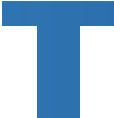
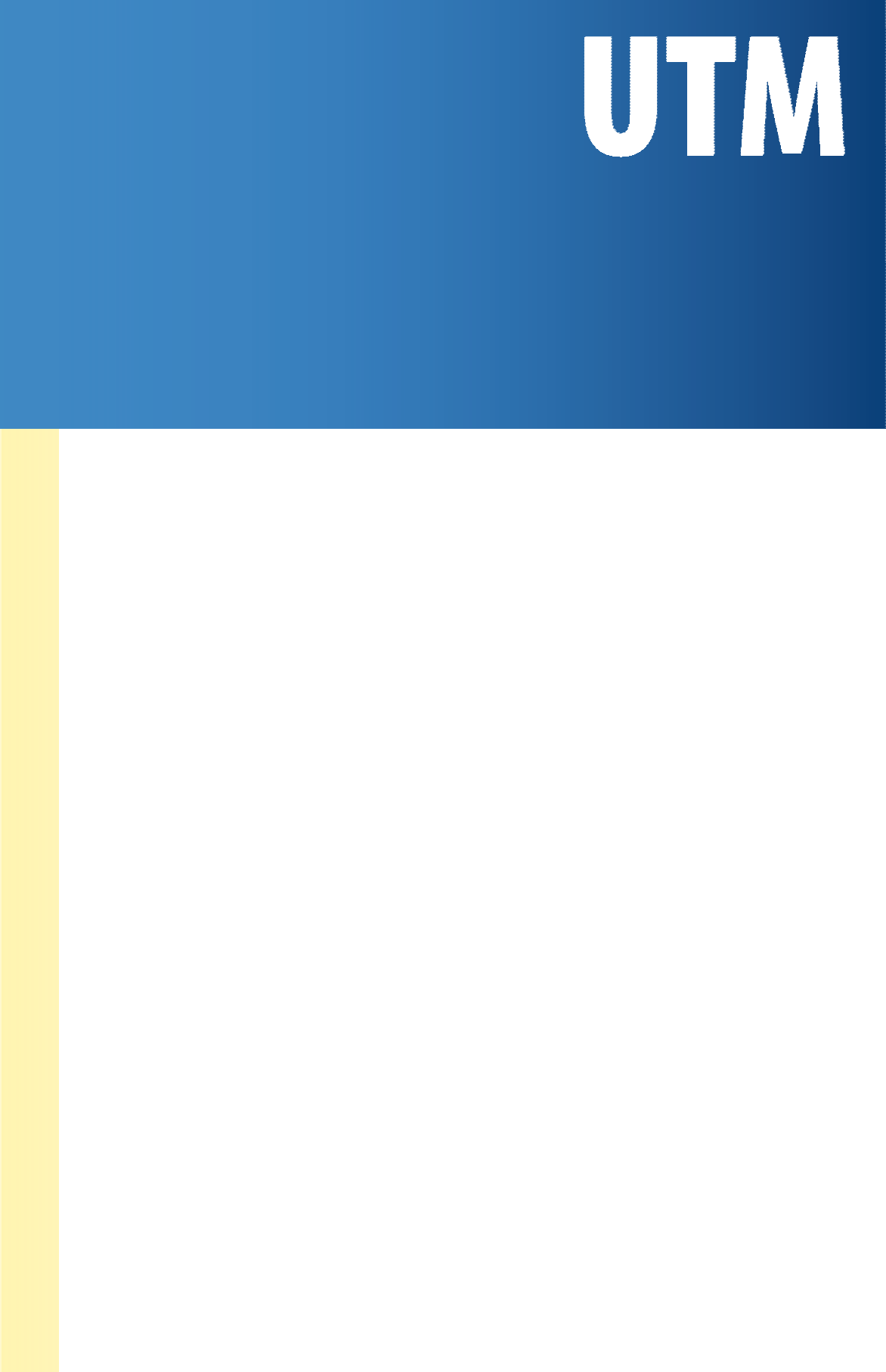
**Undergraduate Texts in Mathematics**



Sheldon Axler

Linear Algebra Done Right

*Third Edition*

Undergraduate Texts in Mathematics

Undergraduate Texts in Mathematics

**Series Editors:**

Sheldon Axler

*San Francisco State University, San Francisco, CA, USA*

Kenneth Ribet

*University of California, Berkeley, CA, USA*

**Advisory Board:**

Colin Adams, *Williams College, Williamstown, MA, USA*

Alejandro Adem, *University of British Columbia, Vancouver, BC, Canada*

Ruth Charney, *Brandeis University, Waltham, MA, USA*

Irene M. Gamba, *The University of Texas at Austin, Austin, TX, USA*

Roger E. Howe, *Yale University, New Haven, CT, USA*

David Jerison, *Massachusetts Institute of Technology, Cambridge, MA, USA*

Jeffrey C. Lagarias, *University of Michigan, Ann Arbor, MI, USA*

Jill Pipher, *Brown University, Providence, RI, USA*

Fadil Santosa, *University of Minnesota, Minneapolis, MN, USA*

Amie Wilkinson, *University of Chicago, Chicago, IL, USA*

**Undergraduate Texts in Mathematics** are generally aimed at third- and fourth- year undergraduate mathematics students at North American universities. These texts strive to provide students and teachers with new perspectives and novel approaches. The books include motivation that guides the reader to an appreciation of interrelations among different aspects of the subject. They feature examples that illustrate key concepts as well as exercises that strengthen understanding.

For further volumes: <http://www.springer.com/series/666>

Sheldon Axler

Linear Algebra Done Right

Third edition

1 3

Sheldon Axler

Department of Mathematics San Francisco State University San Francisco, CA, USA

ISSN 0172-6056 ISSN 2197-5604 (electronic)

ISBN 978-3-319-11079-0 ISBN 978-3-319-11080-6 (eBook) DOI 10.1007/978-3-319-11080-6

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014954079

Mathematics Subject Classification (2010): 15-01, 15A03, 15A04, 15A15, 15A18, 15A21 c Springer International Publishing 2015

*Q*

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part

of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher’s location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this

publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typeset by the author in LaTeX

*Cover figure*: For a statement of Apollonius’s Identity and its connection to linear algebra, see the last exercise in Section 6.A.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

***Preface for the Instructor* xi *Preface for the Student* xv *Acknowledgments* xvii**

*Contents*

1. ***Vector Spaces* 1**
   1. **R**n and **C**n **2**

Complex Numbers **2**

Lists **5**

**F**n **6**

Digression on Fields **10**

Exercises 1.A **11**

* 1. Deﬁnition of Vector Space **12**

Exercises 1.B **17**

* 1. Subspaces **18**

Sums of Subspaces **20**

Direct Sums **21**

Exercises 1.C **24**

1. ***Finite-Dimensional Vector Spaces* 27**
   1. Span and Linear Independence **28** Linear Combinations and Span **28** Linear Independence **32** Exercises 2.A **37**

v

**vi** Contents

* 1. Bases **39**

Exercises 2.B **43**

* 1. Dimension **44**

Exercises 2.C **48**

1. ***Linear Maps* 51**
   1. The Vector Space of Linear Maps **52**

Deﬁnition and Examples of Linear Maps **52** Algebraic Operations on *L*.V; W / **55** Exercises 3.A **57**

* 1. Null Spaces and Ranges **59** Null Space and Injectivity **59** Range and Surjectivity **61**

Fundamental Theorem of Linear Maps **63**

Exercises 3.B **67**

* 1. Matrices **70**

Representing a Linear Map by a Matrix **70** Addition and Scalar Multiplication of Matrices **72** Matrix Multiplication **74**

Exercises 3.C **78**

* 1. Invertibility and Isomorphic Vector Spaces **80**

Invertible Linear Maps **80**

Isomorphic Vector Spaces **82**

Linear Maps Thought of as Matrix Multiplication **84**

Operators **86**

Exercises 3.D **88**

* 1. Products and Quotients of Vector Spaces **91**

Products of Vector Spaces **91** Products and Direct Sums **93** Quotients of Vector Spaces **94** Exercises 3.E **98**

Contents vii

* 1. Duality 101

[The Dual Space and the Dual Map **101**](#_TOC_250000)

The Null Space and Range of the Dual of a Linear Map **104**

The Matrix of the Dual of a Linear Map **109**

The Rank of a Matrix **111**

Exercises 3.F **113**

1. ***Polynomials* 117**

Complex Conjugate and Absolute Value **118** Uniqueness of Coefﬁcients for Polynomials **120** The Division Algorithm for Polynomials **121** Zeros of Polynomials **122**

Factorization of Polynomials over **C 123** Factorization of Polynomials over **R 126** Exercises 4 **129**

1. ***Eigenvalues, Eigenvectors, and Invariant Subspaces* 131**
   1. Invariant Subspaces **132**

Eigenvalues and Eigenvectors **133** Restriction and Quotient Operators **137** Exercises 5.A **138**

* 1. Eigenvectors and Upper-Triangular Matrices **143**

Polynomials Applied to Operators **143**

Existence of Eigenvalues **145**

Upper-Triangular Matrices **146**

Exercises 5.B **153**

* 1. Eigenspaces and Diagonal Matrices **155**

Exercises 5.C **160**

1. ***Inner Product Spaces* 163**
   1. Inner Products and Norms **164**

Inner Products **164**

Norms **168**

Exercises 6.A **175**

**viii** Contents

* 1. Orthonormal Bases **180**

Linear Functionals on Inner Product Spaces **187**

Exercises 6.B **189**

* 1. Orthogonal Complements and Minimization Problems **193**

Orthogonal Complements **193**

Minimization Problems **198**

Exercises 6.C **201**

1. ***Operators on Inner Product Spaces* 203**
   1. Self-Adjoint and Normal Operators **204**

Adjoints **204**

Self-Adjoint Operators **209**

Normal Operators **212**

Exercises 7.A **214**

* 1. The Spectral Theorem **217**

The Complex Spectral Theorem **217** The Real Spectral Theorem **219** Exercises 7.B **223**

* 1. Positive Operators and Isometries **225**

Positive Operators **225**

Isometries **228**

Exercises 7.C **231**

* 1. Polar Decomposition and Singular Value Decomposition **233**

Polar Decomposition **233**

Singular Value Decomposition **236**

Exercises 7.D **238**

1. ***Operators on Complex Vector Spaces* 241**
   1. Generalized Eigenvectors and Nilpotent Operators **242**

Null Spaces of Powers of an Operator **242**

Generalized Eigenvectors **244**

Nilpotent Operators **248**

Exercises 8.A **249**

* 1. Decomposition of an Operator **252**

Contents **ix**

Description of Operators on Complex Vector Spaces **252**

Multiplicity of an Eigenvalue **254** Block Diagonal Matrices **255** Square Roots **258**

Exercises 8.B **259**

* 1. Characteristic and Minimal Polynomials **261**

The Cayley–Hamilton Theorem **261** The Minimal Polynomial **262** Exercises 8.C **267**

* 1. Jordan Form **270**

Exercises 8.D **274**

1. ***Operators on Real Vector Spaces* 275**
   1. Complexiﬁcation **276**

Complexiﬁcation of a Vector Space **276**

Complexiﬁcation of an Operator **277**

The Minimal Polynomial of the Complexiﬁcation **279** Eigenvalues of the Complexiﬁcation **280** Characteristic Polynomial of the Complexiﬁcation **283** Exercises 9.A **285**

* 1. Operators on Real Inner Product Spaces **287** Normal Operators on Real Inner Product Spaces **287** Isometries on Real Inner Product Spaces **292** Exercises 9.B **294**

1. ***Trace and Determinant* 295**
   1. Trace **296**

Change of Basis **296**

Trace: A Connection Between Operators and Matrices **299**

Exercises 10.A **304**

**x** Contents

* 1. Determinant **307**

Determinant of an Operator **307** Determinant of a Matrix **309** The Sign of the Determinant **320** Volume **323**

Exercises 10.B **330**

***Photo Credits* 333**

***Symbol Index* 335**

***Index* 337**

You are about to teach a course that will probably give students their second exposure to linear algebra. During their ﬁrst brush with the subject, your students probably worked with Euclidean spaces and matrices. In contrast, this course will emphasize abstract vector spaces and linear maps.

*Preface for the Instructor*

The audacious title of this book deserves an explanation. Almost all linear algebra books use determinants to prove that every linear operator on a ﬁnite-dimensional complex vector space has an eigenvalue. Determinants are difﬁcult, nonintuitive, and often deﬁned without motivation. To prove the theorem about existence of eigenvalues on complex vector spaces, most books

must deﬁne determinants, prove that a linear map is not invertible if and only if its determinant equals 0, and then deﬁne the characteristic polynomial. This tortuous (torturous?) path gives students little feeling for why eigenvalues

exist.

In contrast, the simple determinant-free proofs presented here (for example, see 5.21) offer more insight. Once determinants have been banished to the end of the book, a new route opens to the main goal of linear algebra— understanding the structure of linear operators.

This book starts at the beginning of the subject, with no prerequisites other than the usual demand for suitable mathematical maturity. Even if your students have already seen some of the material in the ﬁrst few chapters, they may be unaccustomed to working exercises of the type presented here, most of which require an understanding of proofs.

Here is a chapter-by-chapter summary of the highlights of the book:

Chapter 1: Vector spaces are deﬁned in this chapter, and their basic proper- ties are developed.

•

Chapter 2: Linear independence, span, basis, and dimension are deﬁned in this chapter, which presents the basic theory of ﬁnite-dimensional vector spaces.

•

xi

**xii** Preface for the Instructor

Chapter 3: Linear maps are introduced in this chapter. The key result here is the Fundamental Theorem of Linear Maps (3.22): if T is a linear map on V, then dim V dim null T dim range T. Quotient spaces and duality

•

D C

are topics in this chapter at a higher level of abstraction than other parts of the book; these topics can be skipped without running into problems elsewhere in the book.

Chapter 4: The part of the theory of polynomials that will be needed to understand linear operators is presented in this chapter. This chapter contains no linear algebra. It can be covered quickly, especially if your students are already familiar with these results.

•

Chapter 5: The idea of studying a linear operator by restricting it to small subspaces leads to eigenvectors in the early part of this chapter. The highlight of this chapter is a simple proof that on complex vector spaces, eigenvalues always exist. This result is then used to show that each linear operator on a complex vector space has an upper-triangular matrix with respect to some basis. All this is done without deﬁning determinants or characteristic polynomials!

•

Chapter 6: Inner product spaces are deﬁned in this chapter, and their basic properties are developed along with standard tools such as orthonormal bases and the Gram–Schmidt Procedure. This chapter also shows how orthogonal projections can be used to solve certain minimization problems.

•

Chapter 7: The Spectral Theorem, which characterizes the linear operators for which there exists an orthonormal basis consisting of eigenvectors, is the highlight of this chapter. The work in earlier chapters pays off here with especially simple proofs. This chapter also deals with positive operators, isometries, the Polar Decomposition, and the Singular Value Decomposition.

•

Chapter 8: Minimal polynomials, characteristic polynomials, and gener- alized eigenvectors are introduced in this chapter. The main achievement of this chapter is the description of a linear operator on a complex vector space in terms of its generalized eigenvectors. This description enables one to prove many of the results usually proved using Jordan Form. For example, these tools are used to prove that every invertible linear operator on a complex vector space has a square root. The chapter concludes with a proof that every linear operator on a complex vector space can be put into Jordan Form.

•

Preface for the Instructor **xiii**

Chapter 9: Linear operators on real vector spaces occupy center stage in this chapter. Here the main technique is complexiﬁcation, which is a natural extension of an operator on a real vector space to an operator on a complex vector space. Complexiﬁcation allows our results about complex vector spaces to be transferred easily to real vector spaces. For example, this technique is used to show that every linear operator on a real vector space

•

has an invariant subspace of dimension 1 or 2. As another example, we

show that that every linear operator on an odd-dimensional real vector space

has an eigenvalue.

Chapter 10: The trace and determinant (on complex vector spaces) are deﬁned in this chapter as the sum of the eigenvalues and the product of the eigenvalues, both counting multiplicity. These easy-to-remember deﬁni- tions would not be possible with the traditional approach to eigenvalues, because the traditional method uses determinants to prove that sufﬁcient eigenvalues exist. The standard theorems about determinants now become much clearer. The Polar Decomposition and the Real Spectral Theorem are used to derive the change of variables formula for multivariable integrals in a fashion that makes the appearance of the determinant there seem natural.

