} \underline{x} \geq 0$ for all \underline{x} .

Corollary. If \underline{A} is real negative definite, then the principal minors of even order are all positive and the principal minors of odd order are all negative. Similarly, if \underline{A} is real and negative semidefinite, then the even principal minors are all nonnegative, and those of odd order are all nonpositive. ♣

Some important extensions of definite and semidefinite matrices will be discussed next. First special matrices with certain sign pattern on the principal minors are introduced.

Definition 7.6. Let \underline{A} be an $n \times n$ real matrix.

- (i) \underline{A} is called a *P-matrix* if all principal minors of \underline{A} are positive;
- (ii) \underline{A} is called an *N-matrix* if all principal minors of \underline{A} are negative;
- (iii) \underline{A} is called an *N-P-matrix* if all principal minors of \underline{A} of odd orders are negative and those of even orders are positive;
- (iv) \underline{A} is called a *P-N-matrix* if all principal minors of \underline{A} of odd orders are positive and those of even orders are negative.

In the economic literature, *N-P*-matrices are sometimes called *Hicksian*. It is obvious that \underline{A} is an *N-P*-matrix if and only if $-\underline{A}$ is a *P*-matrix, and \underline{A} is a *P-N*-matrix if and only if $-\underline{A}$ is an *N*-matrix. These facts are the consequences of the simple properties of determinants that for any $m \times m$ real matrix \underline{B} , $\det(-\underline{B}) = (-1)^m \det(\underline{B})$. In addition, Definition 7.6 implies that every principal submatrix of a matrix \underline{A} is a *P*-matrix, an *N*-matrix, an

N - P -matrix, or a P - N -matrix, according to \underline{A} is a P -matrix, an N -matrix, an N - P -matrix, or a P - N -matrix. Since the principal minors of the principal submatrices of \underline{A} are principal minors of \underline{A} we conclude that the following assertion holds. Let $S \subseteq \{1, 2, \dots, n\}$ be any set of integers, and let \underline{E} be the matrix obtained by replacing the j^{th} column e_j of the identity matrix by $-e_j$ for all $j \in S$. Then $\underline{E}^{-1} \underline{A} \underline{E}$ is a P -matrix, an N -matrix, an N - P -matrix, or a P - N -matrix, respectively, according to \underline{A} is a P -matrix, an N -matrix, an N - P -matrix, or a P - N -matrix.

In the definition of definite and semidefinite matrices we have always assumed that the matrix is symmetric. By dropping this assumption, more general matrix classes can be introduced.

Definition 7.7. Let \underline{A} be an $n \times n$ real matrix

- (i) \underline{A} is called *positive quasidefinite*, if $(\underline{A}\underline{x}, \underline{x}) > 0$ for all $\underline{x} \neq \underline{0}$; $\underline{x} \in R^n$;
- (ii) \underline{A} is called *positive semi-quasidefinite*, if $(\underline{A}\underline{x}, \underline{x}) \geq 0$ for all $\underline{x} \in R^n$;
- (iii) \underline{A} is called *negative quasidefinite*, if $(\underline{A}\underline{x}, \underline{x}) < 0$ for all $\underline{x} \neq \underline{0}$; $\underline{x} \in R^n$;
- (iv) \underline{A} is called *negative semi-quasidefinite*, if $(\underline{A}\underline{x}, \underline{x}) \leq 0$ for all $\underline{x} \in R^n$.

It is clear from the definition that \underline{A} is negative quasidefinite if and only if $-\underline{A}$ is positive quasidefinite, and \underline{A} is negative semi-quasidefinite if and only if $-\underline{A}$ is positive semi-quasidefinite. If \underline{A} is an $n \times n$ real matrix, then define

$$\underline{A}^{(1)} = \frac{1}{2}(\underline{A} + \underline{A}^T)$$

and

$$\underline{A}^{(2)} = \frac{1}{2}(\underline{A} - \underline{A}^T).$$

Obviously, $\underline{A}^{(1)}$ is symmetric and $\underline{A}^{(2)}$ is skew-symmetric. Matrices $\underline{A}^{(1)}$ and $\underline{A}^{(2)}$ are therefore called the *symmetric* and *skew-symmetric parts* of \underline{A} . Since for all \underline{x} ,

$$(\underline{A}^{(2)} \underline{x}, \underline{x}) = (\underline{x}, \underline{A}^{(2)T} \underline{x}) = (\underline{x}, -\underline{A}^{(2)} \underline{x}) = (-\underline{A}^{(2)} \underline{x}, \underline{x}) = -(\underline{A}^{(2)} \underline{x}, \underline{x}),$$

we conclude that $(\underline{A}^{(2)} \underline{x}, \underline{x}) = 0$. This relation implies that

$$(\underline{A} \underline{x}, \underline{x}) = ((\underline{A}^{(1)} + \underline{A}^{(2)}) \underline{x}, \underline{x}) = (\underline{A}^{(1)} \underline{x}, \underline{x}) + (\underline{A}^{(2)} \underline{x}, \underline{x}) = (\underline{A}^{(1)} \underline{x}, \underline{x}).$$

From this observation we conclude that an $n \times n$ real matrix is positive quasidefinite, negative quasidefinite, positive semi-quasidefinite, or negative semi-quasidefinite if and only if the symmetric part of \underline{A} is positive definite, negative definite, positive semidefinite, or negative semidefinite, respectively.

7.6 Nonnegative Matrices

In many linear models the nonnegativity of certain matrices has an important role. For example, in linear input-output models the input \underline{x} and output \underline{y} satisfy equation

$$\underline{x} = (\underline{I} - \underline{A})^{-1} \underline{y},$$

where \underline{A} is the input coefficient matrix. If $(\underline{I} - \underline{A})^{-1}$ is nonnegative (which means that all elements of the matrix are nonnegative), then for arbitrary nonnegative output vector \underline{y} there exists nonnegative input which produces that output. In this section the most important properties of nonnegative matrices will be discussed. First, the fundamentals of the Perron-Frobenius theory will be outlined, and then, matrices with nonnegative inverses will be discussed.

Let \underline{A} be an $n \times n$ matrix such that $\underline{A} > \underline{O}$. This assumption means that $a_{ij} > 0$ for all i and j . Let $\lambda_1, \lambda_2, \dots, \lambda_r$ denote the distinct eigenvalues of \underline{A} and assume that they are numbered as

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_r|$$

Our first theorem below presents the most important properties of the dominant eigenvalue λ_1 . In the matrix theory literature this result is known as the famous Perron-Frobenius theorem.

- Theorem 7.7.** (i) $\lambda_1 > 0$;
(ii) λ_1 has a positive associated eigenvector \underline{x}_1 ;
(iii) The eigenvector associated to λ_1 is unique up to a scalar multiplier;
(iv) For all $l \geq 2$, $|\lambda_l| < \lambda_1$.

Proof. Consider the set

$$\Lambda = \{\lambda \in R \mid \underline{Ax} \geq \lambda \underline{x} \text{ with some nonzero vector } \underline{x} \geq \underline{0}\}. \quad (7.10)$$

Note first that $0 \in \Lambda$, since $\underline{Ax} \geq \underline{0}$ with all $\underline{x} \geq \underline{0}$. We prove next that Λ is bounded from above. If $\lambda \in \Lambda$, then for all i ,

$$\lambda x_i \leq \sum_{j=1}^n a_{ij} x_j \quad (7.11)$$

with some nonzero vector $\underline{x} \geq \underline{0}$. Let $x_{i_0} = \max_j \{x_j\}$, then inequality (7.11)

with $i = i_0$ implies that

$$\lambda \leq \sum_{j=1}^n a_{i_0 j} \frac{x_j}{x_{i_0}} \leq \sum_{j=1}^n a_{i_0 j}.$$

Since the defining inequality of Λ is closed and Λ is nonempty and bounded, set Λ has a maximal element λ_0 . We prove that this maximal element is the dominant eigenvalue of matrix \underline{A} , and it satisfies properties (i)–(iv).

(a) First we show that $\lambda_0 > 0$. Note first that if $\underline{x} \geq \underline{0}$ is nonzero, then it has at least one positive component, and therefore $\underline{Ax} > \underline{0} = \underline{0x}$. Therefore

selecting a sufficiently small positive λ , \underline{Ax} is still greater than $\lambda \underline{x}$. Since λ_0 is the largest of such λ values, $\lambda_0 > 0$.

(b) Next we prove that λ_0 is an eigenvalue of \underline{A} . Let \underline{x}_0 be a nonnegative, nonzero vector such that $\underline{Ax}_0 \geq \lambda_0 \underline{x}_0$. Assume that $\underline{Ax}_0 \neq \lambda_0 \underline{x}_0$, then $\underline{Ax}_0 - \lambda_0 \underline{x}_0$ is nonnegative and nonzero. Therefore

$$0 < \underline{A}(\underline{Ax}_0 - \lambda_0 \underline{x}_0) = \underline{A}(\underline{Ax}_0) - \lambda_0 (\underline{Ax}_0).$$

Denote $\underline{y} = \underline{Ax}_0$, then $\underline{y} > \underline{0}$ and $\underline{Ay} > \lambda_0 \underline{y}$. Therefore with sufficiently small $\varepsilon > 0$, $\underline{Ay} > (\lambda_0 + \varepsilon) \underline{y}$, which contradicts the selection of λ_0 . Therefore $\underline{Ax}_0 = \lambda_0 \underline{x}_0$, that is, λ_0 is an eigenvalue with associated eigenvector \underline{x}_0 .

(c) In the above argument we have shown that $\underline{Ax}_0 > \underline{0}$, $\lambda_0 > 0$, and $\underline{Ax}_0 = \lambda_0 \underline{x}_0$, which imply that

$$\underline{x}_0 = \frac{1}{\lambda_0} \underline{Ax}_0 > \underline{0}.$$

(d) Let λ_i be an eigenvalue of \underline{A} . Then

$$\underline{Ay} = \lambda_i \underline{y}$$

with some $\underline{y} \neq \underline{0}$. Define $\underline{z} = \underline{Ay}$, $\underline{z}^* = (|z_i|)$, and $\underline{y}^* = (|y_i|)$, where z_i and y_i are the i^{th} elements of vectors \underline{z} and \underline{y} , respectively. Then

$$\begin{aligned} \underline{A} \cdot \underline{y}^* &= \left(\sum_{j=1}^n a_{ij} |y_j| \right)_{i=1}^n = \left(\sum_{j=1}^n |a_{ij} y_j| \right)_{i=1}^n \\ &\geq \left(\sum_{j=1}^n a_{ij} y_j \right)_{i=1}^n = (\underline{z}_i)_{i=1}^n = (\lambda_i \underline{y}_i)_{i=1}^n = |\lambda_i| \cdot \underline{y}^*, \end{aligned}$$

which implies that $|\lambda_i| \in \Lambda$. Since λ_0 is the maximal element in Λ , $|\lambda_i| \leq \lambda_0$.

(e) We prove next that $|\lambda_i| < \lambda_0$. If $\delta > 0$ is sufficiently small, then $\underline{A}_\delta = \underline{A} - \delta \underline{I} > \underline{0}$. The previous part of the proof implies that

$$|\lambda_i - \delta| \leq \lambda_0 - \delta, \quad (7.12)$$

since the eigenvalues of $\underline{A} - \delta \underline{I}$ are $\lambda_1 - \delta, \dots, \lambda_r - \delta$. Contrary to the assertion that $|\lambda_i| = \lambda_0$. Let $\lambda_i = a + ib$, where a and b are the real and imaginary parts of λ_i , respectively. Since $|\lambda_i| = \sqrt{a^2 + b^2}$, $a^2 + b^2 = \lambda_0^2$. Therefore, if $\lambda \neq \lambda_0$, then $a < \lambda_0$. This inequality implies that

$$\begin{aligned} |\lambda_i - \delta|^2 &= |(a - \delta) + ib|^2 = (a - \delta)^2 + b^2 = a^2 + b^2 + \delta^2 - 2a\delta \\ &> a^2 + b^2 + \delta^2 - 2\delta\lambda_0 = \lambda_0^2 + \delta^2 - 2\delta\lambda_0 = (\lambda_0 - \delta)^2, \end{aligned}$$

which contradicts relation (7.12).

(f) Assume finally that \underline{x}_0 and \underline{x} are linearly independent eigenvectors associated to λ_0 . Since $\underline{x}_0 > \underline{0}$, there exists a scalar α such that vector $\underline{z} = \alpha \underline{x}_0 + \underline{x}$ is nonnegative, nonzero, and at least one element z_i of \underline{z} is

zero. Observe that selection $\alpha = \max \left\{ -\frac{x_i}{x_{0i}} \right\}$ is satisfactory, where $\underline{x} = (x_i)$ and $\underline{x}_0 = (x_{0i})$. Then

$$\begin{aligned} \lambda_0 \underline{z} &= \alpha \lambda_0 \underline{x}_0 + \lambda_0 \underline{x} = \alpha \underline{A} \underline{x}_0 + \underline{A} \underline{x} \\ &= \underline{A}(\alpha \underline{x}_0 + \underline{x}) = \underline{A} \underline{z} > \underline{0}, \end{aligned}$$

which implies that $\underline{z} > \underline{0}$, which contradicts the assumption that at least one element of \underline{z} is zero.

+

Hence the proof is complete.

Remark. λ_1 is usually called the *Perron eigenvalue* of \underline{A} .

Corollary. Assume that \underline{A} is an $n \times n$ real matrix such that \underline{A}^{-1} exists and is positive. Since the distinct eigenvalues of \underline{A}^{-1} are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_r}$ and they satisfy conditions (i)-(iv) of the theorem, we conclude that

(i) $\lambda_r > 0$;

(ii) λ_r has a positive associated eigenvector \underline{x}_r ;

(iii) The eigenvector associated to λ_r is unique up to a scalar multiplier;

(iv) For all $l \leq r-1, |\lambda_l| > \lambda_r$.

Assume now that $\underline{A} \geq \underline{0}$. Since the eigenvalues of real matrices continuously depend on the matrix elements, Theorem 7.7 implies the following result.

Theorem 7.8. Let \underline{A} be a nonnegative matrix with distinct eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_r|$. Then

- (i) $\lambda_1 \geq 0$;
- (ii) λ_1 has a nonnegative associated eigenvector \underline{x}_1 ;
- (iii) For all $l \geq 2, |\lambda_l| \leq \lambda_1$.