•

This book usually develops linear algebra simultaneously for real and complex vector spaces by letting **F** denote either the real or the complex numbers. If you and your students prefer to think of **F** as an arbitrary ﬁeld, then see the comments at the end of Section 1.A. I prefer avoiding arbitrary ﬁelds at this level because they introduce extra abstraction without leading to any new linear algebra. Also, students are more comfortable thinking of polynomials as functions instead of the more formal objects needed for polynomials with coefﬁcients in ﬁnite ﬁelds. Finally, even if the beginning part of the theory were developed with arbitrary ﬁelds, inner product spaces would push consideration back to just real and complex vector spaces.

You probably cannot cover everything in this book in one semester. Going through the ﬁrst eight chapters is a good goal for a one-semester course. If you must reach Chapter 10, then consider covering Chapters 4 and 9 in ﬁfteen minutes each, as well as skipping the material on quotient spaces and duality in Chapter 3.

A goal more important than teaching any particular theorem is to develop in students the ability to understand and manipulate the objects of linear algebra. Mathematics can be learned only by doing. Fortunately, linear algebra has many good homework exercises. When teaching this course, during each class I usually assign as homework several of the exercises, due the next class. Going over the homework might take up a third or even half of a typical class.

**xiv** Preface for the Instructor

Major changes from the previous edition:

This edition contains 561 exercises, including 337 new exercises that were not in the previous edition. Exercises now appear at the end of each section, rather than at the end of each chapter.

•

Many new examples have been added to illustrate the key ideas of linear algebra.

•

Beautiful new formatting, including the use of color, creates pages with an unusually pleasant appearance in both print and electronic versions. As a visual aid, deﬁnitions are in beige boxes and theorems are in blue boxes (in color versions of the book).

•

* Each theorem now has a descriptive name.

New topics covered in the book include product spaces, quotient spaces, and duality.

•

Chapter 9 (Operators on Real Vector Spaces) has been completely rewritten to take advantage of simpliﬁcations via complexiﬁcation. This approach allows for more streamlined presentations in Chapters 5 and 7 because those chapters now focus mostly on complex vector spaces.

•

Hundreds of improvements have been made throughout the book. For example, the proof of Jordan Form (Section 8.D) has been simpliﬁed.

•

Please check the website below for additional information about the book. I may occasionally write new sections on additional topics. These new sections will be posted on the website. Your suggestions, comments, and corrections are most welcome.

Best wishes for teaching a successful linear algebra class!

Sheldon Axler Mathematics Department

San Francisco State University San Francisco, CA 94132, USA

website: linear.axler.net e-mail: [linear@axler.net](mailto:linear@axler.net) Twitter: @AxlerLinear

You are probably about to begin your second exposure to linear algebra. Unlike your ﬁrst brush with the subject, which probably emphasized Euclidean spaces and matrices, this encounter will focus on abstract vector spaces and linear maps. These terms will be deﬁned later, so don’t worry if you do not know what they mean. This book starts from the beginning of the subject, assuming no knowledge of linear algebra. The key point is that you are about to immerse yourself in serious mathematics, with an emphasis on attaining a deep understanding of the deﬁnitions, theorems, and proofs.

*Preface for the Student*

You cannot read mathematics the way you read a novel. If you zip through a page in less than an hour, you are probably going too fast. When you encounter the phrase “as you should verify”, you should indeed do the veriﬁcation, which will usually require some writing on your part. When steps are left out, you need to supply the missing pieces. You should ponder and internalize each deﬁnition. For each theorem, you should seek examples to show why each hypothesis is necessary. Discussions with other students should help.

As a visual aid, deﬁnitions are in beige boxes and theorems are in blue boxes (in color versions of the book). Each theorem has a descriptive name. Please check the website below for additional information about the book. I may occasionally write new sections on additional topics. These new sections will be posted on the website. Your suggestions, comments, and corrections

are most welcome.

Best wishes for success and enjoyment in learning linear algebra!

Sheldon Axler Mathematics Department

San Francisco State University San Francisco, CA 94132, USA

website: linear.axler.net e-mail: [linear@axler.net](mailto:linear@axler.net) Twitter: @AxlerLinear

xv

I owe a huge intellectual debt to the many mathematicians who created linear algebra over the past two centuries. The results in this book belong to the common heritage of mathematics. A special case of a theorem may ﬁrst have been proved in the nineteenth century, then slowly sharpened and improved by many mathematicians. Bestowing proper credit on all the contributors would be a difﬁcult task that I have not undertaken. In no case should the reader assume that any theorem presented here represents my original contribution. However, in writing this book I tried to think about the best way to present lin- ear algebra and to prove its theorems, without regard to the standard methods and proofs used in most textbooks.

*Acknowledgments*

Many people helped make this a better book. The two previous editions of this book were used as a textbook at about 300 universities and colleges. I received thousands of suggestions and comments from faculty and students who used the second edition. I looked carefully at all those suggestions as I was working on this edition. At ﬁrst, I tried keeping track of whose suggestions I used so that those people could be thanked here. But as changes were made and then replaced with better suggestions, and as the list grew longer, keeping track of the sources of each suggestion became too complicated. And lists are boring to read anyway. Thus in lieu of a long list of people who contributed good ideas, I will just say how truly grateful I am to everyone who sent me suggestions and comments. Many many thanks!

Special thanks to Ken Ribet and his giant (220 students) linear algebra class at Berkeley that class-tested a preliminary version of this third edition and that sent me more suggestions and corrections than any other group.

Finally, I thank Springer for providing me with help when I needed it and for allowing me the freedom to make the ﬁnal decisions about the content and appearance of this book. Special thanks to Elizabeth Loew for her wonderful work as editor and David Kramer for unusually skillful copyediting.

Sheldon Axler

xvii

# *Vector Spaces*



*René Descartes explaining his work to Queen Christina of Sweden. Vector spaces are a generalization of the description of a plane using two coordinates, as published by Descartes in 1637.*

1

CHAPTER

Linear algebra is the study of linear maps on ﬁnite-dimensional vector spaces. Eventually we will learn what all these terms mean. In this chapter we will deﬁne vector spaces and discuss their elementary properties.

In linear algebra, better theorems and more insight emerge if complex numbers are investigated along with real numbers. Thus we will begin by introducing the complex numbers and their basic properties.

We will generalize the examples of a plane and ordinary space to **R**n

and **C**n, which we then will generalize to the notion of a vector space. The elementary properties of a vector space will already seem familiar to you.

Then our next topic will be subspaces, which play a role for vector spaces analogous to the role played by subsets for sets. Finally, we will look at sums of subspaces (analogous to unions of subsets) and direct sums of subspaces (analogous to unions of disjoint sets).

LEARNING OBJECTIVES FOR THIS CHAPTER

basic properties of the complex numbers

**R**n and **C**n vector spaces subspaces

sums and direct sums of subspaces

© Springer International Publishing 2015 1

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

DOI 10.1007/978-3-319-11080-6 1

**2** CHAPTER 1 Vector Spaces

**R**n *and* **C**n

1.A

### Complex Numbers

You should already be familiar with basic properties of the set **R** of real numbers. Complex numbers were invented so that we can take square roots of

negative numbers. The idea is to assume we have a square root of 1, denoted

—

i, that obeys the usual rules of arithmetic. Here are the formal deﬁnitions:

1.1 **Deﬁnition *complex numbers***

* A ***complex number*** is an ordered pair .a; b/, where a; b 2 **R**, but

we will write this as a C bi.

* The set of all complex numbers is denoted by **C**:

**C** D fa C bi W a; b 2 **R**g:

* ***Addition and multiplication*** on **C** are deﬁned by

.a C bi/ C .c C di/ D .a C c/ C .b C d/i;

.a C bi/.c C di/ D .ac — bd/ C .ad C bc/iI

here a; b; c; d 2 **R**.

If a 2 **R**, we identify a C 0i with the real number a. Thus we can think of **R** as a subset of **C**. We also usually write 0 C bi as just bi, and we usually

Using multiplication as deﬁned

*The symbol* i *was ﬁrst used to de-*

p

*Leonhard Euler in 1777.*

*note* —1 *by Swiss mathematician*

write 0 C 1i as just i.

above, you should verify that i2 D —1.

Do not memorize the formula for the product of two complex numbers; you can always rederive it by recalling that

i2 1 and then using the usual rules

D —

of arithmetic (as given by 1.3).

1.2 **Example** Evaluate .2 C 3i/.4 C 5i/.

Solution .2 C 3i/.4 C 5i/ D 2 . 4 C 2 . .5i/ C .3i/ . 4 C .3i/.5i/

D 8 C 10i C 12i — 15

D —7 C 22i

SECTION 1.A **R**n and **C**n **3**

1.3 Properties of complex arithmetic

**commutativity**

˛ C ˇ D ˇ C ˛ and ˛ˇ D ˇ˛ for all ˛; ˇ 2 **C**;

**associativity**

.˛Cˇ/C入 D ˛C.ˇC入/ and .˛ˇ/入 D ˛.ˇ入/ for all ˛; ˇ; 入 2 **C**;

**identities**

入C 0 D 入and 入1 D 入for all 入 2 **C**;

**additive inverse**

for every ˛ 2 **C**, there exists a unique ˇ 2 **C** such that ˛ C ˇ D 0;

**multiplicative inverse**

˛ˇ D 1;

**distributive property**

入.˛ C ˇ/ D 入˛ C 入ˇ for all 入; ˛; ˇ 2 **C**.

for every ˛ 2 **C** with ˛ ¤ 0, there exists a unique ˇ 2 **C** such that

The properties above are proved using the familiar properties of real numbers and the deﬁnitions of complex addition and multiplication. The next example shows how commutativity of complex multiplication is proved. Proofs of the other properties above are left as exercises.

1.4 **Example** Show that ˛ˇ D ˇ˛ for all ˛; ˇ; 入 2 **C**.

Solution Suppose ˛ a bi and ˇ c di, where a; b; c; d **R**. Then the deﬁnition of multiplication of complex numbers shows that

D C D C 2

˛ˇ D .a C bi/.c C di/

D .ac — bd/ C .ad C bc/i

and

ˇ˛ D .c C di/.a C bi/

D .ca — db/ C .cb C da/i:

The equations above and the commutativity of multiplication and addition of real numbers show that ˛ˇ D ˇ˛.

**4** CHAPTER 1 Vector Spaces

1.5 **Deﬁnition** —˛***, subtraction,*** 1=˛***, division***

Let ˛; ˇ 2 **C**.

* Let —˛ denote the additive inverse of ˛. Thus —˛ is the unique

complex number such that

˛ C .—˛/ D 0:

* ***Subtraction*** on **C** is deﬁned by

ˇ — ˛ D ˇ C .—˛/:

is the unique complex number such that

˛.1=˛/ D 1:

* ***Division*** on **C** is deﬁned by

ˇ=˛ D ˇ.1=˛/:

* For ˛ ¤ 0, let 1=˛ denote the multiplicative inverse of ˛. Thus 1=˛

So that we can conveniently make deﬁnitions and prove theorems that apply to both real and complex numbers, we adopt the following notation:

1.6 **Notation F**

Throughout this book, **F** stands for either **R** or **C**.

Thus if we prove a theorem involving **F**, we will know that it holds when **F** is replaced with **R** and when **F** is replaced with **C**.

*The letter* **F** *is used because* **R** *and* **C** *are examples of what are called* ***ﬁelds****.*

Elements of **F** are called ***scalars***. The word “scalar”, a fancy word for “number”, is often used when we want to emphasize that an object is a number, as opposed to a vector (vectors will be deﬁned soon).

For ˛ **F** and m a positive integer, we deﬁne ˛m to denote the product of

2

˛ with itself m times:

˛m D ˛ . . . ˛ :

„ƒ‚…

m times

Clearly .˛m/n ˛mn and .˛ˇ/m ˛mˇm for all ˛; ˇ **F** and all positive integers m; n.

D D 2

### Lists

SECTION 1.A **R**n and **C**n **5**

Before deﬁning **R**n and **C**n, we look at two important examples.

1.7 **Example R**2 ***and* R**3

The set **R**2, which you can think of as a plane, is the set of all ordered pairs of real numbers:

•

**R**2 D f.x; y/ W x; y 2 **R**g:

The set **R**3, which you can think of as ordinary space, is the set of all ordered triples of real numbers:

•

**R**3 D f.x; y; z/ W x; y; z 2 **R**g:

To generalize **R**2 and **R**3 to higher dimensions, we ﬁrst need to discuss the concept of lists.

1.8 **Deﬁnition *list, length***

Suppose n is a nonnegative integer. A ***list*** of ***length*** n is an ordered collection of n elements (which might be numbers, other lists, or more abstract entities) separated by commas and surrounded by parentheses. A

list of length n looks like this:

.x1;:::; xn/:

Two lists are equal if and only if they have the same length and the same elements in the same order.

Thus a list of length 2 is an ordered pair, and a list of length 3 is an ordered triple.

Sometimes we will use the word ***list*** without specifying its length. Re- member, however, that by deﬁnition each list has a ﬁnite length that is a nonnegative integer. Thus an object that looks like

*Many mathematicians call a list of length* n *an* n***-tuple****.*

.x1; x2;::: /;

which might be said to have inﬁnite length, is not a list.

A list of length 0 looks like this: . /. We consider such an object to be a list so that some of our theorems will not have trivial exceptions.

Lists differ from sets in two ways: in lists, order matters and repetitions have meaning; in sets, order and repetitions are irrelevant.

**6** CHAPTER 1 Vector Spaces

1.9 **Example *lists versus sets***

The lists .3; 5/ and .5; 3/ are not equal, but the sets 3; 5 and 5; 3 are equal.

* f g f g
* The lists .4; 4/ and .4; 4; 4/ are not equal (they do not have the same length), although the sets f4; 4g and f4; 4; 4g both equal the set f4g.

**F***n*

To deﬁne the higher-dimensional analogues of **R**2 and **R**3, we will simply replace **R** with **F** (which equals **R** or **C**) and replace theFana 2 or 3 with an arbitrary positive integer. Speciﬁcally, ﬁx a positive integer n for the rest of

this section.

1.10 **Deﬁnition**

**F**n

**F**n is the set of all lists of length n of elements of **F**:

**F**n D f.x1;:::; xn/ W xj 2 **F** for j D 1; :::; ng:

For .x1;:::; xn/ 2 **F**n and j 2 f1; :::; ng, we say that xj is the j th

***coordinate*** of .x1;:::; xn/.

If **F R** and n equals 2 or 3, then this deﬁnition of **F**n agrees with our previous notions of **R**2 and **R**3.

D

1.11 **Example C**4 is the set of all lists of four complex numbers:

**C**4 D f.z1; z2; z3; z4/ W z1; z2; z3; z4 2 **C**g:

*For an amusing account of how* **R**3 *would be perceived by crea- tures living in* **R**2*, read* ***Flatland: A Romance of Many Dimensions****, by Edwin A. Abbott. This novel, published in 1884, may help you imagine a physical space of four or more dimensions.*

乏

If n 乏 4, we cannot visualize **R**n as a physical object. Similarly, **C**1 can be thought of as a plane, but for n 2, the human brain cannot provide a full image of **C**n. However, even if n is large, we can perform algebraic manip- ulations in **F**n as easily as in **R**2 or **R**3. For example, addition in **F**n is deﬁned

as follows:

SECTION 1.A **R**n and **C**n **7**

1.12 **Deﬁnition *addition in* F**n

***Addition*** in **F**n is deﬁned by adding corresponding coordinates:

.x1;:::; xn/ C .y1;:::; yn/ D .x1 C y1;:::; xn C yn/:

Often the mathematics of **F**n becomes cleaner if we use a single letter to denote a list of n numbers, without explicitly writing the coordinates. For example, the result below is stated with x and y in **F**n even though the proof requires the more cumbersome notation of .x1;:::; xn/ and .y1;:::; yn/.

1.13 Commutativity of addition in **F**n

If x; y 2 **F**n, then x C y D y C x.

Proof Suppose x D .x1;:::; xn/ and y D .y1;:::; yn/. Then

x C y D .x1;:::; xn/ C .y1;:::; yn/

D .x1 C y1;:::; xn C yn/

D .y1 C x1;:::; yn C xn/

D .y1;:::; yn/ C .x1;:::; xn/

D y C x;

where the second and fourth equalities above hold because of the deﬁnition of addition in **F**n and the third equality holds because of the usual commutativity of addition in **F**.

If a single letter is used to denote an element of **F**n, then the same letter with appropriate subscripts is often used

*The symbol means “end of the proof”.*

when coordinates must be displayed. For example, if x **F**n, then letting x equal .x1;:::; xn/ is good notation, as shown in the proof above. Even better, work with just x and avoid explicit coordinates when possible.