An alternative extension of Theorem 7.7 can be formulated as follows.

Theorem 7.9. Assume that $\underline{A} \geq \underline{O}$, and $\underline{A}^m > \underline{O}$ for some positive integer m . Then all conditions (i)-(iv) of Theorem 7.7 apply to \underline{A} .

Proof. Let λ_1 be the dominant eigenvalue of \underline{A} with associated eigenvector \underline{x}_1 . Then the assertion follows immediately from the fact that λ_1^m is the dominant eigenvalue of \underline{A}^m with the same associated eigenvector.

♣

In many applications it is very useful to compute lower and upper bounds for the dominant eigenvalues λ_1 of nonnegative matrices. Such a result is presented in the following theorem.

Theorem 7.10. Let $\underline{A} = (a_{ij})$ be an $n \times n$ nonnegative matrix with dominant eigenvalue λ_1 . Then

$$\min_j c_j \leq \lambda_1 \leq \max_j c_j, \quad (7.13)$$

where

$$c_j = \sum_{i=1}^n a_{ij}, \quad j = 1, 2, \dots, n.$$

Proof. Assume that the nonnegative eigenvector $\underline{x}_1 (= x_{1i})$ associated to λ_1 is normalized so that $\underline{1}^T \underline{x}_1 = 1$, where $\underline{1}^T = (1, 1, \dots, 1)$. Multiply the eigenvector equation $\underline{A}\underline{x}_1 = \lambda_1 \underline{x}_1$ by $\underline{1}^T$ to get

$$\underline{1}^T \underline{A}\underline{x}_1 = \lambda_1 \underline{1}^T \underline{x}_1 = \lambda_1.$$

Observe that

$$\underline{1}^T \underline{A} = (c_1, c_2, \dots, c_n),$$

and therefore

$$\lambda_1 = c_1 x_{11} + c_2 x_{12} + \dots + c_n x_{1n}.$$

Since all components x_{1i} of \underline{x}_1 are nonnegative,

$$\lambda_1 \geq \left(\min_j c_j \right) \cdot (x_{11} + x_{12} + \dots + x_{1n}) = \min_j c_j$$

and

$$\lambda_1 \leq \left(\max_j c_j \right) \cdot (x_{11} + x_{12} + \dots + x_{1n}) = \max_j c_j.$$

+

Corollary. Let $\underline{A} = (a_{ij})$ be an $n \times n$ nonnegative matrix with dominant eigenvalue λ_1 . Then

$$\min_i r_i \leq \lambda_1 \leq \max_i r_i, \quad (7.14)$$

where

$$r_i = \sum_{j=1}^n a_{ij}, \quad i = 1, 2, \dots, n.$$

Proof. The eigenvalues of \underline{A} and \underline{A}^T are the same. Apply relation (7.13) to the transpose \underline{A}^T of \underline{A} to get the desired bounds.

Relations (7.13) and (7.14) are illustrated in the following examples. +

Example 7.5. In the case of matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$$

$c_1 = c_2 = 5$, and therefore

$$\min_j c_j = \max_j c_j = 5.$$

Hence relation (7.13) implies that $\lambda_1 = 5$. This result can also be verified by solving the characteristic polynomial

$$\varphi(\lambda) = \det \begin{pmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{pmatrix} = (1-\lambda)(3-\lambda) - 8 = \lambda^2 - 4\lambda - 5.$$

The roots are $\lambda_1 = 5$ and $\lambda_2 = -1$.

Example 7.6. Consider next matrix

$$\underline{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 4 \\ 0 & 1 & 3 \end{pmatrix}$$

where

$$c_1 = 2, c_2 = 2, c_3 = 7 \quad \text{and} \quad r_1 = 1, r_2 = 6, r_3 = 4.$$

Therefore relations (7.13) and (7.14) imply that

$$2 \leq \lambda_1 \leq 7 \quad \text{and} \quad 1 \leq \lambda_1 \leq 6.$$

Note that from these inequalities we conclude that

$$2 \leq \lambda_1 \leq 6.$$

As a comparison we mention that the true value is

$$\lambda_1 = 2 + \sqrt{5} \approx 4.236.$$

◆

In many applications the nonnegativity of inverse matrices is assumed. In the next part of this section conditions will be introduced which guarantee that $\underline{A}^{-1} \geq \underline{0}$ for a given square matrix \underline{A} . We begin our analysis with a simple result.

Theorem 7.11. Let \underline{B} be an $n \times n$ real matrix. The inverse \underline{B}^{-1} exists and is nonnegative if and only if there exists an $n \times n$ real matrix \underline{D} such that

- (i) $\underline{D} \geq \underline{O}$;
- (ii) $\underline{I} - \underline{DB} \geq \underline{O}$;
- (iii) All eigenvalues of $\underline{I} - \underline{DB}$ are inside the unit circle.

Proof. (a) If \underline{B}^{-1} exists and is nonnegative, then select $\underline{D} = \underline{B}^{-1}$. Then $\underline{I} - \underline{DB} = \underline{O}$, and therefore conditions (i), (ii) and (iii) are obviously satisfied.

(b) Assume now the existence of matrix \underline{D} satisfying conditions (i), (ii) and (iii). Note first that condition (iii) and the results of the third application of section 6.8 imply that $\underline{I} - (\underline{I} - \underline{DB}) = \underline{DB}$ is invertible. Therefore \underline{D} and \underline{B} are both nonsingular. From the same application we also know that

$$(\underline{DB})^{-1} = \underline{I} + (\underline{I} - \underline{DB}) + (\underline{I} - \underline{DB})^2 + \dots \geq \underline{O},$$

since each term of the right-hand side is nonnegative as the consequence of condition (ii). Hence, from (i) we conclude that

$$\underline{B}^{-1} = (\underline{D}^{-1} \underline{DB})^{-1} = (\underline{DB})^{-1} \underline{D} \geq \underline{O},$$

which completes the proof. +

Definition 7.8. An $n \times n$ real matrix $\underline{B} = (b_{ij})$ is called an *M-matrix*, if $\underline{D} = \text{diag}(b_{11}, b_{22}, \dots, b_{nn})^{-1}$ satisfies conditions (i), (ii) and (iii) of Theorem 7.11.

Note that in the case of an *M-matrix*,

$$\underline{I} - \underline{DB} = \begin{pmatrix} 0 & -\frac{b_{12}}{b_{11}} & \cdots & -\frac{b_{1n}}{b_{11}} \\ -\frac{b_{21}}{b_{22}} & 0 & \cdots & -\frac{b_{2n}}{b_{22}} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{b_{n1}}{b_{nn}} & -\frac{b_{n2}}{b_{nn}} & \cdots & 0 \end{pmatrix} \quad (7.15)$$

therefore $b_{ii} > 0$ for all i , and $b_{ij} \leq 0$ for all $j \neq i$.

Example 7.7. We will show that matrix

$$\underline{B} = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

is an M -matrix. In this case,

$$\underline{D} = \text{diag}(5, 5, 5)^{-1} = \text{diag}\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right),$$

and

$$\underline{I} - \underline{DB} = \begin{pmatrix} 0 & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & 0 & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & 0 \end{pmatrix}$$

Conditions (i) and (ii) of Theorem 7.11 are obviously satisfied. By using relation (7.13) we see that $c_1 = c_2 = c_3 = \frac{2}{5}$, therefore for all eigenvalues λ_i of $\underline{I} - \underline{DB}$, $|\lambda_i| \leq \frac{2}{5}$. (The true eigenvalues are $\lambda_1 = \lambda_2 = -\frac{1}{5}$ and $\lambda_3 = \frac{2}{5}$.)

◆

In the conclusion of this section a new class of special matrices will be introduced, which has a very close relation to nonnegative matrices.

Definition 7.9. An $n \times n$ real matrix \underline{A} is called a *Metzler-matrix*, if $a_{ij} \geq 0$ for all $i \neq j$.

Notice that Metzler-matrices have nonnegative elements outside the diagonal, and there is no constraint on the diagonal elements. A square matrix \underline{A} is a Metzler matrix if and only if there is a real number α such that $\underline{A} + \alpha \underline{I} \geq \underline{0}$. This strong relation between nonnegative and Metzler matrices makes us able to extend the Perron-Frobenius theorem to Metzler matrices.

Theorem 7.12. Assume that \underline{A} is an $n \times n$ real Metzler matrix. Then there exists a real eigenvalue λ_1 of \underline{A} such that

- (i) λ_1 has a nonnegative associated eigenvector;
- (ii) For any other eigenvalue λ_i of \underline{A} , $\operatorname{Re} \lambda_i < \lambda_1$.

Proof. (i) Select an α such that $\underline{A} + \alpha \underline{I} \geq \underline{0}$, and let μ_1 denote the Perron eigenvalue of $\underline{A} + \alpha \underline{I}$ with associated eigenvector $\underline{x}_1 \geq \underline{0}$. Then $\lambda_1 = \mu_1 - \alpha$ is an eigenvalue of \underline{A} with associated eigenvector \underline{x}_1 .

(ii) Let λ_i denote an eigenvalue of \underline{A} such that $\lambda_i \neq \lambda_1$. Note that $\lambda_i + \alpha$ is nonnegative and is the Perron eigenvalue of $\underline{A} + \alpha \underline{I}$, and for any

other eigenvalue λ_1 of \underline{A} , $\lambda_1 + \alpha$ is also an eigenvalue of $\underline{A} + \alpha \underline{I}$ such that $\lambda_1 + \alpha \neq \lambda_1 + \alpha$.

If $\lambda_1 + \alpha$ is real, then from Theorem 7.8 we know that $|\lambda_1 + \alpha| \leq \lambda_1 + \alpha$. If $\lambda_1 + \alpha$ is nonnegative, then $\lambda_1 + \alpha < \lambda_1 + \alpha$, which implies that $\lambda_1 < \lambda_1$. If $\lambda_1 + \alpha$ is negative, then obviously $\lambda_1 + \alpha < \lambda_1 + \alpha$, that is $\lambda_1 < \lambda_1$, again. Hence, $\text{Re} \lambda_1 = \lambda_1 < \lambda_1$.

If $\lambda_1 = a + ib$ with $b \neq 0$, then from Theorem 7.8 we conclude that

$$|\lambda_1 + \alpha| \leq \lambda_1 + \alpha,$$

that is,

$$|(a + \alpha) + ib| \leq \lambda_1 + \alpha.$$

This relation is equivalent to inequality

$$(a + \alpha)^2 + b^2 \leq (\lambda_1 + \alpha)^2,$$

which implies that

$$(a + \alpha)^2 < (\lambda_1 + \alpha)^2. \quad (7.16)$$

This inequality can be rewritten as

$$(a - \lambda_1)(a + 2\alpha + \lambda_1) < 0. \quad (7.17)$$

Note first that (7.16) implies that

$$|a + \alpha| < \lambda_1 + \alpha,$$

and therefore

$$a + 2\alpha + \lambda_1 > 0.$$

And hence, (7.17) holds only if $a - \lambda_1 < 0$, that is, when $a = \operatorname{Re} \lambda_1 < \lambda_1$. Thus the proof is complete. \clubsuit

Assume again that \underline{A} , is a Metzler matrix. We know that $\underline{A} + \alpha \underline{I}$ is nonnegative with some nonnegative α . It is also known, that the eigenvalues of $\underline{A} + \alpha \underline{I}$ are $\lambda_1 + \alpha$ and the Perron eigenvalue of $\underline{A} + \alpha \underline{I}$ is $\lambda_1 + \alpha$, where λ_1 is the eigenvalue of \underline{A} , which satisfies properties (i) and (ii) of the above theorem. If $\underline{A} = (a_{ij})$, then

$$\underline{A} + \alpha \underline{I} = \begin{cases} a_{ij} + \alpha & \text{if } i = j \\ a_{ij} & \text{if } i \neq j. \end{cases}$$

Use relations (7.13) and (7.14) to this matrix to get inequalities

$$\min_j (c_j + \alpha) \leq \lambda_1 + \alpha \leq \max_j (c_j + \alpha),$$

and

$$\min_j (r_j + \alpha) \leq \lambda_1 + \alpha \leq \max_j (r_j + \alpha)$$

which implies that Theorem 7.10 and its corollary remain true for Metzler matrices.

Example 7.8. Consider matrix

$$\underline{A} = \begin{pmatrix} -2 & 1 \\ 2 & -1 \end{pmatrix}$$

where $c_1 = c_2 = 0$, and therefore $\lambda_1 = 0$. Note that the characteristic polynomial of \underline{A} is

$$\varphi(\lambda) = \det \begin{pmatrix} -2-\lambda & 1 \\ 2 & -1-\lambda \end{pmatrix} = (-2-\lambda)(-1-\lambda) - 2 = \lambda^2 + 3\lambda.$$

The eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = -3$.

Example 7.9. Consider the next matrix

$$\underline{A} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 4 \\ 0 & 1 & 3 \end{pmatrix}$$

where

$$c_1 = 0, c_2 = 2, c_3 = 7 \quad \text{and} \quad r_1 = -1, r_2 = 6, r_3 = 4.$$

Therefore relations (7.13) and (7.14) imply that

$$0 \leq \lambda_1 \leq 7 \quad \text{and} \quad -1 \leq \lambda_1 \leq 6.$$

Hence, $0 \leq \lambda_1 \leq 6$. Note that the true value is

$$\lambda_1 = 2 + \sqrt{5} \approx 4.236.$$

◆

Consider now a Markov chain with n states and the $n \times n$ transition matrix \underline{P} . In section 3.8 we have seen that \underline{P} is a nonnegative matrix, and the sum of the elements in each column is equal to one. Therefore we can use Theorems 7.8 and 7.10 to see that the dominant eigenvalue of \underline{P} is unity, and this eigenvalue has a nonnegative associated eigenvector. The eigenvalue equation with $\lambda_1 = 1$ and eigenvector \underline{x} has the special form

$$\underline{P}\underline{x} = 1 \cdot \underline{x}.$$

Since $\underline{x} \geq \underline{0}$ and $\underline{x} \neq \underline{0}$, $\sum_{k=1}^n x_k > 0$ and vector

$$\underline{x}^* = \frac{1}{\sum_{k=1}^n x_k} \cdot \underline{x}$$

satisfies relations

$$\begin{aligned} \underline{P}\underline{x}^* &= \underline{x}^* \\ x_1^* + x_2^* + \dots + x_n^* &= 1, \end{aligned} \tag{7.18}$$

which is equivalent to relation (3.27). Hence \underline{x}^* is the steady state probability vector.