2

1.14 **Deﬁnition** 0

Let 0 denote the list of length n whose coordinates are all 0:

0 D .0; : : : ; 0/:

**8** CHAPTER 1 Vector Spaces

Here we are using the symbol 0 in two different ways—on the left side of the equation in 1.14, the symbol 0 denotes a list of length n, whereas on the right side, each 0 denotes a number. This potentially confusing practice actually

causes no problems because the context always makes clear what is intended.

1.15 **Example** Consider the statement that 0 is an additive identity for **F**n:

x C 0 D x for all x 2 **F**n:

Is the 0 above the number 0 or the list 0?

Solution Here 0 is a list, because we have not deﬁned the sum of an element of **F**n (namely, x) and the number 0.

*Elements of* **R**2 *can be thought of as points*

*or as vectors.*

*A vector.*

A picture can aid our intuition. We will draw pictures in **R**2 because we can sketch this space on 2-dimensional surfaces such as paper and blackboards. A typical element of **R**2 is a point x

.x1; x2/. Sometimes we think of x not

*(x*1 *, x*2*)*

*x*

D

as a point but as an arrow starting at the origin and ending at .x1; x2/, as shown here. When we think of x as an arrow,

we refer to it as a ***vector***.

When we think of vectors in **R**2 as arrows, we can move an arrow parallel to itself (not changing its length or di- rection) and still think of it as the same vector. With that viewpoint, you will often gain better understanding by dis- pensing with the coordinate axes and the explicit coordinates and just think- ing of the vector, as shown here.

*x*

*x*

Whenever we use pictures in **R**2

*Mathematical models of the econ- omy can have thousands of vari-*

*ables, say* x1;:::; x5000*, which*

*means that we must operate in*

**R**5000*.*

*Such a space cannot be*

*dealt with geometrically. However, the algebraic approach works well. Thus our subject is called* ***linear algebra****.*

or use the somewhat vague language of points and vectors, remember that these are just aids to our understand- ing, not substitutes for the actual math- ematics that we will develop. Although we cannot draw good pictures in high- dimensional spaces, the elements of these spaces are as rigorously deﬁned as elements of **R**2.

SECTION 1.A **R**n and **C**n **9**

For example, .2; 3; 17; 兀; p2/ is an element of **R**5, and we may casually refer to it as a point in **R**5 or a vector in **R**5 without worrying about whether the geometry of **R**5 has any physical meaning.

—

Recall that we deﬁned the sum of two elements of **F**n to be the element of **F**n obtained by adding corresponding coordinates; see 1.12. As we will now see, addition has a simple geometric interpretation in the special case of **R**2.

Suppose we have two vectors x and y in **R**2 that we want to add. Move the vector y parallel to itself so that its

*y*

*x* + *y*

*x*

initial point coincides with the end point of the vector x, as shown here. The sum x y then equals the vector whose

C

initial point equals the initial point of

x and whose end point equals the end point of the vector y, as shown here.

*The sum of two vectors.*

In the next deﬁnition, the 0 on the right side of the displayed equation below is the list 0 2 **F**n.

1.16 **Deﬁnition *additive inverse in* F**n

For x 2 **F**n, the ***additive inverse*** of x, denoted —x, is the vector —x 2 **F**n

such that

x C .—x/ D 0:

In other words, if x D .x1;:::; xn/, then —x D .—x1;:::; —xn/.

For a vector x **R**2, the additive in- verse x is the vector parallel to x and with the same length as x but pointing in

—

2

the opposite direction. The ﬁgure here illustrates this way of thinking about the additive inverse in **R**2.

Having dealt with addition in **F**n, we now turn to multiplication. We could

*x*

一*x*

*A vector and its additive inverse.*

deﬁne a multiplication in **F**n in a similar fashion, starting with two elements of **F**n and getting another element of **F**n by multiplying corresponding coor- dinates. Experience shows that this deﬁnition is not useful for our purposes. Another type of multiplication, called scalar multiplication, will be central to our subject. Speciﬁcally, we need to deﬁne what it means to multiply an element of **F**n by an element of **F**.

**10** CHAPTER 1 Vector Spaces

1.17 **Deﬁnition *scalar multiplication in* F**n

The ***product*** of a number 入and a vector in **F**n is computed by multiplying each coordinate of the vector by 入:

入.x1;:::; xn/ D .入x1;:::; 入xn/I

here 入 2 **F** and .x1;:::; xn/ 2 **F**n.

Scalar multiplication has a nice ge- ometric interpretation in **R**2. If 入 is a positive number and x is a vector in **R**2, then 入 x is the vector that points in the same direction as x and whose length is 入 times the length of x. In other words, to get 入 x, we shrink or stretch x by a factor of 入, depending on whether 入<1 or 入> 1.

If 入is a negative number and x is a vector in **R**2, then 入x is the vector that

*In scalar multiplication, we multi- ply together a scalar and a vector, getting a vector. You may be famil- iar with the dot product in* **R**2 *or* **R**3*, in which we multiply together two vectors and get a scalar. Gen- eralizations of the dot product will become important when we study inner products in Chapter 6.*



*x*

*(*一3*/*2*) x*

*(*1*/*2*) x*

points in the direction opposite to that of x and whose length is 入 times the length of x, as shown here.

j j

*Scalar multiplication.*

### Digression on Fields

A ***ﬁeld*** is a set containing at least two distinct elements called 0 and 1, along with operations of addition and multiplication satisfying all the properties

listed in 1.3. Thus **R** and **C** are ﬁelds, as is the set of rational numbers along with the usual operations of addition and multiplication. Another example of

a ﬁeld is the set 0; 1 with the usual operations of addition and multiplication except that 1 1 is deﬁned to equal 0.

C

f g

In this book we will not need to deal with ﬁelds other than **R** and **C**. However, many of the deﬁnitions, theorems, and proofs in linear algebra that work for both **R** and **C** also work without change for arbitrary ﬁelds. If you prefer to do so, throughout Chapters 1, 2, and 3 you can think of **F** as denoting an arbitrary ﬁeld instead of **R** or **C**, except that some of the examples and

exercises require that for each positive integer n we have 1 C 1 C . .. C 1 ¤ 0.

„ ƒ‚ …

n times

EXERCISES 1.A

SECTION 1.A **R**n and **C**n **11**

1. Suppose a and b are real numbers, not both 0. Find real numbers c and

d such that

1=.a C bi/ D c C di:

1. Show that —1 C p3i 2

is a cube root of 1 (meaning that its cube equals 1).

1. Find two distinct square roots of i.
2. Show that ˛ C ˇ D ˇ C ˛ for all ˛; ˇ 2 **C**.
3. Show that .˛ C ˇ/ C 入 D ˛ C .ˇ C 入/ for all ˛; ˇ; 入 2 **C**.
4. Show that .˛ˇ/入 D ˛.ˇ入/ for all ˛; ˇ; 入 2 **C**.
5. Show that for every ˛ 2 **C**, there exists a unique ˇ 2 **C** such that

˛ C ˇ D 0.

1. Show that for every ˛ 2 **C** with ˛ ¤ 0, there exists a unique ˇ 2 **C** such that ˛ˇ D 1.
2. Show that 入.˛ C ˇ/ D 入˛ C 入ˇ for all 入; ˛; ˇ 2 **C**.
3. Find x 2 **R**4 such that

.4; —3; 1; 7/ C 2x D .5; 9; —6; 8/:

1. Explain why there does not exist 入 2 **C** such that

入.2 — 3i; 5 C 4i; —6 C 7i/ D .12 — 5i; 7 C 22i; —32 — 9i/:

1. Show that .x C y/ C z D x C .y C z/ for all x; y; z 2 **F**n.
2. Show that .ab/x D a.bx/ for all x 2 **F**n and all a; b 2 **F**.
3. Show that 1x D x for all x 2 **F**n.
4. Show that 入.x C y/ D 入x C 入y for all 入 2 **F** and all x; y 2 **F**n.
5. Show that .a C b/x D ax C bx for all a; b 2 **F** and all x 2 **F**n.

**12** CHAPTER 1 Vector Spaces

*Deﬁnition of Vector Space*

1.B

The motivation for the deﬁnition of a vector space comes from properties of addition and scalar multiplication in **F**n: Addition is commutative, associative, and has an identity. Every element has an additive inverse. Scalar multiplica-

tion is associative. Scalar multiplication by 1 acts as expected. Addition and

scalar multiplication are connected by distributive properties.

We will deﬁne a vector space to be a set V with an addition and a scalar multiplication on V that satisfy the properties in the paragraph above.

1.18 **Deﬁnition *addition, scalar multiplication***

* An ***addition*** on a set V is a function that assigns an element uC*v* 2 V

to each pair of elements u; *v* 2 V.

* A ***scalar multiplication*** on a set V is a function that assigns an ele-

ment 入*v* 2 V to each 入 2 **F** and each *v* 2 V.

Now we are ready to give the formal deﬁnition of a vector space.

1.19 **Deﬁnition *vector space***

A ***vector space*** is a set V along with an addition on V and a scalar multi- plication on V such that the following properties hold:

#### commutativity

u C *v* D *v* C u for all u; *v* 2 V ;

#### associativity

.u C *v*/ C *w* D u C .*v* C *w*/ and .ab/*v* D a.b*v*/ for all u; *v*; *w* 2 V

and all a; b 2 **F**;

#### additive identity

there exists an element 0 2 V such that *v* C 0 D *v* for all *v* 2 V ;

#### additive inverse

for every *v* 2 V, there exists *w* 2 V such that *v* C *w* D 0;

#### multiplicative identity

1*v* D *v* for all *v* 2 V ;

#### distributive properties

a.u C *v*/ D au C a*v* and .a C b/*v* D a*v* C b*v* for all a; b 2 **F** and all u; *v* 2 V.

SECTION 1.B Deﬁnition of Vector Space **13**

The following geometric language sometimes aids our intuition.

1.20 **Deﬁnition *vector, point***

Elements of a vector space are called ***vectors*** or ***points***.

The scalar multiplication in a vector space depends on **F**. Thus when we need to be precise, we will say that V is a ***vector space over* F** instead of saying simply that V is a vector space. For example, **R**n is a vector space over **R**, and **C**n is a vector space over **C**.

1.21 **Deﬁnition *real vector space, complex vector space***

* A vector space over **R** is called a ***real vector space***.
* A vector space over **C** is called a ***complex vector space***.

Usually the choice of **F** is either obvious from the context or irrelevant. Thus we often assume that **F** is lurking in the background without speciﬁcally mentioning it.

With the usual operations of addition and scalar multiplication, **F**n is a vector space over **F**, as you should verify. The example of **F**n motivated our deﬁnition of vector space.

*The simplest vector space contains*

*only one point. In other words,* f0g

*is a vector space.*

1.22 **Example F**1 is deﬁned to be the set of all sequences of elements of **F**:

**F**1 D f.x1; x2;:::/ W xj 2 **F** for j D 1; 2; ::: g:

Addition and scalar multiplication on **F**1 are deﬁned as expected:

.x1; x2;:::/ C .y1; y2;:::/ D .x1 C y1; x2 C y2;::: /;

入.x1; x2;:::/ D .入x1; 入x2;::: /:

With these deﬁnitions, **F**1 becomes a vector space over **F**, as you should verify. The additive identity in this vector space is the sequence of all 0’s.

Our next example of a vector space involves a set of functions.

**14** CHAPTER 1 Vector Spaces

1.23 **Notation**

**F**S

* If S is a set, then **F**S denotes the set of functions from S to **F**.
* For f; g 2 **F**S , the ***sum*** f C g 2 **F**S is the function deﬁned by

.f C g/.x/ D f .x/ C g.x/

for all x 2 S.

* For 入 2 **F** and f 2 **F**S , the ***product*** 入 f 2 **F**S is the function

deﬁned by

.入f /.x/ D 入f .x/

for all x 2 S.

As an example of the notation above, if S is the interval Œ0; 1] and **F R**, then **R**Œ0;1] is the set of real-valued functions on the interval Œ0; 1].

You should verify all three bullet points in the next example.

D

1.24 **Example F**S ***is a vector space***

If S is a nonempty set, then **F**S (with the operations of addition and scalar multiplication as deﬁned above) is a vector space over **F**.

•

* The additive identity of **F**S is the function 0 W S ! **F** deﬁned by

0.x/ D 0

for all x 2 S.

For f **F**S , the additive inverse of f is the function f S **F**

* 2 — W !

deﬁned by

.—f /.x/ D —f .x/

for all x 2 S.

Our previous examples of vector spaces, **F**n and **F**1, are special cases

*The elements of the vector space* **R**Œ0;1] *are real-valued functions on* Œ0; 1]*, not lists. In general, a vector space is an abstract entity whose*

*elements might be lists, functions, or weird objects.*

of the vector space **F**S because a list of length n of numbers in **F** can be thought of as a function from 1; 2; :::; n to **F**

f g

and a sequence of numbers in **F** can be thought of as a function from the set of

positive integers to **F**. In other words, we can think of **F**n as **F**f1;2;:::;ng and

we can think of **F**1 as **F**f1;2;::: g.

SECTION 1.B Deﬁnition of Vector Space **15**

Soon we will see further examples of vector spaces, but ﬁrst we need to develop some of the elementary properties of vector spaces.

The deﬁnition of a vector space requires that it have an additive identity.

The result below states that this identity is unique.

1.25 Unique additive identity

A vector space has a unique additive identity.

Proof Suppose 0 and 00 are both additive identities for some vector space V. Then

00 D 00 C 0 D 0 C 00 D 0;

where the ﬁrst equality holds because 0 is an additive identity, the second equality comes from commutativity, and the third equality holds because 00 is an additive identity. Thus 00 0, proving that V has only one additive

D

identity.

Each element *v* in a vector space has an additive inverse, an element *w* in the vector space such that *v w* 0. The next result shows that each element in a vector space has only one additive inverse.

C D

1.26 Unique additive inverse

Every element in a vector space has a unique additive inverse.

Proof Suppose V is a vector space. Let *v* V. Suppose *w* and *w*0 are additive inverses of *v*. Then

*w* D *w* C 0 D *w* C .*v* C *w*0/ D .*w* C *v*/ C *w*0 D 0 C *w*0 D *w*0:

2

Thus *w* D *w*0, as desired.

Because additive inverses are unique, the following notation now makes sense.

1.27 **Notation** —*v****,*** *w* — *v*

Let *v*; *w* 2 V. Then

* —*v* denotes the additive inverse of *v*;
* *w* — *v* is deﬁned to be *w* C .—*v*/.

**16** CHAPTER 1 Vector Spaces

Almost all the results in this book involve some vector space. To avoid having to restate frequently that V is a vector space, we now make the necessary declaration once and for all:

2

1.28 **Notation** V

For the rest of the book, V denotes a vector space over **F**.

In the next result, 0 denotes a scalar (the number 0 **F**) on the left side of the equation and a vector (the additive identity of V ) on the right side of the equation.

1.29 The number 0 times a vector

0*v*D 0 for every *v* 2 V.

Proof For *v* 2 V, we have

0*v* D .0 C 0/*v* D 0*v* C 0*v*:

*Note that 1.29 asserts something about scalar multiplication and the*

*additive identity of* V*. The only*

*part of the deﬁnition of a vector*

*space that connects scalar multi- plication and vector addition is the distributive property. Thus the dis- tributive property must be used in the proof of 1.29.*

Adding the additive inverse of 0*v* to both sides of the equation above gives 0 0*v*, as desired.

D

In the next result, 0 denotes the addi- tive identity of V. Although their proofs

are similar, 1.29 and 1.30 are not identical. More precisely, 1.29 states that the product of the scalar 0 and any vector equals the vector 0, whereas 1.30 states that the product of any scalar and the vector 0 equals the vector 0.

1.30 A number times the vector 0 a0 D 0 for every a 2 **F**.

Proof For a 2 **F**, we have

a0 D a.0 C 0/ D a0 C a0:

Adding the additive inverse of a0 to both sides of the equation above gives

1. D a0, as desired.

Now we show that if an element of V is multiplied by the scalar 1, then the result is the additive inverse of the element of V.

—

SECTION 1.B Deﬁnition of Vector Space **17**

1.31 The number —1 times a vector

.—1/*v* D —*v* for every *v* 2 V.

Proof For *v* 2 V, we have

*v* C .—1/*v* D 1*v* C .—1/*v* D 1 C .—1/ *v* D 0*v* D 0:

（ ）

This equation says that . 1/*v*, when added to *v*, gives 0. Thus . 1/*v* is the additive inverse of *v*, as desired.

— —

EXERCISES 1.B

1. Prove that —.—*v*/ D *v* for every *v* 2 V.
2. Suppose a 2 **F**, *v* 2 V, and a*v* D 0. Prove that a D 0 or *v* D 0.
3. Suppose *v*; *w* 2 V. Explain why there exists a unique x 2 V such that

*v* C 3x D *w*.

1. The empty set is not a vector space. The empty set fails to satisfy only one of the requirements listed in 1.19. Which one?
2. Show that in the deﬁnition of a vector space (1.19), the additive inverse condition can be replaced with the condition that

0*v* D 0 for all *v* 2 V:

Here the 0 on the left side is the number 0, and the 0 on the right side is the additive identity of V. (The phrase “a condition can be replaced” in a deﬁnition means that the collection of objects satisfying the deﬁnition is

unchanged if the original condition is replaced with the new condition.)

1. Let and denote two distinct objects, neither of which is in **R**. Deﬁne an addition and scalar multiplication on **R** as you could guess from the notation. Speciﬁcally, the sum and product of two

[ f1g [ f—1g

1 —1

real numbers is as usual, and for t 2 **R** deﬁne

8

8ˆ<—1 if t < 0;

t 1D ˆ0 if t D 0;

ˆ<

:1 if t > 0;

1 if t < 0; t .—1/ D 0 if t D 0;

—1 if t > 0;

ˆ:

t C 1D1 C t D 1; t C .—1/ D .—1/ C t D —1; 1 C 1D 1; .—1/ C .—1/ D —1; 1C .—1/ D 0:

Is **R** [ f1g [ f—1g a vector space over **R**? Explain.

**18** CHAPTER 1 Vector Spaces

## *Subspaces*

1.C

By considering subspaces, we can greatly expand our examples of vector spaces.

1.32 **Deﬁnition *subspace***

A subset U of V is called a ***subspace*** of V if U is also a vector space (using the same addition and scalar multiplication as on V ).

1.33 **Example** f.x1; x2; 0/ W x1; x2 2 **F**g is a subspace of **F**3.

The next result gives the easiest way to check whether a subset of a vector space is a subspace.

*Some mathematicians use the term* ***linear subspace****, which means the same as subspace.*

1.34 Conditions for a subspace

A subset U of V is a subspace of V if and only if U satisﬁes the following three conditions:

**additive identity**

0 2 U

**closed under addition**

u; *w* 2 U implies u C *w* 2 U ;

**closed under scalar multiplication**

a 2 **F** and u 2 U implies au 2 U.

Proof If U is a subspace of V, then U satisﬁes the three conditions above by the deﬁnition of vector space.

Conversely, suppose U satisﬁes the

*The additive identity condition*

*above could be replaced with the condition that* U *is nonempty* (*then*

*and using the condition that* U *is*

*closed under scalar multiplication*

*taking* u 2 U*, multiplying it by* 0*,*

*would imply that* 0 2 U )*. However,*

*if* U *is indeed a subspace of* V*, then the easiest way to show that* U *is nonempty is to show that* 0 2 U*.*

three conditions above. The ﬁrst con-

dition above ensures that the additive identity of V is in U.

The second condition above ensures that addition makes sense on U. The third condition ensures that scalar mul- tiplication makes sense on U.

SECTION 1.C Subspaces **19**

If u U, then u [which equals . 1/u by 1.31] is also in U by the third condition above. Hence every element of U has an additive inverse in U.

2 — —

The other parts of the deﬁnition of a vector space, such as associativity and commutativity, are automatically satisﬁed for U because they hold on the larger space V. Thus U is a vector space and hence is a subspace of V.

The three conditions in the result above usually enable us to determine quickly whether a given subset of V is a subspace of V. You should verify all the assertions in the next example.

1.35 **Example *subspaces***

1. If b 2 **F**, then

f.x1; x2; x3; x4/ 2 **F** W x3 D 5x4 C bg

4

is a subspace of **F**4 if and only if b D 0.

1. The set of continuous real-valued functions on the interval Œ0; 1] is a subspace of **R**Œ0;1].
2. The set of differentiable real-valued functions on **R** is a subspace of **RR** .
3. The set of differentiable real-valued functions f on the interval .0; 3/

such that f 0.2/ D b is a subspace of **R**.0;3/ if and only if b D 0.

1. The set of all sequences of complex numbers with limit 0 is a subspace of **C**1.

*Clearly* f0g *is the smallest sub-*

*space of* V *and* V *itself is the largest subspace of* V*. The empty set is not a subspace of* V *because*

*a subspace must be a vector space and hence must contain at least one element, namely, an additive identity.*

Verifying some of the items above shows the linear structure underlying parts of calculus. For example, the sec- ond item above requires the result that the sum of two continuous functions is continuous. As another example, the fourth item above requires the result

that for a constant c, the derivative of

cf equals c times the derivative of f .

The subspaces of **R**2 are precisely 0 , **R**2, and all lines in **R**2 through the origin. The subspaces of **R**3 are precisely 0 , **R**3, all lines in **R**3 through the origin, and all planes in **R**3 through the origin. To prove that all these objects

f g

f g

are indeed subspaces is easy—the hard part is to show that they are the only subspaces of **R**2 and **R**3. That task will be easier after we introduce some additional tools in the next chapter.

**20** CHAPTER 1 Vector Spaces

### Sums of Subspaces

*The union of subspaces is rarely a subspace* (*see Exercise 12*)*, which is why we usually work with sums rather than unions.*

When dealing with vector spaces, we are usually interested only in subspaces, as opposed to arbitrary subsets. The notion of the sum of subspaces will be useful.

1.36 **Deﬁnition *sum of subsets***

Suppose U1;:::; Um are subsets of V. The ***sum*** of U1;:::; Um, denoted

U1 C ... C Um, is the set of all possible sums of elements of U1;:::; Um.

More precisely,

U1 C ... C Um D fu1 C ... C um W u1 2 U1;:::; um 2 Umg:

Let’s look at some examples of sums of subspaces.

* 1. **Example** Suppose U is the set of all elements of **F**3 whose second and third coordinates equal 0, and W is the set of all elements of **F**3 whose ﬁrst and third coordinates equal 0:

U D f.x; 0; 0/ 2 **F**3 W x 2 **F**g and W D f.0; y; 0/ 2 **F**3 W y 2 **F**g:

Then

as you should verify.

U C W D f.x; y; 0/ W x; y 2 **F**g;

* 1. **Example** Suppose that U D f.x; x; y; y/ 2 **F**4 W x; y 2 **F**g and

W D f.x; x; x; y/ 2 **F**4 W x; y 2 **F**g. Then

U C W D f.x; x; y; z/ 2 **F**4 W x; y; z 2 **F**g;

as you should verify.

The next result states that the sum of subspaces is a subspace, and is in fact the smallest subspace containing all the summands.

1.39 Sum of subspaces is the smallest containing subspace

Suppose U1;:::; Um are subspaces of V. Then U1 C . . . C Um is the

smallest subspace of V containing U1;:::; Um.

SECTION 1.C Subspaces **21**

Proof It is easy to see that 0 U1 Um and that U1 Um is closed under addition and scalar multiplication. Thus 1.34 implies that U1 Um is a subspace of V.

C ... C

2 C ... C C ... C

Clearly U1;:::; Um are all con-

*Sums of subspaces in the theory of vector spaces are analogous to unions of subsets in set theory. Given two subspaces of a vector space, the smallest subspace con- taining them is their sum. Analo- gously, given two subsets of a set, the smallest subset containing them is their union.*

tained in U1 Um (to see this, consider sums u1 um where all except one of the u’s are 0). Con- versely, every subspace of V contain- ing U1;:::; Um contains U1 Um

C . . .C

C . . .C

C. . . C

(because subspaces must contain all ﬁ- nite sums of their elements). Thus

U1 Um is the smallest subspace of V containing U1;:::; Um.

C ... C

### Direct Sums

Suppose U1;:::; Um are subspaces of V. Every element of U1 Um

C. . .C

can be written in the form

u1 C ... C um;

where each uj is in Uj . We will be especially interested in cases where each vector in U1 Um can be represented in the form above in only one way. This situation is so important that we give it a special name: direct sum.

C. . . C

1.40 **Deﬁnition *direct sum***

Suppose U1;:::; Um are subspaces of V.

* The sum U1 C ... C Um is called a ***direct sum*** if each element

u1 C ... C um, where each uj is in Uj .

of U1 C ... C Um can be written in only one way as a sum

* If U1 C ... C Um is a direct sum, then U1 ˚ ... ˚ Um denotes

U1 C. . .C Um, with the ˚ notation serving as an indication that

this is a direct sum.

* 1. **Example** Suppose U is the subspace of **F**3 of those vectors whose last coordinate equals 0, and W is the subspace of **F**3 of those vectors whose ﬁrst two coordinates equal 0:

U D f.x; y; 0/ 2 **F**3 W x; y 2 **F**g and W D f.0; 0; z/ 2 **F**3 W z 2 **F**g:

Then **F**3 D U ˚ W, as you should verify.

R:: 7JC CHAPTER 1 Vector Spaces

* 1. **Example** Suppose Uj is the subspace of **F**n of those vectors whose coordinates are all 0, except possibly in the j th slot (thus, for example, U2 D f.0; x; 0;:::; 0/ 2 **F**n W x 2 **F**g). Then

**F**n D U1 ˚ ... ˚ Un;

as you should verify.

Sometimes nonexamples add to our understanding as much as examples.

* 1. **Example** Let

U1 D f.x; y; 0/ 2 **F**3 W x; y 2 **F**g; U2 D f.0; 0; z/ 2 **F**3 W z 2 **F**g; U3 D f.0; y; y/ 2 **F**3 W y 2 **F**g:

Show that U1 C U2 C U3 is not a direct sum.

Solution Clearly **F**3 U1 U2 U3, because every vector .x; y; z/ **F**3

D C C 2

can be written as

.x; y; z/ D .x; y; 0/ C .0; 0; z/ C .0; 0; 0/;

where the ﬁrst vector on the right side is in U1, the second vector is in U2, and the third vector is in U3.

However, **F**3 does not equal the direct sum of U1; U2; U3, because the vector .0; 0; 0/ can be written in two different ways as a sum u1 u2 u3, with each uj in Uj . Speciﬁcally, we have

C C

.0; 0; 0/ D .0; 1; 0/ C .0; 0; 1/ C .0; —1; —1/

and, of course,

.0; 0; 0/ D .0; 0; 0/ C .0; 0; 0/ C .0; 0; 0/;

where the ﬁrst vector on the right side of each equation above is in U1, the second vector is in U2, and the third vector is in U3.

*The symbol* ˚*, which is a plus*

*sign inside a circle, serves as a re- minder that we are dealing with a special type of sum of subspaces— each element in the direct sum can be represented only one way as a sum of elements from the speciﬁed subspaces.*

The deﬁnition of direct sum requires that every vector in the sum have a unique representation as an appropriate sum. The next result shows that when deciding whether a sum of subspaces is a direct sum, we need only consider

whether 0 can be uniquely written as an

appropriate sum.

SECTION 1.C Subspaces **23**

1.44 Condition for a direct sum

Suppose U1;:::; Um are subspaces of V. Then U1 C ... C Um is a direct

sum if and only if the only way to write 0 as a sum u1 C ... C um, where

each uj is in Uj , is by taking each uj equal to 0.

Proof First suppose U1 Um is a direct sum. Then the deﬁnition of direct sum implies that the only way to write 0 as a sum u1 um, where each uj is in Uj , is by taking each uj equal to 0.

Now suppose that the only way to write 0 as a sum u1 C ... C um, where each uj is in Uj , is by taking each uj equal to 0. To show that U1 C. . .CUm is a direct sum, let *v* 2 U1 C ... C Um. We can write

C. . . C

C ... C

*v* D u1 C ... C um

for some u1 U1;:::; um Um. To show that this representation is unique, suppose we also have

2 2

*v* D *v*1 C ... C *v*m;

where *v*1 2 U1;:::; *v*m 2 Um. Subtracting these two equations, we have

0 D .u1 — *v*1/ C ... C .um — *v*m/:

Because u1 — *v*1 2 U1;:::; um — *v*m 2 Um, the equation above implies that each uj — *v*j equals 0. Thus u1 D *v*1;:::; um D *v*m, as desired.

The next result gives a simple condition for testing which pairs of sub- spaces give a direct sum.

1.45 Direct sum of two subspaces

Suppose U and W are subspaces of V. Then U C W is a direct sum if

and only if U \ W D f0g.

Proof First suppose that U C W is a direct sum. If *v* 2 U \ W, then 0 D *v* C .—*v*/, where *v* 2 U and —*v* 2 W. By the unique representation of 0 as the sum of a vector in U and a vector in W, we have *v* D 0. Thus U \ W D f0g, completing the proof in one direction.

To prove the other direction, now suppose U \ W D f0g. To prove that

U C W is a direct sum, suppose u 2 U, *w* 2 W, and

0 D u C *w*:

To complete the proof, we need only show that u D *w* D 0 (by 1.44). The equation above implies that u D —*w* 2 W. Thus u 2 U \ W. Hence u D 0, which by the equation above implies that *w* D 0, completing the proof.

**24** CHAPTER 1 Vector Spaces

*Sums of subspaces are analogous to unions of subsets. Similarly, di- rect sums of subspaces are analo- gous to disjoint unions of subsets. No two subspaces of a vector space can be disjoint, because both con-*

*tain* 0*. So disjointness is replaced,*

*at least in the case of two sub-*

*spaces, with the requirement that the intersection equals*f0g*.*

EXERCISES 1.C

The result above deals only with the case of two subspaces. When ask- ing about a possible direct sum with more than two subspaces, it is not enough to test that each pair of the

subspaces intersect only at 0. To see

this, consider Example 1.43. In that

nonexample of a direct sum, we have

U1 \ U2 D U1 \ U3 D U2 \ U3 D f0g.

1. For each of the following subsets of **F**3, determine whether it is a sub- space of **F**3:

(a) f.x1; x2; x3/ 2 **F**3 W x1 C 2x2 C 3x3 D 0g;

(b) f.x1; x2; x3/ 2 **F**3 W x1 C 2x2 C 3x3 D 4g;

(c) f.x1; x2; x3/ 2 **F**3 W x1x2x3 D 0g;

(d) f.x1; x2; x3/ 2 **F**3 W x1 D 5x3g.

1. Verify all the assertions in Example 1.35.
2. Show that the set of differentiable real-valued functions f on the interval

.—4; 4/ such that f 0.—1/ D 3f .2/ is a subspace of **R**. 4;4/.

**4** Suppose b 2 **R**. Show that the sRet of continuous real-valued functions f

on the interval Œ0; 1] such that only if b D 0.

0

1 f D b is a subspace of **R**Œ0;1] if and

1. Is **R**2 a subspace of the complex vector space **C**2?
2. (a) Is f.a; b; c/ 2 **R**3 W a3 D b3g a subspace of **R**3?

(b) Is f.a; b; c/ 2 **C**3 W a3 D b3g a subspace of **C**3?

1. Give an example of a nonempty subset U of **R**2 such that U is closed under addition and under taking additive inverses (meaning —u 2 U whenever u 2 U ), but U is not a subspace of **R**2.
2. Give an example of a nonempty subset U of **R**2 such that U is closed under scalar multiplication, but U is not a subspace of **R**2.

SECTION 1.C Subspaces **25**

1. A function f **R R** is called ***periodic*** if there exists a positive number p such that f .x/ f .x p/ for all x **R**. Is the set of periodic functions from **R** to **R** a subspace of **RR** ? Explain.

D C 2

W !

1. Suppose U1 and U2 are subspaces of V. Prove that the intersection

U1 \ U2 is a subspace of V.

1. Prove that the intersection of every collection of subspaces of V is a subspace of V.
2. Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.
3. Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

[*This exercise is surprisingly harder than the previous exercise, possibly because this exercise is not true if we replace* **F** *with a ﬁeld containing only two elements.*]

1. Verify the assertion in Example 1.38.
2. Suppose U is a subspace of V. What is U C U ?
3. Is the operation of addition on the subspaces of V commutative? In other words, if U and W are subspaces of V, is U C W D W C U ?
4. Is the operation of addition on the subspaces of V associative? In other words, if U1; U2; U3 are subspaces of V, is

.U1 C U2/ C U3 D U1 C .U2 C U3/‹

1. Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?
2. Prove or give a counterexample: if U1; U2;W are subspaces of V such that

then U1 D U2.

1. Suppose

U1 C W D U2 C W;

U D f.x; x; y; y/ 2 **F**4 W x; y 2 **F**g:

Find a subspace W of **F**4 such that **F**4 D U ˚ W.