Consider now an economy with n products and a continuum of consumers of the same type with utility function

$$U(\underline{s}) = \underline{s}^T \underline{Q} \underline{s} + \underline{b}^T \underline{s} + \underline{c}, \quad (7.19)$$

where \underline{Q} is a negative definite matrix. The representative consumer maximizes $U(\underline{s}) - \underline{p}^T \underline{s}$, where \underline{p} is the price vector. This objective function is strictly concave in \underline{s} , and it converges to $-\infty$ if any one of the components of \underline{s} tends to $-\infty$, therefore there is a unique vector $\underline{s} = \underline{d}(\underline{p})$ maximizing the objective function. This vector is called the demand function and can be obtained by simple differentiation. The gradient of $U(\underline{s}) - \underline{p}^T \underline{s}$ is given as $2\underline{Q}\underline{s} + \underline{b} - \underline{p}$, so the optimal \underline{s} can be obtained by solving equation

$$2\underline{Q}\underline{s} + \underline{b} - \underline{p} = 0$$

which implies that

$$\underline{d}(\underline{p}) = \underline{s} = \frac{1}{2} \underline{Q}^{-1} \underline{p} - \frac{1}{2} \underline{Q}^{-1} \underline{b} \quad (7.20)$$

resulting in a linear demand function.

Definite matrices have an important role in determining if a function is convex, or concave. Let $f: \mathbf{R}^n \mapsto \mathbf{R}$ be a twice continuously differentiable

real valued function in an open domain $D \subseteq \mathbf{R}^n$. Construct the Hessian matrix \underline{H} of f in the following way. The (i, j) element of \underline{H} is given by the second-order partial derivative $\partial^2 f / \partial x_i \partial x_j$. Hence

$$\underline{H}(\underline{x}) = \begin{pmatrix} \frac{\partial^2 f(\underline{x})}{\partial x_1^2} & \frac{\partial^2 f(\underline{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\underline{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\underline{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\underline{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\underline{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\underline{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x})}{\partial x_n^2} \end{pmatrix} \quad (7.21)$$

Let now \underline{x}^* be an arbitrary point from D . It is well known from convex analysis, that if $\underline{H}(\underline{x}^*)$ is positive definite, then f is strictly convex in a neighborhood of \underline{x}^* , and if $\underline{H}(\underline{x}^*)$ is negative definite, then f is strictly concave in a neighborhood of \underline{x}^* . In Section 7.5 we have discussed the conditions which have to be checked in order to verify the definiteness of the Hessian.

Let $f: \mathbf{R}^n \mapsto \mathbf{R}$ be a continuously differentiable function on a convex, open domain $D \subseteq \mathbf{R}^n$. We say that \underline{f} is monotone, if for all $\underline{x}, \underline{y} \in D$,

$$(\underline{x} - \underline{y})^T (\underline{f}(\underline{x}) - \underline{f}(\underline{y})) \geq 0. \quad (7.22)$$

In the one-dimensional case this kind of monotonicity coincides with the property that f is nondecreasing. Similarly we say that \underline{f} is strictly monotone, if for all $\underline{x}, \underline{y} \in D$ and $\underline{x} \neq \underline{y}$,

$$(\underline{x} - \underline{y})^T (\underline{f}(\underline{x}) - \underline{f}(\underline{y})) > 0. \quad (7.23)$$

In the one-dimensional case this kind of strict monotonicity coincides with the assumption that f is strictly increasing.

The Jacobian of \underline{f} is the $n \times n$ matrix with (i, j) element $\partial^2 f / \partial x_i \partial x_j$, that is,

$$\underline{J}(\underline{x}) = \begin{pmatrix} \frac{\partial f_1(\underline{x})}{\partial x_1} & \frac{\partial f_1(\underline{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\underline{x})}{\partial x_n} \\ \frac{\partial f_2(\underline{x})}{\partial x_1} & \frac{\partial f_2(\underline{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\underline{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\underline{x})}{\partial x_1} & \frac{\partial f_n(\underline{x})}{\partial x_2} & \cdots & \frac{\partial f_n(\underline{x})}{\partial x_n} \end{pmatrix}$$

where x_i and f_i denote the i^{th} entry of \underline{x} and \underline{f} , respectively. The monotonicity and strict monotonicity of \underline{f} can be checked in the following way. If $\underline{J}(\underline{x}) + \underline{J}(\underline{x})^T$ is nonnegative definite for all $\underline{x} \in D$, then \underline{f} is monotone, and if in addition, $\underline{J}(\underline{x}) + \underline{J}(\underline{x})^T$ is positive definite for all $\underline{x} \in D$, then \underline{f} is strictly monotone.

Monotone and strictly monotone functions have an important role in many fields of optimization. As the conclusion of this section we show a classical result.

Consider a nonlinear complementarity problem

$$\begin{aligned} \underline{x}^T \cdot \underline{f}(\underline{x}) &= 0 \\ \underline{x} &\geq 0, \underline{f}(\underline{x}) \geq 0 \end{aligned}$$

(for details see the earlier section 2.6), and assume that \underline{f} is strictly monotone. Then the nonlinear complementarity problem must not have a multiple solution. In contrary to this assertion assume that \underline{x} and \underline{z} are both solutions, then

$$\begin{aligned} 0 &< (\underline{x} - \underline{z})^T (\underline{f}(\underline{x}) - \underline{f}(\underline{z})) = \underline{x}^T \underline{f}(\underline{x}) - \underline{z}^T \underline{f}(\underline{x}) - \underline{x}^T \underline{f}(\underline{z}) + \underline{z}^T \underline{f}(\underline{z}) \\ &= -\underline{z}^T \underline{f}(\underline{x}) - \underline{x}^T \underline{f}(\underline{z}) \leq 0, \end{aligned}$$

which is an obvious contradiction.

7.7 Applications

In this section two major applications will be first discussed. The first one is an important application of real symmetric matrices to define and compute singular value decompositions of (not necessarily square) matrices. Generalized matrix inverses will then be introduced and some of their applications will be outlined. After the two major applications Markov chains will be revisited and some simple examples will be presented.

1. In this application *singular value decomposition* (SVD) of real matrices will be introduced and examined. Let \underline{A} be a real $m \times n$ matrix. The existence of the SVD is guaranteed by the following theorem.

Theorem 7.13. There exist orthogonal matrices \underline{U} and \underline{V} such that

$$\underline{U}^T \underline{A} \underline{V} = \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix},$$

where \underline{D} is a nonsingular diagonal matrix with positive diagonal elements, which are arranged in nonincreasing order. The size of \underline{D} is the rank of \underline{A} .

Proof. From section 7.5 we know that $\underline{A}^T \underline{A}$ is a positive semidefinite matrix with nonnegative eigenvalues. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$ denote the eigenvalues of $\underline{A}^T \underline{A}$. Denote the corresponding set of orthonormal eigenvectors by $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$, which exists, since $\underline{A}^T \underline{A}$ is symmetric. For $i = 1, 2, \dots, n$,

$$\underline{A}^T \underline{A} \underline{v}_i = \sigma_i^2 \underline{v}_i,$$

where $\lambda_i = \sigma_i^2$ with a nonnegative σ_i . This equation implies that for all i and j ,

$$\underline{v}_j^T \underline{A}^T \underline{A} \underline{v}_i = \sigma_i^2 \underline{v}_j^T \underline{v}_i = \begin{cases} 0 & \text{if } i \neq j \\ \sigma_i^2 & \text{if } i = j, \end{cases} \quad (7.24)$$

from which we also conclude that for $i = r + 1, \dots, n$,

$$\underline{A} \underline{v}_i = \underline{0}, \quad (7.25)$$

since for vector $\underline{z} = \underline{A} \underline{v}_i$, $\underline{z}^T \underline{z} = 0$, and therefore $\sum_{j=1}^n z_j^2 = 0$, where z_j denotes the j^{th} element of \underline{z} . Introduce matrices

$$\underline{V}_1 = (\underline{v}_1, \dots, \underline{v}_r), \quad \underline{V}_2 = (\underline{v}_{r+1}, \dots, \underline{v}_n),$$

and the m -element vectors

$$\underline{u}_i = \frac{1}{\sigma_i} \underline{A}v_i \quad (i=1,2,\dots,r).$$

First we show that these vectors form an orthonormal system. From (7.24) we see that

$$\begin{aligned} \underline{u}_i^T \underline{u}_j &= \frac{1}{\sigma_i} (\underline{A}v_i)^T \frac{1}{\sigma_j} (\underline{A}v_j) = \frac{1}{\sigma_i \sigma_j} (v_i^T \underline{A}^T \underline{A} v_j) \\ &= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \end{aligned}$$

Define $\underline{U}_1 = (\underline{u}_1, \dots, \underline{u}_r)$ and select $\underline{U}_2 = (\underline{u}_{r+1}, \dots, \underline{u}_m)$ such that system $\{\underline{u}_1, \dots, \underline{u}_r, \underline{u}_{r+1}, \dots, \underline{u}_m\}$ is orthonormal. Vectors $\underline{u}_{r+1}, \dots, \underline{u}_m$ can be selected as an orthonormal basis of the orthogonal complementary subspace of the subspace generated by the columns of \underline{U}_1 . Notice that for any $j \geq r+1$ and $i \leq r$,

$$\underline{u}_j^T \underline{A}v_i = \underline{u}_j^T \sigma_i \underline{u}_i = 0,$$

and for any $j \geq r+1$ and $i \geq r+1$,

$$\underline{u}_j^T \underline{A}v_i = 0$$

as a consequence of relation (7.25).

Define matrices

$$\underline{U} = (\underline{U}_1, \underline{U}_2) \quad \text{and} \quad \underline{V} = (\underline{V}_1, \underline{V}_2).$$

Then

$$\begin{aligned} \underline{U}^T \underline{AV} &= \begin{pmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \dots \\ \underline{u}_m^T \end{pmatrix} \underline{A}(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n) = \begin{pmatrix} \frac{1}{\sigma_1} & \underline{v}_1^T \underline{A}^T & & \\ & \dots & & \\ \frac{1}{\sigma_r} & \underline{v}_r^T \underline{A}^T & & \\ & \underline{u}_{r+1}^T & & \\ & \dots & & \\ & \underline{u}_m^T & & \end{pmatrix} (\underline{Av}_1, \underline{Av}_2, \dots, \underline{Av}_n) \\ &= \begin{pmatrix} \frac{1}{\sigma_1} \sigma_1^2 & & & & \\ & \dots & & & \\ & & \frac{1}{\sigma_r} \sigma_r^2 & & \\ & & & 0 & \\ & & & & \dots \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} \underline{D} & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix}. \end{aligned}$$

Notice that the rows of \underline{U}^T as well as the columns of \underline{V} are linearly independent. Therefore (see Exercise 3.10)

$$\text{rank}(\underline{A}) = \text{rank}(\underline{U}^T \underline{A}) = \text{rank}(\underline{U}^T \underline{AV}) = \text{rank}(\underline{D}),$$

which completes the proof. +

$$\varphi(\lambda) = \det \begin{pmatrix} 9 - \lambda & 12 \\ 12 & 16 - \lambda \end{pmatrix} = \lambda^2 - 25\lambda,$$

so the eigenvalues of $\underline{A}^T \underline{A}$ are $\lambda_1 = 25$ and $\lambda_2 = 0$. Therefore $\sigma_1 = 5$ and $\sigma_2 = 0$. Next, the associated eigenvectors will be determined. For eigenvalue λ_1 , the associated eigenvector $\underline{v}_1 = (v_{1i})$ solves equation

$$\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = 25 \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix},$$

that is,

$$\begin{aligned} -16v_{11} + 12v_{12} &= 0 \\ 12v_{11} - 9v_{12} &= 0. \end{aligned}$$

A solution is vector

$$\begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

which can be normalized as

$$\underline{v}_1 = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}.$$

For eigenvalue λ_2 , the associated eigenvector satisfies equation

$$\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = 0 \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix},$$

that is,

$$\begin{aligned} 9v_{21} + 12v_{22} &= 0 \\ 12v_{21} + 16v_{22} &= 0. \end{aligned}$$

A solution is vector

$$\begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix},$$

which can be normalized as

$$\underline{v}_2 = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}.$$

In our case $r = 1$ and

$$\underline{u}_1 = \frac{1}{5} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 0 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Since \underline{u}_1 is the third natural basis vector, we may select

$$\underline{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \underline{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Thus,

$$\underline{u} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \underline{v} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \\ \frac{5}{5} & \frac{5}{5} \end{pmatrix},$$

consequently

$$\begin{aligned} \underline{U}^T \underline{AV} &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \\ \frac{5}{5} & \frac{5}{5} \end{pmatrix} \\ &= \begin{pmatrix} 3 & 4 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \\ \frac{5}{5} & \frac{5}{5} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

which is 3×2 , and $\underline{D} = (5)$ is a 1×1 matrix.

2. The second application will introduce the *pseudoinverses* of (not necessarily square) real matrices. Let \underline{A} be a real $m \times n$ matrix.

Definition 7.10. The pseudoinverse of \underline{A} is an $n \times m$ matrix \underline{X} satisfying the following conditions:

(i) $\underline{AXA} = \underline{A}$;

- (ii) $\underline{XAX} = \underline{X}$;
- (iii) \underline{XA} is symmetric;
- (iv) \underline{AX} is symmetric.

In the special case when \underline{A} is a nonsingular square matrix, $\underline{X} = \underline{A}^{-1}$ satisfies these relations showing that this concept of pseudoinverses of general matrices is a generalization of that of inverses of nonsingular square matrices.

The following theorem guarantees the existence and the uniqueness of the pseudoinverse of any real matrix \underline{A} .

Theorem 7.14. The pseudoinverse of any $m \times n$ real matrix \underline{A} exists and is unique.

Proof. Consider the singular value decomposition (7.20) of \underline{A} , and define matrix

$$\underline{X} = \underline{V} \begin{pmatrix} \frac{1}{\sigma_1} & & & \underline{0} \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ \underline{0} & & & \underline{0} \end{pmatrix} \underline{U}^T, \quad (7.27)$$

where the second factor is an $n \times m$ matrix. We can easily show that this matrix satisfies the conditions of Definition 7.10. By using the notation of Theorem 7.13 we have

$$(i) \quad \underline{AXA} = \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T = \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T = \underline{A},$$

since $\underline{V}^T \underline{V}$ and $\underline{U}^T \underline{U}$ are equal to identity matrices of sizes m and n , respectively.

(ii)

$$\underline{XAX} = \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T = \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T = \underline{X}.$$

$$(iii) \underline{AX} = \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T = \underline{U} \begin{pmatrix} \underline{I} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T,$$

where \underline{I} is the $r \times r$ identity matrix, and the second factor is $m \times m$. This matrix is obviously symmetric.

$$(iv) \underline{XA} = \underline{V} \begin{pmatrix} \underline{D}^{-1} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{U}^T \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T = \underline{V} \begin{pmatrix} \underline{I} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T,$$

where the second factor is now $n \times n$. This matrix is obviously symmetric.

We will next show that matrix (7.27) is the only pseudoinverse of \underline{A} . Assume in contrary that matrix \underline{Y} also satisfies the conditions of Definition 7.10. From conditions (i) and (iv) we see that

$$\underline{A}^T = (\underline{AYA})^T = \underline{A}^T \underline{Y}^T \underline{A}^T = (\underline{YA})^T \underline{A}^T = \underline{YAA}^T. \quad (7.28)$$

Conditions (ii) and (iii) imply that

$$\underline{Y}^T = (\underline{YAY})^T = \underline{Y}^T \underline{A}^T \underline{Y}^T = (\underline{AY})^T \underline{Y}^T = \underline{AYY}^T. \quad (7.29)$$

Since \underline{X} is also a pseudoinverse, it also satisfies the above relation:

$$\underline{A}^T = \underline{X} \underline{A} \underline{A}^T \quad \text{and} \quad \underline{X}^T = \underline{A} \underline{X} \underline{X}^T, \quad (7.30)$$

from which and equation (7.22) we have

$$\underline{0} = \underline{A}^T - \underline{A}^T = \underline{Y} \underline{A} \underline{A}^T - \underline{X} \underline{A} \underline{A}^T = (\underline{Y} - \underline{X}) \underline{A} \underline{A}^T,$$

that is,

$$(\underline{Y} - \underline{X}) \underline{A} \underline{A}^T (\underline{Y} - \underline{X})^T = \underline{0}.$$

This equation implies that

$$(\underline{Y} - \underline{X}) \underline{A} = \underline{0}. \quad (7.31)$$

Notice in addition that from equations (7.23) and (7.24) we see that

$$\underline{Y} - \underline{X} = (\underline{Y} \underline{Y}^T - \underline{X} \underline{X}^T) \underline{A}^T,$$

and therefore

$$(\underline{Y} - \underline{X}) \underline{C} = \underline{0} \quad (7.32)$$

for any matrix \underline{C} the columns of which are orthogonal to the columns of \underline{A} . Finally, observe that equations (7.31) and (7.32) imply that $\underline{Y} - \underline{X} = \underline{0}$, that is $\underline{Y} = \underline{X}$. Hence, the pseudoinverse is unique.

♣

Example 7.11. Consider again the matrix

$$\underline{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix}.$$

From the previous example we know that

$$\underline{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{5}{5} & \frac{5}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix},$$

therefore

$$\begin{aligned} \underline{X} &= \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{3}{25} & 0 & 0 \\ \frac{4}{25} & 0 & 0 \\ \frac{4}{25} & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{4}{25} \end{pmatrix}. \end{aligned}$$

As an exercise, we verify the conditions of Definition 4.10:

(i)

$$\underline{AXA} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \\ \frac{25}{25} & \frac{25}{25} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} = \underline{A};$$

(ii)

$$\underline{XAX} = \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} = \begin{pmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \\ \frac{25}{25} & \frac{25}{25} \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} \\ = \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} = \underline{X};$$

(iii)

$$\underline{AX} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$(iv) \underline{XA} = \begin{pmatrix} 0 & 0 & \frac{3}{25} \\ 0 & 0 & \frac{4}{25} \\ 0 & 0 & \frac{25}{25} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \\ \frac{25}{25} & \frac{25}{25} \end{pmatrix}.$$

Notice that both \underline{AX} and \underline{XA} are symmetric.

7.8 Exercises

1. Which special matrix classes do the following matrices belong to?

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

2. Characterize all real 2×2 symmetric matrices such that $\underline{A}^2 = \underline{A}$.
3. Assume that $\underline{A}^3 = \underline{0}$. Show that

$$(\underline{I} - \underline{A})^{-1} = \underline{I} + \underline{A} + \underline{A}^2.$$

4. If $\underline{A} + \underline{B}$ is symmetric, simplify the following expression:

$$(3(\underline{A} + \underline{B})^T + \underline{A})^T - 3\underline{B}^T.$$

5. Find all values of a and b such that matrix

$$\begin{pmatrix} a-b & a+b \\ a+2b & b \end{pmatrix}$$

is symmetric.

6. Repeat Example 7.1 for system

$$x_1 + 2x_2 = 3$$

$$x_1 + x_2 - x_3 = 2$$

$$x_2 + 2x_3 + x_4 = 1$$

$$x_3 - x_4 = 0$$

7. Find the characteristic polynomial of the coefficient matrix of the previous problem. (Use recursion (7.4).)

8. Prove that if \underline{A} is symmetric, then $a\underline{A}$ is also symmetric.

9. Prove that if \underline{A} and \underline{B} are $n \times n$, symmetric, and $\underline{A}\underline{B} = \underline{B}\underline{A}$, then $\underline{A}\underline{B}$ is symmetric.

10. Prove that if $\underline{A} + \underline{B}$ and $\underline{A} \times \underline{B}$ are symmetric, then \underline{A} and \underline{B} are symmetric.

11. Prove that if \underline{A} is skew symmetric, that \underline{A}^2 is symmetric.

12. Is the sum of two triangular matrices also triangular?

13. Is matrix $\begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}$ unitary?

14. Prove that the inverse of a symmetric real matrix is symmetric.

15. Repeat Example 7.2 for matrix

$$\underline{A} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

16. Repeat Example 7.3 for matrix

$$\underline{A} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix}.$$

17. Check which of the following matrices are positive definite, negative definite, positive semidefinite, or negative semidefinite:

a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix};$

b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix};$

c) $\begin{pmatrix} 8 & 1 \\ 1 & 8 \end{pmatrix};$

d) $\begin{pmatrix} -10 & 0 & 1 \\ 0 & -12 & 0 \\ 1 & 0 & -14 \end{pmatrix}.$

18. Illustrate the Perron-Frobenius Theorem for matrices a), b), and c) of the previous problem.

19. Use Theorem 7.10 and its corollary to bound the Perron-eigenvalues of matrices a), b), and c) of the previous problem.

20. Repeat Example 7.7 for matrix

$$\underline{A} = \begin{pmatrix} 10 & -1 & -1 \\ 0 & 8 & -1 \\ -1 & 0 & 6 \end{pmatrix}.$$

21. Let

$$\underline{A} = \underline{U} \begin{pmatrix} \underline{D} & \underline{O} \\ \underline{O} & \underline{O} \end{pmatrix} \underline{V}^T$$

be the SVD of an $m \times n$ real matrix \underline{A} . Prove that

$$\underline{V}^T (\underline{A}^T \underline{A}) \underline{V} = \text{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0),$$

which is $n \times n$, and

$$\underline{U}^T (\underline{A} \underline{A}^T) \underline{U} = \text{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0),$$

which is $m \times m$.

22. Assume that \underline{A} is an $n \times n$ real symmetric matrix with eigenvalues $\lambda_i, i = 1, 2, \dots, n$. Show that the singular values of \underline{A} are $|\lambda_i|$ for all nonzero eigenvalues.

23. (Continuation of Problem 22) Prove that an $n \times n$ real matrix is nonsingular if and only if all its singular values are nonzero.

24. Find the SVD for matrix

$$\underline{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

25. Find the pseudoinverse of the matrix given in the previous example.

Chapter 8

Elements of Matrix Analysis

8.1 Introduction

In Chapter 2 we have already introduced some elements of matrix analysis such as lengths of vectors, which give the possibility to introduce, for example, convergence and to estimate the speed of convergence. In this chapter we will consider these issues in more detail. We will first generalize vector lengths by introducing vector and matrix norms, and their applications in convergence analysis and perturbation analysis will then be outlined.

8.2 Vector Norms

Let V denote the set of all n -element real (or complex) vectors.

Definition 8.1. A real valued function $\underline{x} \mapsto \|\underline{x}\|$ defined on V is called a vector norm, if it satisfies the following conditions:

- (i) $\|\underline{x}\| \geq 0$ for all $\underline{x} \in V$, and $\|\underline{x}\| = 0$ if and only if $\underline{x} = \underline{0}$;
- (ii) $\|a\underline{x}\| = |a| \cdot \|\underline{x}\|$ for all $\underline{x} \in V$ and scalar a ;
- (iii) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$ for all $\underline{x}, \underline{y} \in V$.

Notice that these properties imply that

$$\|-\underline{x}\| = \|\underline{x}\|$$

and

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|. \quad (8.1)$$

The first relation is a consequence of property (ii) with the selection of $a = -1$. The second inequality can be proved as follows. From property (iii) we have

$$\|x\| = \|y + (x - y)\| \leq \|y\| + \|x - y\|,$$

which implies that

$$\|x\| - \|y\| \leq \|x - y\|. \quad (8.2)$$

By interchanging x and y , we see that

$$\|y\| - \|x\| \leq \|y - x\| = \|(x - y)\| = \|x - y\|. \quad (8.3)$$

If $\|x\| \geq \|y\|$, then from (8.2) we have

$$\|x\| - \|y\| = \|x\| - \|y\| \leq \|x - y\|,$$

and if $\|x\| < \|y\|$, then from (8.3) we obtain relation

$$\|x\| - \|y\| = \|y\| - \|x\| \leq \|x - y\|.$$

Hence inequality (8.1) is verified.

Three particular vector norms will be introduced in the next examples.

Example 8.1. Let x_1, \dots, x_n denote the elements of vector \underline{x} . Consider the vector norm

$$\|\underline{x}\|_{\infty} = \max_i \{|x_i|\},$$

which is called the "infinite-norm" of \underline{x} . We can easily verify that conditions (i), (ii), and (iii) are satisfied.

(i) $\|\underline{x}\|_{\infty} \geq 0$ obviously, and $\|\underline{x}\|_{\infty} = 0$ if and only if all $x_i = 0$ which means that $\underline{x} = \underline{0}$;

(ii) For all i , $|ax_i| = |a| \cdot |x_i|$, which implies the condition;

(iii) For all i , $|x_i + y_i| \leq |x_i| + |y_i|$.

Select i_0 such that

$$|x_{i_0} + y_{i_0}| = \max_i \{|x_i + y_i|\}.$$

Then

$$\begin{aligned} \|\underline{x} + \underline{y}\|_{\infty} &= |x_{i_0} + y_{i_0}| \leq |x_{i_0}| + |y_{i_0}| \leq \max_i \{|x_i|\} + \max_i \{|y_i|\} \\ &= \|\underline{x}\|_{\infty} + \|\underline{y}\|_{\infty}. \end{aligned}$$

For example, if $n = 4$ and $\underline{x} = (-1, 2, -4, 3)^T$, then $\|\underline{x}\|_{\infty} = 4$, since

$$4 = \max\{|-1|; |2|; |-4|; |3|\}.$$

Example 8.2. Consider next the "one-norm" of vectors defined as

$$\|\underline{x}\|_1 = \sum_{i=1}^n |x_i|.$$

This quantity also satisfies the conditions of Definition 8.1, so it can also be considered as a vector norm. Conditions (i), (ii) are trivially satisfied, and (iii) can be proven as follows. For all i ,

$$|x_i + y_i| \leq |x_i| + |y_i|.$$

The sum of this inequality for $i = 1, 2, \dots, n$ yields

$$\sum_i |x_i + y_i| \leq \sum_i |x_i| + \sum_i |y_i|$$

which is equivalent to condition (iii) applied to this norm. For example, if $n = 4$ and $\underline{x} = (-1, 2, -4, 3)^T$, then $\|\underline{x}\|_1 = 1 + 2 + 4 + 3 = 10$.

◆

Example 8.3. In section 2.4 we have introduced the length of n -element vectors, which can also be considered as a vector norm

$$\|\underline{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

This norm is known as the "two-norm" or "Euclidean-norm". Properties (i) and (ii) are obviously satisfied, and condition (iii) has been verified as the Corollary of Theorem 2.10. For example, if $n = 4$ and $\underline{x} = (-1, 2, -4, 3)^T$, then $\|\underline{x}\|_2 = \sqrt{1^2 + 2^2 + 4^2 + 3^2} = \sqrt{30}$.

◆

Notice that all these norms are special cases of the Hölder-norm:

$$\|\underline{x}\|_p = \left\{ \sum_{i=1}^n |x_i|^p \right\}^{1/p} \quad (8.4)$$

with some $p \geq 1$. A relation between the 1, 2, and ∞ -norms can be given as follows.

Theorem 8.1. For all $\underline{x} \in V$,

$$\|\underline{x}\|_2 \leq \|\underline{x}\|_1 \leq \sqrt{n} \|\underline{x}\|_2$$

$$\|\underline{x}\|_\infty \leq \|\underline{x}\|_2 \leq \sqrt{n} \|\underline{x}\|_\infty$$

$$\|\underline{x}\|_\infty \leq \|\underline{x}\|_1 \leq n \|\underline{x}\|_\infty.$$

Proof. The left-hand sides of these relations are obvious, since

$$\|\underline{x}\|_2^2 = \sum_{i=1}^n |x_i|^2 \leq \left(\sum_{i=1}^n |x_i| \right)^2 = \|\underline{x}\|_1^2,$$

$$\|\underline{x}\|_\infty^2 = \max_i \{ |x_i|^2 \} \leq \sum_{i=1}^n |x_i|^2 = \|\underline{x}\|_2^2,$$

and

$$\|\underline{x}\|_\infty = \max_i \{ |x_i| \} \leq \sum_{i=1}^n |x_i| = \|\underline{x}\|_1.$$

The right-hand sides can be shown as follows:

$$\|\underline{x}\|_1^2 = \left(\sum_{i=1}^n |x_i| \right)^2 = \left(\sum_{i=1}^n 1 \cdot |x_i| \right)^2 \leq \left(\sum_{i=1}^n 1^2 \right) \left(\sum_{i=1}^n |x_i|^2 \right) = n \cdot \|\underline{x}\|_2^2,$$

where we used the Cauchy-Schwarz inequality.