**26** CHAPTER 1 Vector Spaces

1. Suppose

U D f.x; y; x C y; x — y; 2x/ 2 **F**5 W x; y 2 **F**g:

Find a subspace W of **F**5 such that **F**5 D U ˚ W.

1. Suppose

U D f.x; y; x C y; x — y; 2x/ 2 **F**5 W x; y 2 **F**g:

Find three subspaces W1; W2; W3 of **F**5, none of which equals f0g, such that **F**5 D U ˚ W1 ˚ W2 ˚ W3.

1. Prove or give a counterexample: if U1; U2;W are subspaces of V such that

V D U1 ˚ W and V D U2 ˚ W;

then U1 D U2.

1. A function f W **R** ! **R** is called ***even*** if

f .—x/ D f .x/

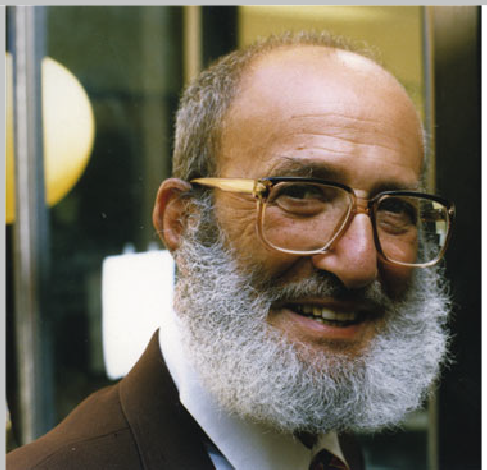
for all x 2 **R**. A function f W **R** ! **R** is called ***odd*** if

f .—x/ D —f .x/

for all x **R**. Let Ue denote the set of real-valued even functions on **R** and let Uo denote the set of real-valued odd functions on **R**. Show that **RR** D Ue ˚ Uo.

2

# *Finite-Dimensional* Vector Spaces



*American mathematician Paul Halmos* (*1916–2006*)*, who in 1942 published the ﬁrst modern linear algebra book. The title of Halmos’s book was the same as the title of this chapter.*

2

CHAPTER

Let’s review our standing assumptions:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a vector space over **F**.

In the last chapter we learned about vector spaces. Linear algebra focuses not on arbitrary vector spaces, but on ﬁnite-dimensional vector spaces, which we introduce in this chapter.

LEARNING OBJECTIVES FOR THIS CHAPTER

span

linear independence bases

dimension

© Springer International Publishing 2015 27

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

DOI 10.1007/978-3-319-11080-6 2

**28** CHAPTER 2 Finite-Dimensional Vector Spaces

## *Span and Linear Independence*

2.A

We have been writing lists of numbers surrounded by parentheses, and we will continue to do so for elements of **F**n; for example, .2; 7; 8/ **F**3. However, now we need to consider lists of vectors (which may be elements of **F**n or of

— 2

other vector spaces). To avoid confusion, we will usually write lists of vectors without surrounding parentheses. For example, .4; 1; 6/; .9; 5; 7/ is a list of length 2 of vectors in **R**3.

2.2 **Notation *list of vectors***

We will usually write lists of vectors without surrounding parentheses.

### Linear Combinations and Span

Adding up scalar multiples of vectors in a list gives what is called a linear combination of the list. Here is the formal deﬁnition:

2.3 **Deﬁnition *linear combination***

A ***linear combination*** of a list *v*1;:::; *v*m of vectors in V is a vector of the form

a1*v*1 C ... C am*v*m;

where a1;:::; am 2 **F**.

* 1. **Example** In **F**3,
     + .17; —4; 2/ is a linear combination of .2; 1; —3/; .1; —2; 4/ because

.17; —4; 2/ D 6.2; 1; —3/ C 5.1; —2; 4/:

* + - .17; —4; 5/ is not a linear combination of .2; 1; —3/; .1; —2; 4/ because there do not exist numbers a1; a2 2 **F** such that

.17; —4; 5/ D a1.2; 1; —3/ C a2.1; —2; 4/:

In other words, the system of equations

17 D 2a1 C a2

—4 D a1 — 2a2

5 D —3a1 C 4a2

has no solutions (as you should verify).

SECTION 2.A Span and Linear Independence **29**

2.5 **Deﬁnition *span***

The set of all linear combinations of a list of vectors *v*1;:::; *v*m in V is called the ***span*** of *v*1;:::; *v*m, denoted span.*v*1;:::; *v*m/. In other words,

span.*v*1;:::; *v*m/ D fa1*v*1 C ... C am*v*m W a1;:::; am 2 **F**g:

The span of the empty list ./ is deﬁned to be f0g.

* 1. **Example** The previous example shows that in **F**3,
     + .17; —4; 2/ 2 span .2; 1; —3/; .1; —2; 4/ ;

（ ）

（ ）

* + - .17; —4; 5/ … span .2; 1; —3/; .1; —2; 4/ .

Some mathematicians use the term ***linear span***, which means the same as span.

2.7 Span is the smallest containing subspace

The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.

Proof Suppose *v*1;:::; *v*m is a list of vectors in V.

First we show that span.*v*1;:::; *v*m/ is a subspace of V. The additive identity is in span.*v*1;:::; *v*m/, because

0 D 0*v*1 C ... C 0*v*m:

Also, span.*v*1;:::; *v*m/ is closed under addition, because

.a1*v*1C. . .Cam*v*m/C.c1*v*1C. . .Ccm*v*m/ D .a1Cc1/*v*1C. . .C.amCcm/*v*m: Furthermore, span.*v*1;:::; *v*m/ is closed under scalar multiplication, because入.a1*v*1 C ... C am*v*m/ D 入a1*v*1 C ... C 入am*v*m:

Thus span.*v*1;:::; *v*m/ is a subspace of V (by 1.34).

Each *v*j is a linear combination of *v*1;:::; *v*m (to show this, set aj 1

D

and let the other a’s in 2.3 equal 0). Thus span.*v*1;:::; *v*m/ contains each *v*j .

Conversely, because subspaces are closed under scalar multiplication and addition, every subspace of V containing each *v*j contains span.*v*1;:::; *v*m/. Thus span.*v*1;:::; *v*m/ is the smallest subspace of V containing all the vectors *v*1;:::; *v*m.

**30** CHAPTER 2 Finite-Dimensional Vector Spaces

2.8 **Deﬁnition *spans***

If span.*v*1;:::; *v*m/ equals V, we say that *v*1;:::; *v*m ***spans*** V.

2.9 **Example** Suppose n is a positive integer. Show that

.1; 0; : : : ; 0/; .0; 1; 0; : : : ; 0/; : : : ; .0; : : : ; 0; 1/

spans **F**n. Here the j th vector in the list above is the n-tuple with 1 in the j th

slot and 0 in all other slots.

Solution Suppose .x1;:::; xn/ 2 **F**n. Then

.x1;:::; xn/ D x1.1; 0; : : : ; 0/ C x2.0; 1; 0; : : : ; 0/ C ... C xn.0; : : : ; 0; 1/:

（ ）2

Thus .x1;:::; xn/ span .1; 0; : : : ; 0/; .0; 1; 0; : : : ; 0/; : : : ; .0; : : : ; 0; 1/ , as

desired.

Now we can make one of the key deﬁnitions in linear algebra.

2.10 **Deﬁnition *ﬁnite-dimensional vector space***

A vector space is called ***ﬁnite-dimensional*** if some list of vectors in it spans the space.

Example 2.9 above shows that **F**n

*Recall that by deﬁnition every list has ﬁnite length.*

is a ﬁnite-dimensional vector space for every positive integer n.

The deﬁnition of a polynomial is no doubt already familiar to you.

2.11 **Deﬁnition *polynomial,*** *P*.**F**/

* A function p W **F** ! **F** is called a ***polynomial*** with coefﬁcients in **F**

if there exist a0;:::; am 2 **F** such that

p.z/ D a0 C a1z C a2z C ... C amz

for all z 2 **F**.

2

m

* *P*.**F**/ is the set of all polynomials with coefﬁcients in **F**.

SECTION 2.A Span and Linear Independence **31**

With the usual operations of addition and scalar multiplication, .**F**/ is a vector space over **F**, as you should verify. In other words, .**F**/ is a subspace of **FF** , the vector space of functions from **F** to **F**.

*P*

*P*

If a polynomial (thought of as a function from **F** to **F**) is represented by two sets of coefﬁcients, then subtracting one representation of the polynomial from the other produces a polynomial that is identically zero as a function on **F** and hence has all zero coefﬁcients (if you are unfamiliar with this fact, just believe it for now; we will prove it later—see 4.7). **Conclusion:** the coefﬁcients of a polynomial are uniquely determined by the polynomial. Thus the next deﬁnition uniquely deﬁnes the degree of a polynomial.

2.12 **Deﬁnition *degree of a polynomial,*** deg p

* A polynomial p 2 *P*.**F**/ is said to have ***degree*** m if there exist

scalars a0; a1;:::; am 2 **F** with am ¤ 0 such that

p.z/ D a0 C a1z C ... C amz

m

for all z 2 **F**. If p has degree m, we write deg p D m.

* The polynomial that is identically 0 is said to have degree —1.

In the next deﬁnition, we use the convention that —1 < m, which means that the polynomial 0 is in *P*m.**F**/.

2.13 **Deﬁnition** *P*m.**F**/

with coefﬁcients in **F** and degree at most m.

For m a nonnegative integer, *P*m.**F**/ denotes the set of all polynomials

To verify the next example, note that m.**F**/ span.1; z; :::; zm/; here we are slightly abusing notation by letting zk denote a function.

*P*

*P* D

2.14 **Example** m.**F**/ is a ﬁnite-dimensional vector space for each non- negative integer m.

2.15 **Deﬁnition *inﬁnite-dimensional vector space***

A vector space is called ***inﬁnite-dimensional*** if it is not ﬁnite-dimensional.

**32** CHAPTER 2 Finite-Dimensional Vector Spaces

2.16 **Example** Show that *P*.**F**/ is inﬁnite-dimensional.

Solution Consider any list of elements of .**F**/. Let m denote the highest degree of the polynomials in this list. Then every polynomial in the span of this list has degree at most m. Thus zmC1 is not in the span of our list. Hence

*P*

no list spans *P*.**F**/. Thus *P*.**F**/ is inﬁnite-dimensional.

### Linear Independence

Suppose *v*1;:::; *v*m 2 V and *v* 2 span.*v*1;:::; *v*m/. By the deﬁnition of span, there exist a1;:::; am 2 **F** such that

*v* D a1*v*1 C ... C am*v*m:

Consider the question of whether the choice of scalars in the equation above is unique. Suppose c1;:::; cm is another set of scalars such that

*v* D c1*v*1 C ... C cm*v*m:

Subtracting the last two equations, we have

0 D .a1 — c1/*v*1 C ... C .am — cm/*v*m:

Thus we have written 0 as a linear combination of .*v*1;:::; *v*m/. If the only way to do this is the obvious way (using 0 for all scalars), then each aj cj equals 0, which means that each aj equals cj (and thus the choice of scalars

—

was indeed unique). This situation is so important that we give it a special name—linear independence—which we now deﬁne.

2.17 **Deﬁnition *linearly independent***

* A list *v*1;:::; *v*m of vectors in V is called ***linearly independent*** if

equal 0 is a1 D . . .D am D 0.

* The empty list ./ is also declared to be linearly independent.

the only choice of a1;:::; am 2 **F** that makes a1*v*1 C. . .C am*v*m

The reasoning above shows that *v*1;:::; *v*m is linearly independent if and only if each vector in span.*v*1;:::; *v*m/ has only one representation as a linear combination of *v*1;:::; *v*m.

SECTION 2.A Span and Linear Independence **33**

2.18 **Example *linearly independent lists***

1. A list *v* of one vector *v* 2 V is linearly independent if and only if *v* ¤ 0.
2. A list of two vectors in V is linearly independent if and only if neither vector is a scalar multiple of the other.

(c) .1; 0; 0; 0/; .0; 1; 0; 0/; .0; 0; 1; 0/ is linearly independent in **F**4.

(d) The list 1; z; :::; zm is linearly independent in .**F**/ for each nonnega- tive integer m.

*P*

If some vectors are removed from a linearly independent list, the remaining list is also linearly independent, as you should verify.

2.19 **Deﬁnition *linearly dependent***

* A list of vectors in V is called ***linearly dependent*** if it is not linearly

independent.

* In other words, a list *v*1;:::; *v*m of vectors in V is linearly de-

a1*v*1 C ... C am*v*m D 0.

pendent if there exist a1;:::; am 2 **F**, not all 0, such that

2.20 **Example *linearly dependent lists***

* .2; 3; 1/; .1; —1; 2/; .7; 3; 8/ is linearly dependent in **F**3 because

2.2; 3; 1/ C 3.1; —1; 2/ C .—1/.7; 3; 8/ D .0; 0; 0/:

* The list .2; 3; 1/; .1; —1; 2/; .7; 3; c/ is linearly dependent in **F**3 if and only if c D 8, as you should verify.

If some vector in a list of vectors in V is a linear combination of the other vectors, then the list is linearly dependent. (Proof: After writing one vector in the list as equal to a linear combination of the other

•

vectors, move that vector to the other side of the equation, where it will be multiplied by —1.)

Every list of vectors in V containing the 0 vector is linearly dependent. (This is a special case of the previous bullet point.)

•

**34** CHAPTER 2 Finite-Dimensional Vector Spaces

The lemma below will often be useful. It states that given a linearly dependent list of vectors, one of the vectors is in the span of the previous ones and furthermore we can throw out that vector without changing the span of the original list.

2.21 Linear Dependence Lemma

Suppose *v*1;:::; *v*m is a linearly dependent list in V. Then there exists

j 2 f1; 2; :::; mg such that the following hold:

(a) *v*j 2 span.*v*1;:::; *v*j 1/;

(b) if the j th term is removed from *v*1;:::; *v*m, the span of the remain- ing list equals span.*v*1;:::; *v*m/.

Proof Because the list *v*1;:::; *v*m is linearly dependent, there exist numbers

a1;:::; am 2 **F**, not all 0, such that

a1*v*1 C ... C am*v*m D 0:

Let j be the largest element of f1; :::; mg such that aj ¤ 0. Then

**2.22** *v* a1 *v* — ... — aj 1 *v* ;

proving (a).

j D — aj 1

aj j 1

To prove (b), suppose u 2 span.*v*1;:::; *v*m/. Then there exist numbers

c1;:::; cm 2 **F** such that

u D c1*v*1 C ... C cm*v*m:

In the equation above, we can replace *v*j with the right side of 2.22, which shows that u is in the span of the list obtained by removing the j th term from *v*1;:::; *v*m. Thus (b) holds.

Choosing j 1 in the Linear Dependence Lemma above means that *v*1 0, because if j 1 then condition (a) above is interpreted to mean that *v*1 span. /; recall that span./ 0 . Note also that the proof of part (b) above needs to be modiﬁed in an obvious way if *v*1 0 and j 1.

D D

2 D f g

D D

D

In general, the proofs in the rest of the book will not call attention to special cases that must be considered involving empty lists, lists of length 1, the subspace 0 , or other trivial cases for which the result is clearly true but

f g

needs a slightly different proof. Be sure to check these special cases yourself.

Now we come to a key result. It says that no linearly independent list in V

is longer than a spanning list in V.

SECTION 2.A Span and Linear Independence **35**

2.23 Length of linearly independent list 三 length of spanning list In a ﬁnite-dimensional vector space, the length of every linearly indepen- dent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof Suppose u1;:::; um is linearly independent in V. Suppose also that *w*1;:::; *w*n spans V. We need to prove that m n. We do so through the multi-step process described below; note that in each step we add one of the

u’s and remove one of the *w*’s.

三

#### Step 1

Let B be the list *w*1;:::; *w*n, which spans V. Thus adjoining any vector in V to this list produces a linearly dependent list (because the newly adjoined vector can be written as a linear combination of the other

vectors). In particular, the list

u1; *w*1;:::; *w*n

is linearly dependent. Thus by the Linear Dependence Lemma (2.21), we can remove one of the *w*’s so that the new list B (of length n) consisting of u1 and the remaining *w*’s spans V.

#### Step j

The list B (of length n) from step j 1 spans V. Thus adjoining any vector to this list produces a linearly dependent list. In particular, the list of length .n C 1/ obtained by adjoining uj to B, placing it just after

—

u1;:::; uj 1, is linearly dependent. By the Linear Dependence Lemma

(2.21), one of the vectors in this list is in the span of the previous ones, and because u1;:::; uj is linearly independent, this vector is one of the *w*’s, not one of the u’s. We can remove that *w* from B so that the new list B (of length n) consisting of u1;:::; uj and the remaining *w*’s spans V.