Similarly,

$$\|\underline{x}\|_2^2 = \sum_{i=1}^n |x_i|^2 \leq n \cdot \max_i \{ |x_i|^2 \} = n \cdot \|\underline{x}\|_\infty^2,$$

and

$$\|\underline{x}\|_1 = \left(\sum_{i=1}^n |x_i| \right) \leq n \cdot \max_i \{ |x_i| \} = n \cdot \|\underline{x}\|_\infty.$$

♣

Remark. It can be proved that all vector norms are equivalent to each other, that is, if $\|\dots\|$ and $\|\dots\|^*$ are two vector norms in V , then there exist constants K_1, K_2 such that for all $\underline{x} \in V$,

$$K_1 \|\underline{x}\| \leq \|\underline{x}\|^* \leq K_2 \|\underline{x}\|.$$

Let \underline{A} be a nonsingular $n \times n$ real or complex matrix, and assume that $\|\dots\|$ is a given vector norm. Consider the mapping: $\underline{x} \mapsto \|\underline{Ax}\|$ for all n -element vectors \underline{x} . We can easily verify that this mapping also satisfies the conditions of Definition 8.1, that is, it can also be considered as a vector norm:

(i) $\|\underline{Ax}\| \geq 0$, since \underline{Ax} is an n -element vector, and $\|\underline{Ax}\| = 0$ if and only if $\underline{Ax} = \underline{0}$, which holds if and only if $\underline{x} = \underline{0}$, since \underline{A} is nonsingular.

(ii) $\|a \underline{Ax}\| = |a| \cdot \|\underline{Ax}\|$, since $\|\dots\|$ is a vector norm.

$$(iii) \quad \|\underline{A}(\underline{x} + \underline{y})\| = \|\underline{A}\underline{x} + \underline{A}\underline{y}\| \leq \|\underline{A}\underline{x}\| + \|\underline{A}\underline{y}\|,$$

where we used the triangle inequality for norm $\|\cdot\|$.

Notice, that in the case of a singular matrix \underline{A} condition (i) might fail, since $\underline{A}\underline{x} = \underline{0}$ does not imply that $\underline{x} = \underline{0}$, therefore $\|\underline{A}\underline{x}\|$ has zero value for all nonzero vectors $\underline{x} \in N(\underline{A})$.

8.3 Matrix Norms

Consider now the set M of all $m \times n$ real (or complex) matrices.

Definition 8.2. A real-valued function $\underline{A} \mapsto \|\underline{A}\|$ defined on M is called a matrix norm, if it satisfies the following conditions:

(i) $\|\underline{A}\| \geq 0$ for all $\underline{A} \in M$, and $\|\underline{A}\| = 0$ if and only if $\underline{A} = \underline{0}$;

(ii) $\|a\underline{A}\| = |a| \cdot \|\underline{A}\|$ for all $\underline{A} \in M$ and scalar a ;

(iii) For all $\underline{A}, \underline{B} \in M$,

$$\|\underline{A} + \underline{B}\| \leq \|\underline{A}\| + \|\underline{B}\|.$$

A special class of matrix norms can be defined in the following way.

Definition 8.3. Let $\|\cdot\|$ be a given vector norm. For all $\underline{A} \in M$, define

$$\|\underline{A}\| = \max_{\underline{x} \neq \underline{0}} \frac{\|\underline{A}\underline{x}\|}{\|\underline{x}\|}, \quad (8.5)$$

which is called the matrix norm subordinate to the given vector norm.

In order to show that (8.5) defines a matrix norm, we have to verify that it satisfies the conditions of Definition 8.2:

(i) $\|A\| \geq 0$, since $\|Ax\| \geq 0$ and for $x \neq 0$, $\|x\| > 0$, furthermore $\|A\| = 0$ if and only if $\|Ax\| = 0$ for all $x \neq 0$, which holds if and only if A is the zero matrix.

(ii)

$$\begin{aligned} \|aA\| &= \max_{x \neq 0} \frac{\|aAx\|}{\|x\|} = \max_{x \neq 0} \frac{|a| \cdot \|Ax\|}{\|x\|} \\ &= |a| \cdot \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = |a| \cdot \|A\|; \end{aligned}$$

(iii) For all A and $B \in M$,

$$\begin{aligned} \|A+B\| &= \max_{x \neq 0} \frac{\|(A+B)x\|}{\|x\|} \leq \max_{x \neq 0} \frac{\|Ax\| + \|Bx\|}{\|x\|} \\ &\leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} + \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\|. \end{aligned}$$

Property (ii) implies that the matrix norm subordinate to a given vector norm can also be defined as

$$\|A\| = \max\{\|Ax\| \mid x \in V, \|x\| = 1\}.$$

In the next three examples the subordinate matrix norms to the 1, 2, and ∞ vector norms will be determined.

Example 8.4. Consider first the ∞ vector norm. Assume that

$$\|\underline{x}\|_{\infty} = \max_i \{x_i\} = 1.$$

Then

$$\|Ax\|_{\infty} = \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_i \sum_{j=1}^n |a_{ij}| \cdot |x_j| \leq \max_i \sum_{j=1}^n |a_{ij}|. \quad (8.6)$$

Therefore

$$\|A\| \leq \max_i \sum_{j=1}^n |a_{ij}|.$$

Assume that maximum is obtained for $i = i_0$ on the right-hand side.

Select vector \underline{x} such that

$$x_j = \begin{cases} 0 & \text{if } a_{i_0 j} = 0 \\ \overline{\frac{a_{i_0 j}}{|a_{i_0 j}|}} & \text{if } a_{i_0 j} \neq 0, \end{cases}$$

where overbar denotes complex conjugate. In the real case we select

$$x_j = \begin{cases} 0 & \text{if } a_{i_0 j} = 0 \\ 1 & \text{if } a_{i_0 j} > 0 \\ -1 & \text{if } a_{i_0 j} < 0. \end{cases}$$

Then

$$a_{i_0j}x_j = |a_{i_0j}|,$$

therefore in inequality (8.6) we have equality. Hence, the subordinate matrix norm to the ∞ vector norm is the following:

$$\|A\|_{\infty} = \max_i \sum_{j=1}^n |a_{ij}|. \quad (8.7)$$

♦

Example 8.5. Consider next the 1 vector norm. Assume that

$$\|x\|_1 = \sum_{j=1}^n |x_j| = 1.$$

Then

$$\begin{aligned} \|Ax\|_1 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| \cdot |x_j| \\ &\leq \sum_{j=1}^n \sum_{i=1}^m |a_{ij}| \cdot |x_j| = \sum_{j=1}^n \left(|x_j| \sum_{i=1}^m |a_{ij}| \right) \\ &\leq \sum_{j=1}^n |x_j| \cdot \left(\max_j \sum_{i=1}^m |a_{ij}| \right) = \max_j \sum_{i=1}^m |a_{ij}|. \end{aligned} \quad (8.8)$$

Therefore,

$$\|A\| \leq \max_j \sum_{i=1}^m |a_{ij}|.$$

Assume next that maximum occurs for $j = j_0$. Select vector $\underline{x} = \underline{e}_{j_0}$, that is,

$$x_j = \begin{cases} 1 & \text{if } j = j_0 \\ 0 & \text{otherwise.} \end{cases}$$

Then in inequality (8.8) we have equality, therefore

$$\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}| \quad (8.9)$$

Example 8.6. Consider next the 2 vector norm

$$\|\underline{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} = \sqrt{\overline{\underline{x}}^T \underline{x}},$$

where overbar denotes complex conjugate. Assume that $\|\underline{x}\|_2 = 1$. Then

$$\|\underline{Ax}\|_2^2 = \overline{\underline{x}}^T \overline{A}^T \underline{Ax} = \overline{\underline{x}}^T \cdot \underline{Hx},$$

where matrix \underline{H} is an $n \times n$ self-adjoint matrix. Theorem 7.2 implies that all eigenvalues of \underline{H} are real. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ denote the eigenvalues of \underline{H} . It is easy to see that \underline{H} is positive semidefinite, therefore $\lambda_1 \geq 0$. From section 7.4 we know that there is a unitary matrix \underline{U} such that

$$\underline{H} = \overline{\underline{U}}^T \underline{D} \underline{U},$$

where \underline{D} is a diagonal matrix with diagonal elements $\lambda_1, \dots, \lambda_n$. Then

$$\|\underline{Ax}\|_2^2 = \bar{\underline{x}}^T \underline{U}^T \underline{DUx} = \bar{\underline{z}}^T \underline{Dz} = \sum_{i=1}^n \lambda_i |z_i|^2 \leq \lambda_1 \sum_{i=1}^n |z_i|^2,$$

where $\underline{z} = \underline{Ux}$. Since \underline{U} is unitary, $\|\underline{z}\|_2 = \|\underline{Ux}\|_2 = \|\underline{x}\|_2 = 1$. Therefore

$$\|\underline{Ax}\|_2^2 \leq \lambda_1,$$

that is,

$$\|\underline{Ax}\|_2 \leq \sqrt{\lambda_1}.$$

Select \underline{x} as an eigenvector of \underline{H} associated to λ_1 such that $\|\underline{x}\|_2 = 1$.

Then

$$\|\underline{Ax}\|_2^2 = \bar{\underline{x}}^T \underline{Hx} = \bar{\underline{x}}^T \lambda_1 \underline{x} = \lambda_1 \|\underline{x}\|_2^2 = \lambda_1,$$

which shows that

$$\|\underline{A}\|_2 = \sqrt{\lambda_1}. \quad (8.10)$$

◆

In the following example we introduce a matrix norm, which differs from the ones shown in the previous examples. In many practical applications it is used instead of the 2 matrix norm.

Example 8.7. For an $m \times n$ real (or complex) matrix introduce the norm

$$\|A\|_F = \left\{ \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right\}^{1/2},$$

which is called the *Frobenius norm* of A . Notice that it is the 2 vector norm of the mn element vector

$$\underline{a} = (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, a_{m1}, \dots, a_{mn})^T,$$

therefore it satisfies the properties of Definition 8.2. We will next show that for all A ,

$$\|A\|_F \geq \|A\|_2.$$

Assume that $\|\underline{x}_2\| = 1$, then using the Cauchy-Schwarz inequality (Theorem 2.10) we have

$$\begin{aligned} \|A\underline{x}\|_2^2 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \leq \sum_{i=1}^m \left\{ \sum_{j=1}^n |a_{ij}|^2 \cdot \sum_{j=1}^n |x_j|^2 \right\} \\ &= \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right) \left(\sum_{j=1}^n |x_j|^2 \right) = \|A\|_F^2, \end{aligned}$$

since the second factor equals one. Therefore

$$\|A\|_2 = \max\{ \|A\underline{x}\|_2 \mid \|\underline{x}\|_2 = 1 \} \leq \|A\|_F.$$

A very useful property of the above introduced matrix norms is given next.

Theorem 8.2. For any real or complex matrices \underline{A} , \underline{B} such that $\underline{A} \cdot \underline{B}$ is defined,

$$\|\underline{A} \cdot \underline{B}\|_p \leq \|\underline{A}\|_p \cdot \|\underline{B}\|_p \quad (p=1,2,\infty,F).$$

Proof. Assume first that $p=1,2$, or ∞ . Then for any vector \underline{x} such that $\underline{B} \cdot \underline{x}$ exists and $\|\underline{x}\|_p = 1$,

$$\begin{aligned} \|(\underline{A} \cdot \underline{B})\underline{x}\|_p &= \|\underline{A} \cdot (\underline{B} \cdot \underline{x})\|_p \leq \|\underline{A}\|_p \cdot \|\underline{B} \cdot \underline{x}\|_p \\ &\leq \|\underline{A}\|_p \cdot \|\underline{B}\|_p \cdot \|\underline{x}\|_p = \|\underline{A}\|_p \cdot \|\underline{B}\|_p, \end{aligned}$$

which implies the assertion. If $p=F$, then assuming that \underline{A} is $m \times n$ and \underline{B} is $n \times r$,

$$\begin{aligned} \|\underline{A} \cdot \underline{B}\|_F^2 &= \sum_{i=1}^m \sum_{j=1}^r \left| \sum_{l=1}^n a_{il} b_{lj} \right|^2 \leq \sum_{i=1}^m \sum_{j=1}^r \left(\sum_{l=1}^n |a_{il}|^2 \cdot \sum_{l=1}^n |b_{lj}|^2 \right) \\ &= \left(\sum_{i=1}^m \sum_{l=1}^n |a_{il}|^2 \right) \left(\sum_{l=1}^n \sum_{j=1}^r |b_{lj}|^2 \right) = \|\underline{A}\|_F^2 \cdot \|\underline{B}\|_F^2. \end{aligned}$$

✦

Remark. The assertion of the theorem holds for any subordinate matrix norm.

Corollary. Let \underline{A} be an $n \times n$ real or complex matrix. Then for $p=1,2,\infty,F$ and $k=1,2,\dots$,

$$\|\underline{A}^k\|_p \leq (\|\underline{A}\|_p)^k. \quad (8.11)$$

Theorem 8.3. Let \underline{A} be any real or complex $n \times n$ matrix, and assume that λ is an eigenvalue of \underline{A} . Then

$$|\lambda| \leq \|\underline{A}\|_p \quad (p = 1, 2, \infty, F).$$

Proof. Let $p = 1, 2$, or ∞ , and assume that $\underline{x} \neq \underline{0}$ is an eigenvector of \underline{A} associated to λ . Then the eigenvector equation implies that

$$|\lambda| \cdot \|\underline{x}\|_p = \|\lambda \underline{x}\|_p = \|\underline{A} \underline{x}\|_p \leq \|\underline{A}\|_p \cdot \|\underline{x}\|_p.$$

Since $\underline{x} \neq \underline{0}$, $\|\underline{x}\|_p \neq 0$. Divide this inequality by $\|\underline{x}\|_p$ to obtain the assertion. For $p = F$, the assertion follows from the fact that $|\lambda| \leq \|\underline{A}\|_2 \leq \|\underline{A}\|_F$.

♣

Remark 1. This theorem provides simple bounds for the eigenvalues of real or complex matrices:

$$|\lambda| \leq \max_i \sum_{j=1}^n |a_{ij}|;$$

$$|\lambda| \leq \max_j \sum_{i=1}^n |a_{ij}|;$$

$$|\lambda| \leq \left\{ \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right\}^{1/2}.$$

Example 8.8. Consider the 3×3 real matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Notice that

$$\|\underline{A}\|_1 = \max\{1+1+1; 2+1+1; 1+1+1\} = 4;$$

$$\|\underline{A}\|_\infty = \max\{1+2+1; 1+1+1; 1+1+1\} = 4;$$

and

$$\|\underline{A}\|_F = \{1+4+1+1+1+1+1+1+1\}^{1/2} = \sqrt{12}.$$

Therefore for all eigenvalues λ of \underline{A} , $|\lambda| \leq \sqrt{12}$. The true eigenvalues are:

$$\lambda_{1,2} = \frac{3 \pm \sqrt{13}}{2}, \lambda_3 = 0.$$

◆

Remark 2. For nonnegative matrices, in Theorem 7.10 we presented lower and upper bounds for the largest eigenvalue. The upper bounds are the ∞ -norms of the matrices. Therefore the above theorem generalizes this upper bound for any $n \times n$ real or complex matrices.

A nice and practically important refinement of the previous theorem for the 1 and ∞ norms is presented in the following result, which is known as the *Gerschgorin-disk theorem*.

Theorem 8.4. Let \underline{A} be a real or complex $n \times n$ matrix, and

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad (i=1, 2, \dots, n)$$

For all i , define the disk

$$B_i = \{z \mid |z - a_{ii}| \leq r_i\}$$

on the complex plane. Then all eigenvalues of \underline{A} can be found in the union $B_1 \cup B_2 \cup \dots \cup B_n$ of these disks.

Proof. The eigenvector equation can be written as

$$\lambda x_i = \sum_{j=1}^n a_{ij} x_j \quad (i=1, 2, \dots, n), \quad (8.12)$$

where λ is an eigenvalue of \underline{A} and $\underline{x} = (x_i)$ is an associated eigenvector to λ . Let i_0 be selected so that

$$|x_{i_0}| = \max_i |x_i|,$$

then $x_{i_0} \neq 0$, and from equation (8.12) we have

$$\lambda = a_{i_0 i_0} + \sum_{\substack{j=1 \\ j \neq i_0}}^n a_{i_0 j} \frac{x_j}{x_{i_0}},$$

which implies that

$$|\lambda - a_{i_0 i_0}| \leq \sum_{\substack{j=1 \\ j \neq i_0}}^n |a_{i_0 j}|,$$

that is, $\lambda \in B_{i_0}$. Hence λ is in the union of all disks B_i .

♣

Corollary. Since the eigenvalues of \underline{A} and \underline{A}^T are the same, we can apply the theorem to \underline{A}^T to obtain new Gerschgorin disks. For $j = 1, 2, \dots, n$, let

$$s_j = \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ji}|,$$

and

$$D_j = \{z \mid |z - a_{jj}| \leq s_j\}$$

Then all eigenvalues of \underline{A} can be found in the union $D_1 \cup D_2 \cup \dots \cup D_n$ of these disks.

The application of the theorem and its corollary is illustrated in the next example.

Example 8.9. Consider again the 3 x 3 matrix

$$\underline{A} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

which was the subject of the previous example. In this case, $a_{11} = a_{22} = a_{33} = 1, r_1 = 3, r_2 = r_3 = 2, s_1 = s_3 = 2, s_2 = 3$.

Therefore

$$B_1 = D_2 = \{z \mid |z - 1| < 3\},$$

$$B_2 = B_3 = D_1 = D_3 = \{z \mid |z - 1| < 2\}.$$

The union of these disk is B_1 , which therefore contains all eigenvalues of \underline{A} . This domain and the true eigenvalues are shown in Figure 8.1.

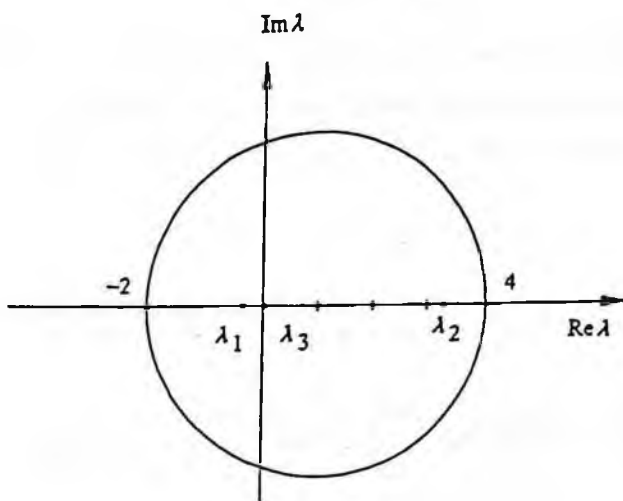


Figure 8.1. Disk containing all eigenvalues of \underline{A} .

8.4 Applications

In this section some useful applications of the properties of vector and matrix norms will be presented.

1. In our first application we derive a *bound for the inverses* of matrices of a special class. Assume that $\|\cdot\|$ is a subordinate matrix norm such that $\|\underline{I}\| = 1$, where \underline{I} is the identity matrix. Let \underline{A} be a real or complex $n \times n$ matrix, and assume that $\|\underline{A}\| < 1$. We will prove that $\underline{I} - \underline{A}$ is nonsingular, and

$$\|(\underline{I} - \underline{A})^{-1}\| \leq \frac{1}{1 - \|\underline{A}\|}. \quad (8.13)$$

Notice first that the assumption implies that for all eigenvalues λ_i of \underline{A} , $|\lambda_i| < 1$, therefore the eigenvalues $1 - \lambda_i$ of $\underline{I} - \underline{A}$ are all nonzero. Thus, $\underline{I} - \underline{A}$ is nonsingular. From Application 3 of section 6.8 we know that

$$(\underline{I} - \underline{A})^{-1} = \underline{I} + \underline{A} + \underline{A}^2 + \underline{A}^3 + \dots$$

Property (iii) of Definition 8.2 and Theorem 8.2 imply that

$$\|(\underline{I} - \underline{A})^{-1}\| \leq \|\underline{I}\| + \sum_{k=1}^{\infty} \|\underline{A}^k\| \leq 1 + \sum_{k=1}^{\infty} \|\underline{A}\|^k = \frac{1}{1 - \|\underline{A}\|}.$$

2. In this application the *relative errors of inverses* of approximating matrices will be examined. Assume that \underline{A} is nonsingular, and $\underline{A} - \underline{E}$ is a small perturbation of \underline{A} such that $\|\underline{A}^{-1}\underline{E}\| < 1$. We will now prove that $\underline{A} - \underline{E}$ is also nonsingular and

$$\frac{\|(\underline{A} - \underline{E})^{-1} - \underline{A}^{-1}\|}{\|\underline{A}^{-1}\|} \leq \frac{\|\underline{A}^{-1} \underline{E}\|}{1 - \|\underline{A}^{-1} \underline{E}\|}. \quad (8.14)$$

Notice first that

$$\underline{A} - \underline{E} = \underline{A}(\underline{I} - \underline{A}^{-1} \underline{E}).$$

The first factor is nonsingular by our assumption, and the second factor is nonsingular as the consequence of the previous application. Thus, $\underline{A} - \underline{E}$ is nonsingular. From the above equation we have

$$(\underline{A} - \underline{E})^{-1} = (\underline{I} - \underline{A}^{-1} \underline{E})^{-1} \underline{A}^{-1} = (\underline{I} + (\underline{A}^{-1} \underline{E}) + (\underline{A}^{-1} \underline{E})^2 + \dots) \underline{A}^{-1}.$$

That is,

$$(\underline{A} - \underline{E})^{-1} - \underline{A}^{-1} = \sum_{k=1}^{\infty} (\underline{A}^{-1} \underline{E})^k \underline{A}^{-1}.$$

Using property (iii) of Definition 8.2 and Theorem 8.2 we obtain the inequality

$$\begin{aligned} \|(\underline{A} - \underline{E})^{-1} - \underline{A}^{-1}\| &\leq \sum_{k=1}^{\infty} \|(\underline{A}^{-1} \underline{E})^k \underline{A}^{-1}\| \leq \left(\sum_{k=1}^{\infty} \|\underline{A}^{-1} \underline{E}\|^k \right) \|\underline{A}^{-1}\| \\ &= \frac{\|\underline{A}^{-1} \underline{E}\|}{1 - \|\underline{A}^{-1} \underline{E}\|} \|\underline{A}^{-1}\| \end{aligned}$$

which implies the assertion. Notice that inequality (8.14) gives an upper bound for the relative error of inverse of an approximating matrix, if we consider \underline{A} as the exact matrix and $\underline{A} - \underline{E}$ as its approximation.

3. Our next application is concerned with the solution of systems of *linear equations with approximating data*. Assume that linear system $\underline{A}\underline{x} = \underline{b}$ is to be solved. Let \underline{x} denote the solution. We discuss first the effect of the perturbation in \underline{b} . Assume that $\underline{b} \neq \underline{0}$ and it is approximated by $\underline{b} + \underline{e}$, where \underline{e} can be considered as the error term. Let \underline{x}^* denote the solution of the approximating system, and let $\underline{\varepsilon} = \underline{x}^* - \underline{x}$. Since \underline{x} solves the exact system,

$$\underline{A}\underline{x} = \underline{b},$$

and since \underline{x}^* is the solution of the approximating system,

$$\underline{A}(\underline{x} + \underline{\varepsilon}) = \underline{b} + \underline{e}.$$

Subtract the two equations to see that

$$\underline{A}\underline{\varepsilon} = \underline{e},$$

that is,

$$\underline{\varepsilon} = \underline{A}^{-1}\underline{e},$$

which implies that

$$\|\underline{\varepsilon}\| \leq \|\underline{A}^{-1}\| \cdot \|\underline{e}\|.$$

It is clear that

$$\| \underline{b} \| = \| \underline{A} \underline{x} \| \leq \| \underline{A} \| \cdot \| \underline{x} \|.$$

Combining the last two inequalities, we have

$$\frac{\| \underline{\epsilon} \|}{\| \underline{x} \|} \leq \frac{\| \underline{A}^{-1} \| \cdot \| \underline{e} \|}{\| \underline{b} \| / \| \underline{A} \|} = \| \underline{A} \| \cdot \| \underline{A}^{-1} \| \cdot \frac{\| \underline{e} \|}{\| \underline{b} \|}.$$

The quantity $\| \underline{A} \| \cdot \| \underline{A}^{-1} \|$ is called the *condition number* of matrix \underline{A} . This inequality shows that the relative change in the solution is bounded by the condition number times the relative change on the right-hand side vector. Therefore small condition number indicates that small relative change on the right-hand side vector implies only a small relative change in the solution. On the other hand, if the condition number is large, then even a small relative change on the right-hand side vector might imply a drastic relative change in the solution. The condition number is denoted by $\text{Cond}(\underline{A})$, and the above inequality can be restated as

$$\frac{\| \underline{\epsilon} \|}{\| \underline{x} \|} \leq \text{Cond}(\underline{A}) \cdot \frac{\| \underline{e} \|}{\| \underline{b} \|}. \quad (8.15)$$

Assume next that the coefficient matrix \underline{A} is approximated by $\underline{A} + \underline{E}$, where \underline{E} is a small matrix, but the right hand side remains the same. Let \underline{x}^* denote again the solution of the approximating equation and let $\underline{\epsilon} = \underline{x}^* - \underline{x}$. Subtract equations

$$(\underline{A} + \underline{E})(\underline{x} + \underline{\epsilon}) = \underline{b}$$

and

$$\underline{A}x = \underline{b}$$

to get

$$(\underline{A} + \underline{E})\underline{\varepsilon} = -\underline{E}x,$$

which implies that

$$\underline{\varepsilon} = -\underline{A}^{-1}\underline{E}(\underline{x} + \underline{\varepsilon}).$$

Therefore

$$\|\underline{\varepsilon}\| \leq \|\underline{A}^{-1}\| \cdot \|\underline{E}\|(\|\underline{x}\| + \|\underline{\varepsilon}\|) = \frac{\text{Cond}(\underline{A})}{\|\underline{A}\|} \|\underline{E}\|(\|\underline{x}\| + \|\underline{\varepsilon}\|).$$

Solving this inequality for $\|\underline{\varepsilon}\|$ we have

$$\|\underline{\varepsilon}\| \leq \frac{\text{Cond}(\underline{A}) \cdot \frac{\|\underline{E}\|}{\|\underline{A}\|} \|\underline{x}\|}{1 - \text{Cond}(\underline{A}) \cdot \frac{\|\underline{E}\|}{\|\underline{A}\|}},$$

where we have to assume that

$$\text{Cond}(\underline{A}) \frac{\|\underline{E}\|}{\|\underline{A}\|} = \|\underline{A}^{-1}\| \cdot \|\underline{E}\| < 1. \quad (8.16)$$

From the above inequality we conclude that

$$\frac{\|\underline{\epsilon}\|}{\|\underline{x}\|} \leq \frac{\text{Cond}(\underline{A}) \frac{\|\underline{E}\|}{\|\underline{A}\|}}{1 - \text{Cond}(\underline{A}) \frac{\|\underline{E}\|}{\|\underline{A}\|}}. \quad (8.17)$$

This inequality shows that if $\text{Cond}(\underline{A})$ is small, then small relative change in \underline{A} indicates only small relative change in the solution. If $\text{Cond}(\underline{A})$ is large, then \underline{E} has to be very small in order to guarantee that condition (8.16) holds. Even in this case, the relative change in \underline{x} might be very large.

Consider next the case, when both \underline{A} and \underline{b} are approximated. Similarly to the previous cases one can easily prove that

$$\frac{\|\underline{\epsilon}\|}{\|\underline{x}\|} \leq \frac{\text{Cond}(\underline{A}) \frac{\|\underline{E}\|}{\|\underline{A}\|}}{1 - \text{Cond}(\underline{A}) \frac{\|\underline{E}\|}{\|\underline{A}\|}} \left(\frac{\|\underline{E}\|}{\|\underline{A}\|} + \frac{\|\underline{e}\|}{\|\underline{b}\|} \right) \quad (8.18)$$

This inequality can be interpreted in the same way as it was done in the previous cases.