After step m, we have added all the u’s and the process stops. At each step as we add a u to B, the Linear Dependence Lemma implies that there is some *w* to remove. Thus there are at least as many *w*’s as u’s.

The next two examples show how the result above can be used to show, without any computations, that certain lists are not linearly independent and that certain lists do not span a given vector space.

**36** CHAPTER 2 Finite-Dimensional Vector Spaces

2.24 **Example** Show that the list .1; 2; 3/; .4; 5; 8/; .9; 6; 7/; . 3; 2; 8/ is not linearly independent in **R**3.

Solution The list .1; 0; 0/; .0; 1; 0/; .0; 0; 1/ spans **R**3. Thus no list of length larger than 3 is linearly independent in **R**3.

—

2.25 **Example** Show that the list .1; 2; 3; 5/; .4; 5; 8; 3/; .9; 6; 7; 1/

does not span **R**4.

— —

Solution The list .1; 0; 0; 0/; .0; 1; 0; 0/; .0; 0; 1; 0/; .0; 0; 0; 1/ is linearly in- dependent in **R**4. Thus no list of length less than 4 spans **R**4.

Our intuition suggests that every subspace of a ﬁnite-dimensional vector space should also be ﬁnite-dimensional. We now prove that this intuition is correct.

2.26 Finite-dimensional subspaces

Every subspace of a ﬁnite-dimensional vector space is ﬁnite-dimensional.

Proof Suppose V is ﬁnite-dimensional and U is a subspace of V. We need to prove that U is ﬁnite-dimensional. We do this through the following multi-step construction.

#### Step 1

If U D f0g, then U is ﬁnite-dimensional and we are done. If U ¤ f0g, then choose a nonzero vector *v*1 2 U.

#### Step j

If U D span.*v*1;:::; *v*j 1/, then U is ﬁnite-dimensional and we are done. If U span.*v*1;:::; *v*j 1/, then choose a vector *v*j U such that

¤ 2

*v*j … span.*v*1;:::; *v*j 1/:

After each step, as long as the process continues, we have constructed a list of vectors such that no vector in this list is in the span of the previous vectors. Thus after each step we have constructed a linearly independent list, by the Linear Dependence Lemma (2.21). This linearly independent list cannot be

longer than any spanning list of V (by 2.23). Thus the process eventually terminates, which means that U is ﬁnite-dimensional.

EXERCISES 2.A

SECTION 2.A Span and Linear Independence **37**

1. Suppose *v*1; *v*2; *v*3; *v*4 spans V. Prove that the list

*v*1 — *v*2; *v*2 — *v*3; *v*3 — *v*4; *v*4

also spans V.

1. Verify the assertions in Example 2.18.
2. Find a number t such that

.3; 1; 4/; .2; —3; 5/; .5; 9; t/

is not linearly independent in **R**3.

1. Verify the assertion in the second bullet point in Example 2.20.
2. (a) Show that if we think of **C** as a vector space over **R**, then the list

.1 C i; 1 — i/ is linearly independent.

(b) Show that if we think of **C** as a vector space over **C**, then the list

.1 C i; 1 — i/ is linearly dependent.

1. Suppose *v*1; *v*2; *v*3; *v*4 is linearly independent in V. Prove that the list

*v*1 — *v*2; *v*2 — *v*3; *v*3 — *v*4; *v*4

is also linearly independent.

1. Prove or give a counterexample: If *v*1; *v*2;:::; *v*m is a linearly indepen- dent list of vectors in V, then

5*v*1 — 4*v*2; *v*2; *v*3;:::; *v*m

is linearly independent.

1. Prove or give a counterexample: If *v*1; *v*2;:::; *v*m is a linearly indepen- dent list of vectors in V and 入 **F** with 入 0, then 入*v*1; 入*v*2;:::; 入*v*m is linearly independent.

2 ¤

1. Prove or give a counterexample: If *v*1;:::; *v*m and *w*1;:::; *w*m are lin- early independent lists of vectors in V, then *v*1 *w*1;:::; *v*m *w*m is linearly independent.

C C

1. Suppose *v*1;:::; *v*m is linearly independent in V and *w* 2 V. Prove that if *v*1 C *w*;:::; *v*m C *w* is linearly dependent, then *w* 2 span.*v*1;:::; *v*m/.

**38** CHAPTER 2 Finite-Dimensional Vector Spaces

1. Suppose *v*1;:::; *v*m is linearly independent in V and *w* V. Show that

2

*v*1;:::; *v*m; *w* is linearly independent if and only if

*w* … span.*v*1;:::; *v*m/:

1. Explain why there does not exist a list of six polynomials that is linearly independent in *P*4.**F**/.
2. Explain why no list of four polynomials spans *P*4.**F**/.
3. Prove that V is inﬁnite-dimensional if and only if there is a sequence *v*1; *v*2;::: of vectors in V such that *v*1;:::; *v*m is linearly independent for every positive integer m.
4. Prove that **F**1 is inﬁnite-dimensional.
5. Prove that the real vector space of all continuous real-valued functions on the interval Œ0; 1] is inﬁnite-dimensional.
6. Suppose p0; p1;:::; pm are polynomials in m.**F**/ such that pj .2/ 0 for each j . Prove that p0; p1;:::; pm is not linearly independent in *P*m.**F**/.

*P* D

## *Bases*

2.B

SECTION 2.B Bases **39**

In the last section, we discussed linearly independent lists and spanning lists. Now we bring these concepts together.

2.27 **Deﬁnition *basis***

A ***basis*** of V is a list of vectors in V that is linearly independent and spans V.

2.28 **Example *bases***

(a) The list .1; 0; : : : ; 0/; .0; 1; 0; : : : ; 0/; : : : ; .0; : : : ; 0; 1/ is a basis of **F**n, called the ***standard basis*** of **F**n.

1. The list .1; 2/; .3; 5/ is a basis of **F**2.
2. The list .1; 2; 4/; .7; 5; 6/ is linearly independent in **F**3 but is not a basis of **F**3 because it does not span **F**3.

— —

1. The list .1; 2/; .3; 5/; .4; 13/ spans **F**2 but is not a basis of **F**2 because it is not linearly independent.
2. The list .1; 1; 0/; .0; 0; 1/ is a basis of f.x; x; y/ 2 **F**3 W x; y 2 **F**g.
3. The list .1; —1; 0/; .1; 0; —1/ is a basis of

3

f.x; y; z/ 2 **F** W x C y C z D 0g:

1. The list 1; z; :::; zm is a basis of *P*m.**F**/.

In addition to the standard basis, **F**n has many other bases. For example,

.7; 5/; . 4; 9/ and .1; 2/; .3; 5/ are both bases of **F**2.

—

The next result helps explain why bases are useful. Recall that “uniquely”

means “in only one way”.

2.29 Criterion for basis

A list *v*1;:::; *v*n of vectors in V is a basis of V if and only if every *v* 2 V

can be written uniquely in the form

**2.30** *v* D a1*v*1 C ... C an*v*n;

where a1;:::; an 2 **F**.

**40** CHAPTER 2 Finite-Dimensional Vector Spaces

Proof First suppose that *v*1;:::; *v*n is a basis of V. Let *v* V. Because

2

*v*1;:::; *v*n spans V, there exist a1;:::; an **F** such that 2.30 holds. To

2

show that the representation in 2.30 is unique, suppose c1;:::; cn are scalars such that we also have

*This proof is essentially a repeti- tion of the ideas that led us to the deﬁnition of linear independence.*

*v* D c1*v*1 C ... C cn*v*n:

Subtracting the last equation from 2.30, we get

0 D .a1 — c1/*v*1 C ... C .an — cn/*v*n:

This implies that each aj cj equals 0 (because *v*1;:::; *v*n is linearly inde- pendent). Hence a1 c1;:::; an cn. We have the desired uniqueness, completing the proof in one direction.

D D

—

For the other direction, suppose every *v* V can be written uniquely in the form given by 2.30. Clearly this implies that *v*1;:::; *v*n spans V. To show that *v*1;:::; *v*n is linearly independent, suppose a1;:::; an 2 **F** are such that

2

0 D a1*v*1 C ... C an*v*n:

The uniqueness of the representation 2.30 (taking *v* 0) now implies that a1 an 0. Thus *v*1;:::; *v*n is linearly independent and hence is a basis of V.

D . . .D D

D

A spanning list in a vector space may not be a basis because it is not linearly independent. Our next result says that given any spanning list, some (possibly none) of the vectors in it can be discarded so that the remaining list is linearly independent and still spans the vector space.

As an example in the vector space **F**2, if the procedure in the proof below is applied to the list .1; 2/; .3; 6/; .4; 7/; .5; 9/, then the second and fourth vectors will be removed. This leaves .1; 2/; .4; 7/, which is a basis of **F**2.

2.31 Spanning list contains a basis

Every spanning list in a vector space can be reduced to a basis of the vector space.

Proof Suppose *v*1;:::; *v*n spans V. We want to remove some of the vectors from *v*1;:::; *v*n so that the remaining vectors form a basis of V. We do this through the multi-step process described below.

Start with B equal to the list *v*1;:::; *v*n.

SECTION 2.B Bases **41**

#### Step 1

If *v*1 D 0, delete *v*1 from B. If *v*1 ¤ 0, leave B unchanged.

#### Step j

If *v*j is in span.*v*1;:::; *v*j 1/, delete *v*j from B. If *v*j is not in span.*v*1;:::; *v*j 1/, leave B unchanged.

Stop the process after step n, getting a list B. This list B spans V because our original list spanned V and we have discarded only vectors that were already in the span of the previous vectors. The process ensures that no vector in B is in the span of the previous ones. Thus B is linearly independent, by the Linear Dependence Lemma (2.21). Hence B is a basis of V.

Our next result, an easy corollary of the previous result, tells us that every ﬁnite-dimensional vector space has a basis.

2.32 Basis of ﬁnite-dimensional vector space

Every ﬁnite-dimensional vector space has a basis.

Proof By deﬁnition, a ﬁnite-dimensional vector space has a spanning list. The previous result tells us that each spanning list can be reduced to a basis.

Our next result is in some sense a dual of 2.31, which said that every spanning list can be reduced to a basis. Now we show that given any linearly independent list, we can adjoin some additional vectors (this includes the possibility of adjoining no additional vectors) so that the extended list is still linearly independent but also spans the space.

2.33 Linearly independent list extends to a basis

Every linearly independent list of vectors in a ﬁnite-dimensional vector space can be extended to a basis of the vector space.

Proof Suppose u1;:::; um is linearly independent in a ﬁnite-dimensional vector space V. Let *w*1;:::; *w*n be a basis of V. Thus the list

u1;:::; um; *w*1;:::; *w*n

spans V. Applying the procedure of the proof of 2.31 to reduce this list to a basis of V produces a basis consisting of the vectors u1;:::; um (none of the u’s get deleted in this procedure because u1;:::; um is linearly independent)

and some of the *w*’s.

**42** CHAPTER 2 Finite-Dimensional Vector Spaces

As an example in **F**3, suppose we start with the linearly independent list .2; 3; 4/; .9; 6; 8/. If we take *w*1; *w*2; *w*3 in the proof above to be the standard basis of **F**3, then the procedure in the proof above produces the list

.2; 3; 4/; .9; 6; 8/; .0; 1; 0/, which is a basis of **F**3.

As an application of the result above,

*Using the same basic ideas but considerably more advanced tools, the next result can be proved with-*

*out the hypothesis that* V *is ﬁnite-*

*dimensional.*

we now show that every subspace of a ﬁnite-dimensional vector space can be paired with another subspace to form a direct sum of the whole space.

2.34 Every subspace of V is part of a direct sum equal to V

Suppose V is ﬁnite-dimensional and U is a subspace of V. Then there is a subspace W of V such that V D U ˚ W.

Proof Because V is ﬁnite-dimensional, so is U (see 2.26). Thus there is a basis u1;:::; um of U (see 2.32). Of course u1;:::; um is a linearly in- dependent list of vectors in V. Hence this list can be extended to a basis u1;:::; um; *w*1;:::; *w*n of V (see 2.33). Let W D span.*w*1;:::; *w*n/.

To prove that V D U ˚ W, by 1.45 we need only show that

V D U C W and U \ W D f0g:

To prove the ﬁrst equation above, suppose *v* V. Then, because the list u1;:::; um; *w*1;:::; *w*n spans V, there exist a1;:::; am; b1;:::; bn **F** such that

2

2

*v* D „a1u1 C .ƒ. ‚. C amum… C b„1*w*1 C ƒ. .‚. C bn*w*…n :

*w*

u

In other words, we have *v* D u C *w*, where u 2 U and *w* 2 W are deﬁned as above. Thus *v* 2 U C W, completing the proof that V D U C W.

To show that U \ W D f0g, suppose *v* 2 U \ W. Then there exist scalars

a1;:::; am; b1;:::; bn 2 **F** such that

*v* D a1u1 C ... C amum D b1*w*1 C ... C bn*w*n:

Thus

a1u1 C ... C amum — b1*w*1 — ... — bn*w*n D 0:

Because u1;:::; um; *w*1;:::; *w*n is linearly independent, this implies that a1 D . . .D am D b1 D . . .D bn D 0. Thus *v* D 0, completing the proof that U \ W D f0g.

EXERCISES 2.B

SECTION 2.B Bases **43**

1. Find all vector spaces that have exactly one basis.
2. Verify all the assertions in Example 2.28.
3. (a) Let U be the subspace of **R**5 deﬁned by

U D f.x1; x2; x3; x4; x5/ 2 **R**5 W x1 D 3x2 and x3 D 7x4g:

Find a basis of U.

* 1. Extend the basis in part (a) to a basis of **R**5.
  2. Find a subspace W of **R**5 such that **R**5 D U ˚ W.

1. (a) Let U be the subspace of **C**5 deﬁned by

U D f.z1; z2; z3; z4; z5/ 2 **C**5 W 6z1 D z2 and z3 C2z4 C3z5 D 0g:

Find a basis of U.

1. Extend the basis in part (a) to a basis of **C**5.
2. Find a subspace W of **C**5 such that **C**5 D U ˚ W.
3. Prove or disprove: there exists a basis p0; p1; p2; p3 of 3.**F**/ such that none of the polynomials p0; p1; p2; p3 has degree 2.

*P*

1. Suppose *v*1; *v*2; *v*3; *v*4 is a basis of V. Prove that

*v*1 C *v*2; *v*2 C *v*3; *v*3 C *v*4; *v*4

is also a basis of V.

1. Prove or give a counterexample: If *v*1; *v*2; *v*3; *v*4 is a basis of V and U is a subspace of V such that *v*1; *v*2 U and *v*3 U and *v*4 U, then *v*1; *v*2 is a basis of U.

2 … …

1. Suppose U and W are subspaces of V such that V U W. Suppose also that u1;:::; um is a basis of U and *w*1;:::; *w*n is a basis of W. Prove that

D ˚

u1;:::; um; *w*1;:::; *w*n

is a basis of V.

**44** CHAPTER 2 Finite-Dimensional Vector Spaces

## *Dimension*

2.C

Although we have been discussing ﬁnite-dimensional vector spaces, we have not yet deﬁned the dimension of such an object. How should dimension be

deﬁned? A reasonable deﬁnition should force the dimension of **F**n to equal n.

Notice that the standard basis

.1; 0; : : : ; 0/; .0; 1; 0; : : : ; 0/; : : : ; .0; : : : ; 0; 1/

of **F**n has length n. Thus we are tempted to deﬁne the dimension as the length of a basis. However, a ﬁnite-dimensional vector space in general has many different bases, and our attempted deﬁnition makes sense only if all bases in a

given vector space have the same length. Fortunately that turns out to be the case, as we now show.

2.35 Basis length does not depend on basis

Any two bases of a ﬁnite-dimensional vector space have the same length.

Proof Suppose V is ﬁnite-dimensional. Let B1 and B2 be two bases of V. Then B1 is linearly independent in V and B2 spans V, so the length of B1 is at most the length of B2 (by 2.23). Interchanging the roles of B1 and B2, we also see that the length of B2 is at most the length of B1. Thus the length of B1 equals the length of B2, as desired.

Now that we know that any two bases of a ﬁnite-dimensional vector space have the same length, we can formally deﬁne the dimension of such spaces.

2.36 **Deﬁnition *dimension,*** dim V

any basis of the vector space.

* The dimension of V (if V is ﬁnite-dimensional) is denoted by dim V.
* The ***dimension*** of a ﬁnite-dimensional vector space is the length of

2.37 **Example *dimensions***

* dim **F**n D n because the standard basis of **F**n has length n.
* dim *P*m.**F**/ D m C 1 because the basis 1; z; :::; zm of *P*m.**F**/ has length m C 1.

SECTION 2.C Dimension **45**

Every subspace of a ﬁnite-dimensional vector space is ﬁnite-dimensional (by 2.26) and so has a dimension. The next result gives the expected inequality about the dimension of a subspace.

2.38 Dimension of a subspace

If V is ﬁnite-dimensional and U is a subspace of V, then dim U 三 dim V.

Proof Suppose V is ﬁnite-dimensional and U is a subspace of V. Think of a basis of U as a linearly independent list in V, and think of a basis of V as a spanning list in V. Now use 2.23 to conclude that dim U 三 dim V.

To check that a list of vectors in V

*The real vector space* **R**2 *has di- mension* 2*; the complex vector space* **C** *has dimension* 1*. As sets,* **R**2 *can be identiﬁed with* **C**

*(and addition is the same on both spaces, as is scalar multiplication by real numbers). Thus when we talk about the dimension of a vec- tor space, the role played by the choice of* **F** *cannot be neglected.*

is a basis of V, we must, according to

the deﬁnition, show that the list in ques-

tion satisﬁes two properties: it must be linearly independent and it must span

V. The next two results show that if the

list in question has the right length, then

we need only check that it satisﬁes one of the two required properties. First we prove that every linearly independent list with the right length is a basis.

D

2.39 Linearly independent list of the right length is a basis

Suppose V is ﬁnite-dimensional. Then every linearly independent list of vectors in V with length dim V is a basis of V.

Proof Suppose dim V n and *v*1;:::; *v*n is linearly independent in V. The list *v*1;:::; *v*n can be extended to a basis of V (by 2.33). However, every basis of V has length n, so in this case the extension is the trivial one, meaning that no elements are adjoined to *v*1;:::; *v*n. In other words, *v*1;:::; *v*n is a basis of V, as desired.

* 1. **Example** Show that the list .5; 7/; .4; 3/ is a basis of **F**2.

Solution This list of two vectors in **F**2 is obviously linearly independent (because neither vector is a scalar multiple of the other). Note that **F**2 has dimension 2. Thus 2.39 implies that the linearly independent list .5; 7/; .4; 3/

of length 2 is a basis of **F**2 (we do not need to bother checking that it spans **F**2).

**46** CHAPTER 2 Finite-Dimensional Vector Spaces

* 1. **Example** Show that 1; .x — 5/2; .x — 5/3 is a basis of the subspace

U of *P*3.**R**/ deﬁned by

U D fp 2 *P*3.**R**/ W p0.5/ D 0g:

Solution Clearly each of the polynomials 1, .x — 5/2, and .x — 5/3 is in U.

Suppose a; b; c 2 **R** and

a C b.x — 5/2 C c.x — 5/3 D 0

for every x **R**. Without explicitly expanding the left side of the equation above, we can see that the left side has a cx3 term. Because the right side has no x3 term, this implies that c D 0. Because c D 0, we see that the left side has a bx2 term, which implies that b D 0. Because b D c D 0, we can also conclude that a D 0.

2

Thus the equation above implies that a D b D c D 0. Hence the list

1; .x — 5/2; .x — 5/3 is linearly independent in U.

Thus dim U 乏 3. Because U is a subspace of *P*3.**R**/, we know that dim U 三 dim *P*3.**R**/ D 4 (by 2.38). However, dim U cannot equal 4, because otherwise when we extend a basis of U to a basis of *P*3.**R**/ we would get a list with length greater than 4. Hence dim U D 3. Thus 2.39 implies that the linearly independent list 1; .x — 5/2; .x — 5/3 is a basis of U.

Now we prove that a spanning list with the right length is a basis.

2.42 Spanning list of the right length is a basis

Suppose V is ﬁnite-dimensional. Then every spanning list of vectors in V

with length dim V is a basis of V.

Proof Suppose dim V n and *v*1;:::; *v*n spans V. The list *v*1;:::; *v*n can be reduced to a basis of V (by 2.31). However, every basis of V has length n, so in this case the reduction is the trivial one, meaning that no elements are deleted from *v*1;:::; *v*n. In other words, *v*1;:::; *v*n is a basis of V, as

desired.

D

The next result gives a formula for the dimension of the sum of two subspaces of a ﬁnite-dimensional vector space. This formula is analogous to a familiar counting formula: the number of elements in the union of two ﬁnite sets equals the number of elements in the ﬁrst set, plus the number of elements in the second set, minus the number of elements in the intersection of the two sets.

SECTION 2.C Dimension **47**

2.43 Dimension of a sum

If U1 and U2 are subspaces of a ﬁnite-dimensional vector space, then dim.U1 C U2/ D dim U1 C dim U2 — dim.U1 \ U2/:

Proof Let u1;:::; um be a basis of U1 U2; thus dim.U1 U2/ m. Be- cause u1;:::; um is a basis of U1 U2, it is linearly independent in U1. Hence this list can be extended to a basis u1;:::; um; *v*1;:::; *v*j of U1 (by 2.33). Thus dim U1 m j . Also extend u1;:::; um to a basis u1;:::; um; *w*1;:::; *w*k of U2; thus dim U2 m k.

We will show that

\

D C

D C

\ \ D

u1;:::; um; *v*1;:::; *v*j ; *w*1;:::; *w*k

is a basis of U1 CU2. This will complete the proof, because then we will have dim.U1 C U2/ D m C j C k

D .m C j/ C .m C k/ — m

D dim U1 C dim U2 — dim.U1 \ U2/:

Clearly span.u1;:::; um; *v*1;:::; *v*j ; *w*1;:::; *w*k/ contains U1 and U2 and hence equals U1 U2. So to show that this list is a basis of U1 U2 we need only show that it is linearly independent. To prove this, suppose

C C

a1u1 C ... C amum C b1*v*1 C ... C bj *v*j C c1*w*1 C ... C ck*w*k D 0;

where all the a’s, b’s, and c’s are scalars. We need to prove that all the a’s, b’s, and c’s equal 0. The equation above can be rewritten as

c1*w*1 C ... C ck*w*k D —a1u1 — ... — amum — b1*v*1 — ... — bj *v*j ;

which shows that c1*w*1 C . . .C ck*w*k 2 U1. All the *w*’s are in U2, so this implies that c1*w*1 C ... C ck*w*k 2 U1 \ U2. Because u1;:::; um is a basis of U1 \ U2, we can write

c1*w*1 C ... C ck*w*k D d1u1 C ... C dmum

for some choice of scalars d1;:::; dm. But u1;:::; um; *w*1;:::; *w*k is linearly independent, so the last equation implies that all the c’s (and d’s) equal 0.

Thus our original equation involving the a’s, b’s, and c’s becomes

a1u1 C ... C amum C b1*v*1 C ... C bj *v*j D 0:

Because the list u1;:::; um; *v*1;:::; *v*j is linearly independent, this equation implies that all the a’s and b’s are 0. We now know that all the a’s, b’s, and c’s equal 0, as desired.

**48** CHAPTER 2 Finite-Dimensional Vector Spaces

EXERCISES 2.C

1. Suppose V is ﬁnite-dimensional and U is a subspace of V such that dim U D dim V. Prove that U D V.
2. Show that the subspaces of **R**2 are precisely 0 , **R**2, and all lines in **R**2 through the origin.

f g

1. Show that the subspaces of **R**3 are precisely 0 , **R**3, all lines in **R**3

f g

through the origin, and all planes in **R**3 through the origin.

1. (a) Let U D fp 2 *P*4.**F**/ W p.6/ D 0g. Find a basis of U.
2. Extend the basis in part (a) to a basis of *P*4.**F**/.
3. Find a subspace W of *P*4.**F**/ such that *P*4.**F**/ D U ˚ W.
4. (a) Let U D fp 2 *P*4.**R**/ W p00.6/ D 0g. Find a basis of U.
5. Extend the basis in part (a) to a basis of *P*4.**R**/.
6. Find a subspace W of *P*4.**R**/ such that *P*4.**R**/ D U ˚ W.
7. (a) Let U D fp 2 *P*4.**F**/ W p.2/ D p.5/g. Find a basis of U.
8. Extend the basis in part (a) to a basis of *P*4.**F**/.
9. Find a subspace W of *P*4.**F**/ such that *P*4.**F**/ D U ˚ W.
10. (a) Let U D fp 2 *P*4.**F**/ W p.2/ D p.5/ D p.6/g. Find a basis of U.
11. Extend the basis in part (a) to a basis of *P*4.**F**/.
12. Find a subspace W of *P*4.**F**/ such that *P*4.**F**/ D U ˚ W.
13. (a) Let U D fp 2 *P*4.**R**/ W R 1 p D 0g. Find a basis of U.

1

1. Extend the basis in part (a) to a basis of *P*4.**R**/.
2. Find a subspace W of *P*4.**R**/ such that *P*4.**R**/ D U ˚ W.
3. Suppose *v*1;:::; *v*m is linearly independent in V and *w* 2 V. Prove that dim span.*v*1 C *w*;:::; *v*m C *w*/ 乏 m — 1:
4. Suppose p0; p1;:::; pm 2 *P*.**F**/ are such that each pj has degree j . Prove that p0; p1;:::; pm is a basis of *P*m.**F**/.
5. Suppose that U and W are subspaces of **R**8 such that dim U D 3, dim W D 5, and U C W D **R**8. Prove that **R**8 D U ˚ W.

SECTION 2.C Dimension **49**

1. Suppose U and W are both ﬁve-dimensional subspaces of **R**9. Prove that U \ W ¤ f0g.
2. Suppose U and W are both 4-dimensional subspaces of **C**6. Prove that there exist two vectors in U W such that neither of these vectors is a scalar multiple of the other.

\

1. Suppose U1;:::; Um are ﬁnite-dimensional subspaces of V. Prove that

U1 C C Um is ﬁnite-dimensional and

dim.U1 C ... C Um/ 三 dim U1 C C dim Um:

1. Suppose V is ﬁnite-dimensional, with dim V n 1. Prove that there exist 1-dimensional subspaces U1;:::; Un of V such that

D 乏

V D U1 ˚ ˚ Un:

1. Suppose U1;:::; Um are ﬁnite-dimensional subspaces of V such that U1 Um is a direct sum. Prove that U1 Um is ﬁnite- dimensional and

C . . . C ˚ . . . ˚

dim U1 ˚ ... ˚ Um D dim U1 C C dim Um:

[*The exercise above deepens the analogy between direct sums of sub- spaces and disjoint unions of subsets. Speciﬁcally, compare this exercise to the following obvious statement: if a set is written as a disjoint union of ﬁnite subsets, then the number of elements in the set equals the sum of the numbers of elements in the disjoint subsets.*]

1. You might guess, by analogy with the formula for the number of ele- ments in the union of three subsets of a ﬁnite set, that if U1; U2; U3 are subspaces of a ﬁnite-dimensional vector space, then

dim.U1 C U2 C U3/

D dim U1 C dim U2 C dim U3

— dim.U1 \ U2/ — dim.U1 \ U3/ — dim.U2 \ U3/

C dim.U1 \ U2 \ U3/:

Prove this or give a counterexample.

# *Linear Maps*



*German mathematician Carl Friedrich Gauss* (*1777–1855*)*, who in 1809 published a method for solving systems of linear equations. This method, now called Gaussian elimination, was also used in a Chinese book published over 1600 years earlier.*

3

CHAPTER

So far our attention has focused on vector spaces. No one gets excited about vector spaces. The interesting part of linear algebra is the subject to which we now turn—linear maps.

In this chapter we will frequently need another vector space, which we will call W, in addition to V. Thus our standing assumptions are now as follows:

* 1. **Notation F*,*** V***,*** W
     + **F** denotes **R** or **C**.
     + V and W denote vector spaces over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

Fundamental Theorem of Linear Maps

the matrix of a linear map with respect to given bases isomorphic vector spaces

product spaces quotient spaces

the dual space of a vector space and the dual of a linear map

© Springer International Publishing 2015 51

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

DOI 10.1007/978-3-319-11080-6 3

**52** CHAPTER 3 Linear Maps

*Some mathematicians use the term* ***linear transformation****, which means the same as linear map.*

## *The Vector Space of Linear Maps*

3.A

### Deﬁnition and Examples of Linear Maps

Now we are ready for one of the key deﬁnitions in linear algebra.

3.2 **Deﬁnition *linear map***

A ***linear map*** from V to W is a function T W V ! W with the following

properties:

**additivity**

T .u C *v*/ D Tu C T *v* for all u; *v* 2 V ;

**homogeneity**

T .入*v*/ D 入.T *v*/ for all 入 2 **F** and all *v* 2 V.

Note that for linear maps we often use the notation T *v* as well as the more standard functional notation T .*v*/.

3.3 **Notation** *L*.V; W /

The set of all linear maps from V to W is denoted *L*.V; W /.

Let’s look at some examples of linear maps. Make sure you verify that each of the functions deﬁned below is indeed a linear map:

3.4 **Example *linear maps***

#### zero

In addition to its other uses, we let the symbol 0 denote the function that takes each element of some vector space to the additive identity of another vector space. To be speciﬁc, 0 2 *L*.V; W / is deﬁned by

0*v* D 0:

The 0 on the left side of the equation above is a function from V to W, whereas the 0 on the right side is the additive identity in W. As usual, the context should allow you to distinguish between the many uses of the symbol 0.

#### identity

The ***identity map***, denoted I, is the function on some vector space that takes each element to itself. To be speciﬁc, I 2 *L*.V; V / is deﬁned by

I *v* D *v*:

SECTION 3.A The Vector Space of Linear Maps **53**

#### differentiation

（ ）

Deﬁne D 2 *L P*.**R**/; *P*.**R**/ by

Dp D p0:

The assertion that this function is a linear map is another way of stating a basic result about differentiation: .f g/0 f 0 g0 and .入f /0 入f 0 whenever f; g are differentiable and 入is a constant.

C D C D

#### integration

Deﬁne T 2 *L*（*P*.**R**/; **R**） by Z

1

Tp D

p.x/ dx:

0

The assertion that this function is linear is another way of stating a basic result about integration: the integral of the sum of two functions equals the sum of the integrals, and the integral of a constant times a function equals the constant times the integral of the function.

#### multiplication by x2

（ ）

Deﬁne T 2 *L P*.**R**/; *P*.**R**/ by

.Tp/.x/ D x2p.x/

for x 2 **R**.

#### backward shift

Recall that **F**1 denotes the vector space of all sequences of elements of **F**. Deﬁne T 2 *L*.**F**1; **F**1/ by

T .x1; x2; x3;:::/ D .x2; x3;::: /:

**from R**3 **to R**2

Deﬁne T 2 *L*.**R**3; **R**2/ by

T .x; y; z/ D .2x — y C 3z; 7x C 5y — 6z/:

**from F**n **to F**m

Generalizing the previous example, let m and n be positive integers, let Aj;k 2 **F** for j D 1; :::;m and k D 1; :::; n, and deﬁne T 2 *L*.**F**n; **F**m/ by T .x1;:::; xn/ D .A1;1x1 C ... C A1;nxn;:::; Am;1x1 C ... C Am;nxn/:

Actually every linear map from **F**n to **F**m is of this form.

The existence part of the next result means that we can ﬁnd a linear map that takes on whatever values we wish on the vectors in a basis. The uniqueness part of the next result means that a linear map is completely determined by its values on a basis.

**54** CHAPTER 3 Linear Maps

3.5 Linear maps and basis of domain

Suppose *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*n 2 W. Then there exists

a unique linear map T W V ! W such that

T *v*j D *w*j

for each j D 1; :::; n.

Proof First we show the existence of a linear map T with the desired property. Deﬁne T W V ! W by

T .c1*v*1 C ... C cn*v*n/ D c1*w*1 C ... C cn*w*n;

where c1;:::; cn are arbitrary elements of **F**. The list *v*1;:::; *v*n is a basis of V, and thus the equation above does indeed deﬁne a function T from V to W (because each element of V can be uniquely written in the form c1*v*1 C ... C cn*v*n).

For each j , taking cj D 1 and the other c’s equal to 0 in the equation above shows that T *v*j D *w*j .