4. In section 3.8 we have discussed the *least squares method* to find the best polynomial fit for a given data set. In this application we will consider this problem under more general conditions. We will offer a solution, and two alternative formulations based on other than the 2-vector norm will be presented.

Assume that the value of a quantity y depends on variables x_1, x_2, \dots, x_n . For example, x_1, x_2, \dots, x_n may represent the energy usage,

manpower, technology level, etc. of a firm, with y being the output value of the firm in a certain time period. It is assumed that simultaneous values of these variables are measured, for example, at different time periods. Let the real values of the k^{th} ($k = 1, 2, \dots, N$) measurement be denoted as

$$x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}, y^{(k)}.$$

The most simple functional relation is linear, so for the sake of simplicity assume that we are looking for a function of the form

$$y = c_0 + c_1 x_1 + c_2 x_2 + \dots + c_n x_n,$$

where the coefficients c_0, c_1, \dots, c_n are unknown. We will determine these values based on the condition that they should provide best overall fit with respect to the measurement data. The function values give the vector

$$\underline{Y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{pmatrix}$$

where the k^{th} element is the k^{th} measurement for y . With fixed values of c_0, c_1, \dots, c_n , the linear function gives the vector

$$\begin{pmatrix} c_0 + c_1 x_1^{(1)} + \dots + c_n x_n^{(1)} \\ c_0 + c_1 x_1^{(2)} + \dots + c_n x_n^{(2)} \\ \dots \\ c_0 + c_1 x_1^{(N)} + \dots + c_n x_n^{(N)} \end{pmatrix}$$

where the k^{th} element shows what would be the function value at the k^{th} measurement if the relation were linear and c_0, c_1, \dots, c_n were the right coefficient values. This vector can be rewritten as $\underline{X}\underline{c}$, where

$$\underline{X} = \begin{pmatrix} 1 & x_1^{(1)} & \dots & x_n^{(1)} \\ 1 & x_1^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots \\ 1 & x_1^{(N)} & \dots & x_n^{(N)} \end{pmatrix} \quad \text{and} \quad \underline{c} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Hence, a logical way to select the values of c_0, c_1, \dots, c_n that minimize the overall discrepancy between vectors \underline{Y} and $\underline{X}\underline{c}$. This concept can be mathematically formulated as minimizing $\|\underline{X}\underline{c} - \underline{Y}\|$, where $\|\cdot\|$ is a given vector norm. Depending on the norm selection different approximation methods are obtained.

The selection of the 2-vector norm is known as the *least squares method*. The objective function can be rewritten as

$$\begin{aligned} \|\underline{X}\underline{c} - \underline{Y}\|_2 &= \sqrt{(\underline{X}\underline{c} - \underline{Y})^T (\underline{X}\underline{c} - \underline{Y})} \\ &= \sqrt{(\underline{c}^T \underline{X}^T - \underline{Y}^T) (\underline{X}\underline{c} - \underline{Y})} = \sqrt{\underline{c}^T \underline{X}^T \underline{X} \underline{c} - 2\underline{c}^T \underline{X}^T \underline{Y} + \underline{Y}^T \underline{Y}}. \end{aligned}$$

This norm is minimal if and only if the expression under the square root is minimal. Consider therefore the $(n+1)$ -variable function

$$g(\underline{c}) = \underline{c}^T \underline{X}^T \underline{X} \underline{c} - 2\underline{c}^T \underline{X}^T \underline{Y} + \underline{Y}^T \underline{Y}.$$

The gradient vector of g is the following:

$$\nabla g(\underline{c}) = 2\underline{X}^T \underline{X} \underline{c} - 2\underline{X}^T \underline{Y},$$

and the Hessian matrix is

$$H(\underline{c}) = 2\underline{X}^T \underline{X}.$$

Since $\underline{X}^T \underline{X}$ is positive semidefinite, function g is convex. Therefore the stationary points provide global optimum. Equating the gradient to zero leads to the so-called *normal equations*.

$$\underline{X}^T \underline{X} \underline{c} = \underline{X}^T \underline{Y}. \quad (8.19)$$

Notice that this is a system of linear equations with coefficient matrix $\underline{X}^T \underline{X}$ and right-hand side vector $\underline{X}^T \underline{Y}$.

In the special case of $n=1$, the above problem reduces to *linear regression*. In this case,

$$\underline{X} = \begin{pmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \dots & \dots \\ 1 & x^{(N)} \end{pmatrix} \quad \text{and} \quad \underline{Y} = \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(N)} \end{pmatrix},$$

therefore the normal equations can be written as

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x^{(1)} & x^{(2)} & \dots & x^{(N)} \end{pmatrix} \begin{pmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \dots & \dots \\ 1 & x^{(N)} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x^{(1)} & x^{(2)} & \dots & x^{(N)} \end{pmatrix} \begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(N)} \end{pmatrix}$$

Simple calculation shows that this equation is equivalent to the system

$$\begin{aligned} Nc_0 + \left(\sum_{k=1}^N x^{(k)} \right) c_1 &= \sum_{k=1}^N y^{(k)} \\ \left(\sum_{k=1}^N x^{(k)} \right) c_0 + \left(\sum_{k=1}^N x^{(k)^2} \right) c_1 &= \sum_{k=1}^N x^{(k)} y^{(k)}. \end{aligned}$$

In order to simplify these equations, introduce the following notation:

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x^{(k)}, \quad \bar{y} = \frac{1}{N} \sum_{k=1}^N y^{(k)}, \quad \overline{x^2} = \frac{1}{N} \sum_{k=1}^N x^{(k)^2},$$

and

$$\overline{xy} = \frac{1}{N} \sum_{k=1}^N x^{(k)} y^{(k)}.$$

Notice that these quantities are the averages of the values of $x^{(k)}$, $y^{(k)}$, $x^{(k)^2}$, and the products $x^{(k)} y^{(k)}$, respectively. Dividing both equations by N we have

$$\begin{aligned}c_0 + \bar{x}c_1 &= \bar{y} \\ \bar{x}c_0 + \bar{x}^2c_1 &= \bar{xy}.\end{aligned}$$

Subtracting the \bar{x} -multiple of the first equation from the second equation eliminates c_0 :

$$c_1(\bar{x}^2 - \bar{x}^2) = \bar{xy} - \bar{x} \cdot \bar{y},$$

which implies that

$$c_1 = \frac{\bar{xy} - \bar{x} \cdot \bar{y}}{\bar{x}^2 - \bar{x}^2}. \quad (8.20)$$

From the first equation we conclude that

$$c_0 = \bar{y} - \bar{x} \cdot c_1. \quad (8.21)$$

The least squares method is illustrated in the next examples.

Example 8.10. Consider the following data values: $N = 5$,

$$\underline{X} = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \quad \text{and} \quad \underline{Y} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}$$

That is, the measurements are as follows:

$$\begin{aligned}x_1^{(1)} &= -2, & x_2^{(1)} &= 4, & y^{(1)} &= 0, \\x_1^{(2)} &= -1, & x_2^{(2)} &= 1, & y^{(2)} &= 1, \\x_1^{(3)} &= 0, & x_2^{(3)} &= 0, & y^{(3)} &= 2, \\x_1^{(4)} &= 1, & x_2^{(4)} &= 1, & y^{(4)} &= 1, \\x_1^{(5)} &= 2, & x_2^{(5)} &= 4, & y^{(5)} &= 0.\end{aligned}$$

Simple calculation shows that

$$\underline{X}^T \underline{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}$$

$$\underline{X}^T \underline{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$$

Hence the normal equations can be written as

$$\begin{pmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$$

The application of the elimination method gives the solution

$$c_0 = \frac{58}{35}, c_1 = 0, \quad \text{and} \quad c_2 = -\frac{3}{7}.$$

Hence, the least squares function is:

$$y = \frac{58}{35} - \frac{3}{7}x_2.$$

Example 8.11. Consider next a simple regression problem. Assume that a function $y = c_0 + c_1x$ is to be determined with data: $N = 4$,

$$\begin{aligned} x^{(1)} &= 0 & y^{(1)} &= 0 \\ x^{(2)} &= 1 & y^{(2)} &= 0 \\ x^{(3)} &= 2 & y^{(3)} &= 1 \\ x^{(4)} &= 3 & y^{(4)} &= 1. \end{aligned}$$

In this case,

$$\bar{x} = \frac{3}{2}, \bar{x^2} = \frac{7}{2}, \bar{y} = \frac{1}{2}, \quad \text{and} \quad \overline{xy} = \frac{5}{4},$$

therefore equations (8.20) and (8.21) imply that

$$c_1 = \frac{\frac{5}{4} - \frac{3}{2} \cdot \frac{1}{2}}{\frac{7}{2} - \frac{9}{4}} = \frac{2}{5},$$

and

$$c_0 = \frac{1}{2} - \frac{2}{5} \cdot \frac{3}{2} = -\frac{1}{10}.$$

Hence, the regression line is the following:

$$y = -\frac{1}{10} + \frac{2}{5}x.$$

The original data points and the regression line are illustrated in Figure 8.2.

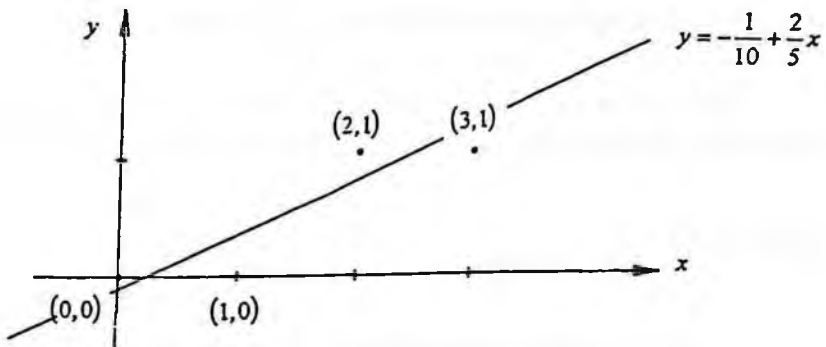


Figure 8.2 Illustration of a Regression Line.

An alternative approach to least squares is derived if we select the ∞ vector norm. In this case the maximal discrepancy between the components of vectors \underline{Xc} and \underline{Y} is minimized. Mathematically this concept can be formulated as the unconstrained optimization problem:

$$\text{Minimize } \max_k \{ |c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} - y^{(k)}| \}. \quad (8.22)$$

If E denotes the objective function, then this nonlinear optimization problem can be rewritten as a linear programming problem:

$$\text{Minimize } E \quad (8.23)$$

$$\text{subject to } \left. \begin{array}{l} c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} + E \geq y^{(k)} \\ c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} - E \leq y^{(k)} \end{array} \right\} (1 \leq k \leq N),$$

where the unknowns are c_0, c_1, \dots, c_n , and E . The solution of this linear programming problem provides the best fitting linear function.

Example 8.12. Consider again the problem of Example 8.10. Based on the data values, the linear programming formulation has now the form:

minimize E

$$\begin{aligned}
 & c_0 - 2c_1 + 4c_2 + E \geq 0 \\
 & c_0 - 2c_1 + 4c_2 - E \leq 0 \\
 & c_0 - c_1 + c_2 + E \geq 1 \\
 & c_0 - c_1 + c_2 - E \leq 1 \\
 \text{subject to } & c_0 + E \geq 2 \\
 & c_0 - E \leq 2 \\
 & c_0 + c_1 + c_2 + E \geq 1 \\
 & c_0 + c_1 + c_2 - E \leq 1 \\
 & c_0 + 2c_1 + 4c_2 + E \geq 0 \\
 & c_0 + 2c_1 + 4c_2 - E \leq 0.
 \end{aligned}$$

Another alternative approach to least squares is derived if we select the 1-vector norm. In this case the sum of the absolute discrepancies between the components of vectors \underline{X}_c and \underline{Y} is minimized. This concept can be mathematically formulated as the unconstrained optimization problem

$$\text{minimize} \quad \sum_{k=1}^N |c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} - y^{(k)}|, \quad (8.24)$$

which can be rewritten as the following linear programming problem:

$$\begin{aligned}
 & \text{minimize} \quad \sum_{k=1}^N E_k \quad (8.25) \\
 & \text{subject to} \quad \left. \begin{aligned} & c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} + E_k \geq y^{(k)} \\ & c_0 + c_1 x_1^{(k)} + \dots + c_n x_n^{(k)} - E_k \leq y^{(k)} \end{aligned} \right\} (1 \leq k \leq N),
 \end{aligned}$$

where E_k denotes the k^{th} term of the objective function (8.24).

Example 8.13. In the case of the data values of the previous example, problem (8.25) has the form:

$$\begin{array}{ll}
 \text{minimize} & \sum_{k=1}^5 E_k \\
 & c_0 - 2c_1 + 4c_2 + E_1 \geq 0 \\
 & c_0 - 2c_1 + 4c_2 - E_1 \leq 0 \\
 & c_0 - c_1 + c_2 + E_2 \geq 1 \\
 & c_0 - c_1 + c_2 - E_2 \leq 1 \\
 \text{subject to} & c_0 + E_3 \geq 2 \\
 & c_0 - E_3 \leq 2 \\
 & c_0 + c_1 + c_2 + E_4 \geq 1 \\
 & c_0 + c_1 + c_2 - E_4 \leq 1 \\
 & c_0 + 2c_1 + 4c_2 + E_5 \geq 0 \\
 & c_0 + 2c_1 + 4c_2 - E_5 \leq 0.
 \end{array}$$

In many applications in economics, for example, in computing production functions, linear relations are usually replaced by function forms as

$$y = a_0 x_1^{a_1} x_2^{a_2} \dots x_n^{a_n},$$

where the constants $a_0, a_1, a_2, \dots, a_n$ are the unknowns. The most commonly used technique to find the values of these unknowns is based on the following linearization idea. Take the logarithms of both sides to have

$$\log y = \log a_0 + a_1 \log x_1 + a_2 \log x_2 + \dots + a_n \log x_n.$$

That is, $\log y$ is a linear function of $\log x_1, \dots, \log x_n$. In the first step of the procedure, the logarithms of all data values are taken, and in the second step, the best linear fit is determined based on the new data set consisting of the logarithm values.

8.5 Exercises

1. Prove that if $\underline{x}, \underline{y} \in R^n$ are orthogonal, then

$$\|\underline{x} + \underline{y}\|_2^2 = \|\underline{x}\|_2^2 + \|\underline{y}\|_2^2.$$

2. Let $\underline{A} = (a_{ij})$ be an $m \times n$ real matrix. Define

$$\|\underline{A}\| = \max\{|a_{ij}| \mid 1 \leq i \leq m, 1 \leq j \leq n\}.$$

Is this quantity a matrix norm?

3. Let

$$\underline{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Give the particular form of the induced vector norms $\underline{x} \mapsto \|\underline{Ax}\|_p$

for $p = 1, 2, \infty$.

4. Give an $n \times n$ real matrix \underline{A} such that $\|\underline{A}\|_1 < 1$ but $\|\underline{A}\|_2 > 1$ and $\|\underline{A}\|_\infty > 1$.

5. Give an $n \times n$ real matrix \underline{A} such that $\|\underline{A}\|_2 < 1$ but $\|\underline{A}\|_1 > 1$ and $\|\underline{A}\|_\infty > 1$.

6. Give an $n \times n$ real matrix \underline{A} such that $\|\underline{A}\|_\infty < 1$ but $\|\underline{A}\|_1 > 1$ and $\|\underline{A}\|_2 > 1$.

7. Repeat Example 8.8 for matrix

$$\underline{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

8. Repeat Example 8.9 for the matrix of the previous problem.

9. Select matrix

$$\underline{A} = \begin{pmatrix} 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 \end{pmatrix}$$

and compare the right and left hand sides of inequality (8.13) with the ∞ matrix norm.

10. Repeat the previous problem with the 1 matrix norm.

11. Select $\underline{A} = \varepsilon \cdot \underline{I}$, where $0 < \varepsilon < 1$ is a given constant. Show that inequality (8.13) becomes an equality for any subordinate matrix norm.

12. Select

$$\underline{A} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{and} \quad \underline{E} = \begin{pmatrix} 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 \end{pmatrix}$$

and compare the right and left hand sides of inequality (8.14) with the ∞ matrix norm.

13. Repeat the previous problem with the 1-matrix norm.

14. Select $\underline{A} = \underline{I}$ and $\underline{E} = \varepsilon \cdot \underline{I}$ with $0 < \varepsilon < 1$. Show that in this special case inequality (8.14) becomes an equation.

15. Show that in the special case of $\underline{A} = \underline{I}$, inequality (8.15) becomes an equality.

16. Illustrate inequality (8.17) in the special case of $\underline{A} = \underline{I}$ and $\underline{E} = \varepsilon \cdot \underline{I}$ (with $0 < \varepsilon < 1$).

17. Prove that for any $n \times n$ real matrix \underline{A} , $\|\underline{A}^T \underline{A}\|_2 = \|\underline{A}\|_2^2$.

18. Let \underline{A} be an $n \times n$ real matrix with columns $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$. Prove that

$$\|\underline{A}\|_F^2 = \sum_{i=1}^n \|\underline{a}_i\|_2^2.$$

19. Let \underline{A} be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Assume that the eigenvalues are ordered so that $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$. Prove that

$$\|\underline{A}\|_2 = |\lambda_1|.$$

20. Repeat Example 8.10 with data values

$$\underline{X} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix} \quad \text{and} \quad \underline{Y} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 16 \end{pmatrix}$$

21. Repeat Example 8.11 with data values

$$\begin{aligned} x^{(1)} &= -2 & y^{(1)} &= 1 \\ x^{(2)} &= -1 & y^{(2)} &= 2 \\ x^{(3)} &= 0 & y^{(3)} &= 4 \\ x^{(4)} &= 1 & y^{(4)} &= 5 \\ x^{(5)} &= 2 & y^{(5)} &= 7. \end{aligned}$$

22. Repeat Example 8.12 with data values of Problem 20.

23. Repeat Example 8.13 with data values of Problem 20.

24. Let \underline{B} be a given $n \times n$ real nonsingular matrix. Consider mapping $\underline{A} \mapsto \|\underline{BA}\|$ for $n \times n$ real matrices \underline{A} with some matrix norm. Is this mapping a norm of \underline{A} ?

25. Let \underline{B} and \underline{C} be given $n \times n$ real nonsingular matrices. Consider mapping $\underline{A} \mapsto \|\underline{BAC}\|$ for $n \times n$ real matrices \underline{A} with some matrix norm. Is this mapping a norm of \underline{A} ?

References

- Argyros, I. K. and F. Szidarovszky (1993) The Theory and Applications of Iteration Methods. CRC Press, Boca Raton/London/Tokyo.
- Baumol, W. J. (1970) Economic Dynamics. (3rd edition) Macmillan, New York.
- Bellman, R. E. (1970) Introduction to Matrix Analysis. (3rd edition) McGraw-Hill Publ. Co., New York.
- Friedman, J. W. (1977) The Theory of Games and Oligopoly. North Holland, Amsterdam.
- Gandolfo, G. (1971) Mathematical Methods and Models in Economic Dynamics. North Holland, Amsterdam.
- Gantmacher, F. R. (1959) The Theory of Matrices. (Vols. 1 and 2) Chelsea, New York.
- Goldberg, S. (1958) Introduction to Difference Equations. John Wiley & Sons, New York.
- Golub, G. and C. Van Loan (1983) Matrix Computations. The John Hopkins Univ. Press, Baltimore, Md.
- Halmos, P. R. (1958) Finite Dimensional Vector Spaces. Van Nostrand, Princeton, N. J.
- Henderson, J. M. and R. E. Quandt (1971) Microeconomic Theory: A Mathematical Approach. (2nd edition) McGraw-Hill, New York.
- Herstein, I. (1964) Topics in Algebra. Blaisdell Publ. Co., New York.
- Kaplan, W. (1952) Advanced Calculus. Addison-Wesley Publ. Co., Reading, Mass.
- Lancaster, P. (1969) Theory of Matrices. Academic Press, New York.
- Leontief, W. (1973) Input-Output Economics. Oxford Univ. Press, New York.
- Liu, D. and F. Szidarovszky (1990) Block-M-Matrices and Their Properties. Pure Math. and Appl., Ser. B, Vol. 1, No. 2-3, pp. 99-107.
- Luenberger, D. G. (1979) Introduction to Dynamic Systems: Theory, Models, and Applications. John Wiley & Sons, New York.
- Molnar, S. (1990) A Special Decomposition of Linear Systems, Belgian Journal of Operations Research, Statistics, and Computer Science, Vol29, No4, pp. 4-37
- Molnar S., Szidarovszky F. (1992) Some notes on discrete Cournot-oligopolies with sequential adjustments, Pure Math and Applications, Ser. B, Vol.3, No 3-4, pp 289-293
- Nikaido, H. (1968) Convex Structures and Economic Theory. Academic Press, New York/London.

- Noble, B. (1968) Applied Linear Algebra. Prentice-Hall, Englewood Cliffs, N. J.
- Rugh, W.J. (1996) Linear System Theory. (2nd edition) Prentice-Hall, Upper Saddle River, N.J.
- Okuguchi, K. (1976) Expectations and Stability in Oligopoly Models. Springer-Verlag, Berlin/Heidelberg/New York.
- Okuguchi, K. and F. Szidarovszky (1990) The Theory of Oligopoly with Multi-Product Firms. Springer-Verlag, Berlin/Heidelberg/New York.
- Ortega, J. M. and W. C. Rheinboldt (1970) Iterative Solutions of Nonlinear Equations in Several Variables. Academic Press, New York.
- Ross, Sh. M. (1987) Introduction to Probability and Statistics for Engineers and Scientists. J. Wiley & Sons, New York/Toronto.
- Stewart, G. (1973) Introduction to Matrix Computations. Academic Press, New York.
- Strang, G. (1976) Linear Algebra and Its Applications. Academic Press, New York.
- Szidarovszky, F. and S. Molnár (1995) A Note on Extrapolative Expectations in a Dynamic Producer-Consumer Market. Keio Econ. Studies, Vol. XXXII, No. 2, pp. 71-73.
- Szidarovszky, F. and S. Yakowitz (1978) Principles and Procedures of Numerical Analysis. Plenum Press, New York/London.
- Szidarovszky, F. and A. T. Bahill (1992) Linear Systems Theory. CRC Press, Boca Raton/London/Tokyo.
- Szidarovszky, F. and S. Molnár (1986) Game Theory with Engineering Applications (in Hungarian) Műszaki Könyvkiadó, Budapest.
- Szidarovszky, F. and S. Molnár (1994) On Discrete Dynamic Producer-Consumer Markets. Keio Econ. Studies, Vol. XXXI, No. 2, pp. 51-63.
- Szidarovszky, F. and S. Molnár (1994) Learning in a Dynamic Producer-Consumer Market. Appl. Math. and Comp., Vol. 62, pp. 223-233.
- Szidarovszky, F. (1989) On Non-Negative Solvability of Nonlinear Input-Output Systems. Econ. Letters, Vol. 30, pp. 319-321.
- Szidarovszky, F. and K. Okuguchi (1989) A Non-Differentiable Input-Output Model. Math. Social Sci., Vol. 18, pp. 187-190.

Index

A

- Addition
 - of matrices, 9
- Adjoint
 - of a matrix, 397
- Augmented matrix, 139

B

- Basis, 80
 - cyclic, 352
 - orthonormal, 96
 - standard, 80
 - transformation equations, 295
- Block matrix, 24

C

- Canonical form
 - Jordan-, 359
- Cauchy-Schwarz inequality, 92
- Cayley-Hamilton theorem, 351
- Characteristic
 - equation, 318
 - polynomial, 318
- Cofactor, 225
- Complex
 - conjugate, 398

inner product space, 104

Condition number, 477

Conjugate

- of a matrix, 398

Cramer's rule, 233

D

- Determinant
 - cofactors, 225
 - multiplication theorem, 224
 - of matrix, 208, 212
 - Diagonable matrix, 342
 - Diagonal
 - block matrix, 383
 - elements, 4
 - matrix, 342
 - Directed edges, 50
 - Direct sum, 110, 112
 - Directing space, 149
 - Dyads, 15
 - Dynamic Cournot-oligopoly, 43
 - Dynamic economic systems, 40
 - Dynamic producer-consumer
 - models, 45
- ## E
- Eigenvalue, 316

Perron~, 422
 Eigenvector, 316
 Elimination method, 165
 Equivalence class, 270
 Euclidean space, 105

F

Fredholm integral equation, 305
 Fundamental theorem of algebra, 326

G

Generating system, 78
 Gerschgorin-disk theorem, 470
 Gram-Schmidt process, 99, 100
 Graphs, 50

H

Hamilton-Cayley Theorem, 617
 Hermitian matrix, 617
 Hicksian matrix, 415
 Homogeneous equation, 143
 Hyperplane, 153

I

Idempotent transformation, 336
 Identity
 matrix, 5, 19
 Image, 254
 Index
 of a permutation, 212
 Inequality

Cauchy-Schwarz, 92
 variational, 127
 Inhomogenous equation, 134
 Inner product, 85
 space, 89
 standard, 85
 Invariant subspace, 313

Inverse,

 closed form representation, 238
 of linear transformation, 279
 of matrix, 21, 160, 185

Invertible linear transformation, 279

Invertible matrix, 21

Isomorphic vector spaces, 257

Isomorphism, 257

J

Jordan

 canonical form, 359
 canonical matrix, 359

K

Kalman matrix, 241

L

Least squares method, 193, 479
 Leontief input-output system, 188
 Linear
 combination, 65
 dependence, 70
 system of \sim equations, 134

functional, 263, 292
 mapping, 255
 manifold, 149
 transformation, 276
 Linearly
 independent, 70
 Lower triangular matrix, 5

M

Mapping, 255
 nullspace, 265
 Markov-chain, 197
 Matrix, 2
 adjoint, 397
 conjugate, 398
 decomposable, 6
 diagonal, 383
 equation, 157
 Hicksian, 415
 idempotent, 393
 identity, 5, 19
 inverse, 21, 160, 185
 invertible, 22
 Jordan, 359
 lower triangular, 5, 392
 Metzler, 429
 minimal polynomial, 334
 nilpotent, 347, 395
 degree of \sim , 395
 nonnegative, 418
 norm, 461
 normal, 399
 of linear mappings, 284
 orthogonal, 399

polynomial, 20, 325
 positive definite, 399
 positive semidefinite, 399
 pseudoinverse, 444
 quasidefinite, 416
 similar matrices, 298
 symmetric, 7, 399
 transformation, 295
 transition, 197
 tridiagonal, 372
 upper triangular, 5, 392

Minimal polynomial, 334

Minor, 235

Multiplication

of a scalar and a matrix, 8
 of a scalar and a vector, 9, 10
 of linear transformations, 276
 of matrices, 14, 17
 of vectors, 15

N

Negative

definite quadratic form, 399
 semidefinite quadratic form,
 399

Nilpotent matrix, 347

Nonsingular

linear transformation, 261,
 262, 294

Norm, 455

Euclidean \sim , 458

Hölder \sim , 459

Frobenius \sim , 466

Nullspace, 265

O

Orthogonal

system, 91

Orthonormal set of vectors, 91

P

Permutation, 212

Perron-Frobenius theorem, 418,
419

Polynomial

interpolation, 191

Positive

definite quadratic form, 399

semidefinite quadratic form,
399

Principal

minor, 413

submatrix, 35

Products

of a scalar and a matrix, 7

of matrices, 14

Q

Quadratic form, 399

Quadratic matrix, 399

R

Range of mapping, 254

Rank of matrix, 139, 155

S

Self-adjoint, 399, 411

Sign of permutation, 212

Similar matrices, 297

Singular value decomposition,
437

Skew-symmetric, 7

Space

complex inner product, 104
dual, 292

Euclidean, 105

Square matrix, 4

Subspace, 63

orthogonal complementary,
114

cyclic, 354

Sum

direct, 110

of matrices, 9

Symmetric matrix, 7

T

Transpose

of a matrix, 3

Tridiagonal matrix, 383

Trigonometric system, 91

U

Unitary space, 105

Upper triangular matrix, 5

V

Vector

norm, 455

space, 57

unit~, 74

Vertices, 50

Volterra integral equation, 306

Series on Concrete and Applicable Mathematics – Vol. 3

Introduction to Matrix Theory

In economic modeling and planning, as well as in business, most problems are linear, or approximated by linear models. Such problems are solved by matrix methods, so the material presented in this book is essential to these fields.