If u; *v* 2 V with u D a1*v*1 C ... C an*v*n and *v* D c1*v*1 C ... C cn*v*n, then

（ ）

T .u C *v*/ D T .a1 C c1/*v*1 C ... C .an C cn/*v*n

D .a1 C c1/*w*1 C ... C .an C cn/*w*n

D .a1*w*1 C ... C an*w*n/ C .c1*w*1 C ... C cn*w*n/

D Tu C T *v*:

Similarly, if 入 2 **F** and *v* D c1*v*1 C ... C cn*v*n, then

T .入*v*/ D T .入c1*v*1 C ... C 入cn*v*n/ D 入c1*w*1 C ... C 入cn*w*n D 入.c1*w*1 C ... C cn*w*n/ D 入T *v*:

Thus T is a linear map from V to W.

To prove uniqueness, now suppose that T 2 *L*.V; W / and that T *v*j D *w*j for j D 1; :::; n. Let c1;:::; cn 2 **F**. The homogeneity of T implies that T .cj *v*j / D cj *w*j for j D 1; :::; n. The additivity of T now implies that

T .c1*v*1 C ... C cn*v*n/ D c1*w*1 C ... C cn*w*n:

Thus T is uniquely determined on span.*v*1;:::; *v*n/ by the equation above. Because *v*1;:::; *v*n is a basis of V, this implies that T is uniquely determined on V.

SECTION 3.A The Vector Space of Linear Maps **55**

**Algebraic Operations on *L***.V; W /

We begin by deﬁning addition and scalar multiplication on *L*.V; W /.

3.6 **Deﬁnition *addition and scalar multiplication on*** *L*.V; W /

入T are the linear maps from V to W deﬁned by

.S C T /.*v*/ D S *v* C T *v* and .入T /.*v*/ D 入.T *v*/

for all *v* 2 V.

Suppose S; T 2 *L*.V; W / and 入 2 **F**. The ***sum*** S C T and the ***product***

You should verify that S T and 入T as deﬁned above are indeed linear maps. In other words, if S; T 2 *L*.V; W / and

入 2 **F**, then S C T 2 *L*.V; W / and

*Although linear maps are perva- sive throughout mathematics, they are not as ubiquitous as imagined by some confused students who seem to think that* cos *is a linear map from* **R** *to* **R** *when they write*

*that* cos 2x *equals* 2 cos x *and that*

cos.x C y/ *equals* cos x C cos y*.*

C

入 T .V; W /

2 *L*

.

Because we took the trouble to de-

ﬁne addition and scalar multiplication on .V; W /, the next result should not be a surprise.

*L*

3.7 *L*.V; W / is a vector space

With the operations of addition and scalar multiplication as deﬁned above,

*L*.V; W / is a vector space.

The routine proof of the result above is left to the reader. Note that the additive identity of .V; W / is the zero linear map deﬁned earlier in this section.

*L*

Usually it makes no sense to multiply together two elements of a vector space, but for some pairs of linear maps a useful product exists. We will need

a third vector space, so for the rest of this section suppose U is a vector space

over **F**.

3.8 **Deﬁnition *Product of Linear Maps***

If T 2 *L*.U; V / and S 2 *L*.V; W /, then the ***product*** ST 2 *L*.U; W / is

deﬁned by

.ST /.u/ D S.T u/

for u 2 U.

**56** CHAPTER 3 Linear Maps

In other words, ST is just the usual composition S T of two functions, but when both functions are linear, most mathematicians write ST instead of S T. You should verify that ST is indeed a linear map from U to W whenever T .U; V / and S .V; W /.

ı

2 *L* 2 *L*

ı

Note that ST is deﬁned only when T maps into the domain of S.

* 1. Algebraic properties of products of linear maps

#### associativity

.T1T2/T3 D T1.T2T3/

whenever T1, T2, and T3 are linear maps such that the products make sense (meaning that T3 maps into the domain of T2, and T2 maps into the domain of T1).

#### identity

TI D IT D T

whenever T .V; W / (the ﬁrst I is the identity map on V, and the second I is the identity map on W ).

2 *L*

#### distributive properties

.S1 C S2/T D S1T C S2T and S.T1 C T2/ D ST1 C S T2

whenever T; T1; T2 2 *L*.U; V / and S; S1; S2 2 *L*.V; W /.

The routine proof of the result above is left to the reader.

Multiplication of linear maps is not commutative. In other words, it is not necessarily true that ST D TS, even if both sides of the equation make sense.

* 1. **Example** Suppose D 2 *L*（（*P*.**R**/; *P*.**R**/）） is the differentiation map

deﬁned in Example 3.4 and T 2 *L P*.**R**/; *P*.**R**/

is the multiplication by x2

map deﬁned earlier in this section. Show that TD ¤ DT.

Solution We have

（ ） （ ）

.TD/p .x/ D x2p0.x/ but .DT /p .x/ D x2p0.x/ C 2xp.x/:

In other words, differentiating and then multiplying by x2 is not the same as multiplying by x2 and then differentiating.

SECTION 3.A The Vector Space of Linear Maps **57**

3.11 Linear maps take 0 to 0

Suppose T is a linear map from V to W. Then T .0/ D 0.

Proof By additivity, we have

T .0/ D T .0 C 0/ D T .0/ C T .0/:

Add the additive inverse of T .0/ to each side of the equation above to conclude that T .0/ D 0.

EXERCISES 3.A

1. Suppose b; c 2 **R**. Deﬁne T W **R**3 ! **R**2 by

T .x; y; z/ D .2x — 4y C 3z C b; 6x C cxyz/:

Show that T is linear if and only if b D c D 0.

1. Suppose b; c 2 **R**. Deﬁne T W *P*.**R**/ ! **R**2 by

Z

Tp D

3p.4/ C 5p0.6/ C bp.1/p.2/;

2

x3p.x/ dx c sin p.0/ :

C

1

Show that T is linear if and only if b D c D 0.

1. Suppose T 2 *L*.**F**n; **F**m/. Show that there exist scalars Aj;k 2 **F** for

j D 1; :::;m and k D 1; :::;n such that

T .x1;:::; xn/ D .A1;1x1 C... CA1;nxn;:::; Am;1x1 C... CAm;nxn/

for every .x1;:::; xn/ **F**n.

2

[*The exercise above shows that* T *has the form promised in the last item*

*of Example 3.4.*]

1. Suppose T .V; W / and *v*1;:::; *v*m is a list of vectors in V such that T *v*1;:::;T *v*m is a linearly independent list in W. Prove that *v*1;:::; *v*m is linearly independent.

2 *L*

1. Prove the assertion in 3.7.
2. Prove the assertions in 3.9.

**58** CHAPTER 3 Linear Maps

1. Show that every linear map from a 1-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if dim V D 1 and T 2 *L*.V; V /, then there exists 入 2 **F** such that T *v* D 入 *v* for all *v* 2 V.
2. Give an example of a function ' W **R**2 ! **R** such that

'.a*v*/ D a'.*v*/

for all a **R** and all *v* **R**2 but ' is not linear.

2 2

[*The exercise above and the next exercise show that neither homogeneity*

*nor additivity alone is enough to imply that a function is a linear map.*]

1. Give an example of a function ' W **C** ! **C** such that

'.*w* C z/ D '.*w*/ C '.z/

for all *w*;z **C** but ' is not linear. (Here **C** is thought of as a complex vector space.)

2

[*There also exists a function* ' **R R** *such that* ' *satisﬁes the additiv- ity condition above but* ' *is not linear. However, showing the existence of such a function involves considerably more advanced tools.*]

W !

1. Suppose U is a subspace of V with U ¤ V. Suppose S 2 *L*.U; W / and

S ¤ 0 (which means that Su ¤ 0 for some u 2 U ). Deﬁne T W V ! W

by

D (

T *v* S *v* if *v* 2 U;

0 if *v* 2 V and *v* … U:

Prove that T is not a linear map on V.

1. Suppose V is ﬁnite-dimensional. Prove that every linear map on a subspace of V can be extended to a linear map on V. In other words, show that if U is a subspace of V and S 2 *L*.U; W /, then there exists T 2 *L*.V; W / such that Tu D Su for all u 2 U.
2. Suppose V is ﬁnite-dimensional with dim V > 0, and suppose W is inﬁnite-dimensional. Prove that *L*.V; W / is inﬁnite-dimensional.
3. Suppose *v*1;:::; *v*m is a linearly dependent list of vectors in V. Suppose also that W ¤ f0g. Prove that there exist *w*1;:::; *w*m 2 W such that no T 2 *L*.V; W / satisﬁes T *v*k D *w*k for each k D 1; :::; m.
4. Suppose V is ﬁnite-dimensional with dim V 乏 2. Prove that there exist

S; T 2 *L*.V; V / such that ST ¤ TS.

SECTION 3.B Null Spaces and Ranges **59**

## *Null Spaces and Ranges*

3.B

### Null Space and Injectivity

In this section we will learn about two subspaces that are intimately connected with each linear map. We begin with the set of vectors that get mapped to 0.

3.12 **Deﬁnition *null space,*** null T

consisting of those vectors that T maps to 0:

null T D f*v* 2 V W T *v* D 0g:

For T 2 *L*.V; W /, the ***null space*** of T, denoted null T, is the subset of V

3.13 **Example *null space***

* If T is the zero map from V to W, in other words if T *v* D 0 for every

*v* 2 V, then null T D V.

* Suppose ' 2 *L*.**C**3; **F**/ is deﬁned by '.z1; z2; z3/ D z1 C 2z2 C 3z3. Then null ' D f.z1; z2; z3/ 2 **C**3 W z1 C 2z2 C 3z3 D 0g. A basis of null ' is .—2; 1; 0/; .—3; 0; 1/.

0 （ ）

Suppose D .**R**/; .**R**/ is the differentiation map deﬁned by Dp p . The only functions whose derivative equals the zero function are the constant functions. Thus the null space of D equals the set of

D

* 2 *L P P*

constant functions.

（ ）

* Suppose T 2 *L P*.**R**/; *P*.**R**/ is the multiplication by x2 map deﬁned by .Tp/.x/ D x2p.x/. The only polynomial p such that x2p.x/ D 0 for all x 2 **R** is the 0 polynomial. Thus null T D f0g.
* Suppose T 2 *L*.**F**1; **F**1/ is the backward shift deﬁned by

T .x1; x2; x3;:::/ D .x2; x3;::: /:

Clearly T .x1; x2; x3;:::/ equals 0 if and only if x2; x3;::: are all 0.

Thus in this case we have null T D f.a; 0; 0;:::/ W a 2 **F**g.

The next result shows that the null space of each linear map is a subspace

*Some mathematicians use the term* ***kernel*** *instead of null space. The word “null” means zero. Thus the term “null space”should remind*

*you of the connection to* 0*.*

of the domain. In particular, 0 is in the

null space of every linear map.

**60** CHAPTER 3 Linear Maps

3.14 The null space is a subspace

Suppose T 2 *L*.V; W /. Then null T is a subspace of V.

Proof Because T is a linear map, we know that T .0/ D 0 (by 3.11). Thus

0 2 null T.

Suppose u; *v* 2 null T. Then

T .u C *v*/ D Tu C T *v* D 0 C 0 D 0:

Hence u C *v* 2 null T. Thus null T is closed under addition.

Suppose u 2 null T and 入 2 **F**. Then

T .入u/ D 入T u D 入0 D 0:

Hence 入 u null T. Thus null T is closed under scalar multiplication.

2

We have shown that null T contains 0 and is closed under addition and scalar multiplication. Thus null T is a sub- space of V (by 1.34).

*Take another look at the null spaces that were computed in Example*

*3.13 and note that all of them are subspaces.*

As we will soon see, for a linear map the next deﬁnition is closely connected to the null space.

3.15 **Deﬁnition *injective***

A function T W V ! W is called ***injective*** if Tu D T *v* implies u D *v*.

The deﬁnition above could be rephrased to say that T is injective if u *v* implies that Tu T *v*. In other words, T is injective if it maps distinct

*Many mathematicians use the term* ***one-to-one****, which means the same as injective.*

¤ ¤

inputs to distinct outputs.

The next result says that we can check whether a linear map is injective by checking whether 0 is the only vector that gets mapped to 0. As a simple application of this result, we see that of the linear maps whose null spaces we computed in 3.13, only multiplication by x2 is injective (except that the zero

map is injective in the special case V D f0g).

SECTION 3.B Null Spaces and Ranges **61**

3.16 Injectivity is equivalent to null space equals f0g

Let T 2 *L*.V; W /. Then T is injective if and only if null T D f0g.

Proof First suppose T is injective. We want to prove that null T D f0g. We already know that f0gc null T (by 3.11). To prove the inclusion in the other direction, suppose *v* 2 null T. Then

T .*v*/ D 0 D T .0/:

Because T is injective, the equation above implies that *v* D 0. Thus we can conclude that null T D f0g, as desired.

To prove the implication in the other direction, now suppose null T D f0g. We want to prove that T is injective. To do this, suppose u; *v* 2 V and Tu D T *v*. Then

0 D Tu — T *v* D T .u — *v*/:

Thus u — *v* is in null T, which equals f0g. Hence u — *v* D 0, which implies that u D *v*. Hence T is injective, as desired.

### Range and Surjectivity

Now we give a name to the set of outputs of a function.

3.17 **Deﬁnition *range***

For T a function from V to W, the ***range*** of T is the subset of W consisting of those vectors that are of the form T *v* for some *v* 2 V :

range T D fT *v* W *v* 2 V g:

3.18 **Example *range***

* If T is the zero map from V to W, in other words if T *v* D 0 for every

*v* 2 V, then range T D f0g.

* Suppose T 2 *L*.**R**2; **R**3/ is deﬁned by T .x; y/ D .2x; 5y; x C y/, then range T D f.2x; 5y; x C y/ W x; y 2 **R**g. A basis of range T is

.

.2; 0; 1/; .0; 5; 1/

（ ）

* Suppose D 2 *L P*.**R**/; *P*.**R**/ is the differentiation map deﬁned by Dp D p0. Because for every polynomial q 2 *P*.**R**/ there exists a polynomial p 2 *P*.**R**/ such that p0 D q, the range of D is *P*.**R**/.

**62** CHAPTER 3 Linear Maps

The next result shows that the range of each linear map is a subspace of the vector space into which it is being mapped.

*Some mathematicians use the word* ***image****, which means the same as range.*

3.19 The range is a subspace

If T 2 *L*.V; W /, then range T is a subspace of W.

Proof Suppose T 2 *L*.V; W /. Then T .0/ D 0 (by 3.11), which implies that

0 2 range T.

If *w*1; *w*2 2 range T, then there exist *v*1; *v*2 2 V such that T *v*1 D *w*1 and

T *v*2 D *w*2. Thus

T .*v*1 C *v*2/ D T *v*1 C T *v*2 D *w*1 C *w*2:

Hence *w*1 *w*2 range T. Thus range T is closed under addition.

C 2

If *w* range T and 入 **F**, then there exists *v* V such that T *v w*.

2 2 2 D

Thus

T .入*v*/ D 入T *v* D 入*w*:

Hence 入*w* range T. Thus range T is closed under scalar multiplication.

2

We have shown that range T contains 0 and is closed under addition and scalar multiplication. Thus range T is a subspace of W (by 1.34).

3.20 **Deﬁnition *surjective***

A function T W V ! W is called ***surjective*** if its range equals W.

To illustrate the deﬁnition above, note that of the ranges we computed in

3.18, only the differentiation map is surjective (except that the zero map is surjective in the special case W D f0g.

Whether a linear map is surjective

*Many mathematicians use the term* ***onto****, which means the same as sur- jective.*

depends on what we are thinking of as the vector space into which it maps.

3.21 **Example** The differentiation map D 2 *L*（*P*5.**R**/; *P*5.**R**/） deﬁned by Dp D p0 is not surjective, because the polyn（omial x5 is not）in the range

of D. However, the differentiation map S 2 *L P*5.**R**/; *P*4.**R**/

deﬁned by

D *P*

Sp p0 is surjective, because its range equals 4.**R**/, which is now the

vector space into which S maps.

SECTION 3.B Null Spaces and Ranges **63**

### Fundamental Theorem of Linear Maps

The next result is so important that it gets a dramatic name.

3.22 Fundamental Theorem of Linear Maps

Suppose V is ﬁnite-dimensional and T 2 *L*.V; W /. Then range T is

ﬁnite-dimensional and

dim V D dim null T C dim range T:

Proof Let u1;:::; um be a basis of null T ; thus dim null T m. The linearly independent list u1;:::; um can be extended to a basis

u1;:::; um; *v*1;:::; *v*n

D

of V (by 2.33). Thus dim V m n. To complete the proof, we need only show that range T is ﬁnite-dimensional and dim range T n. We will do this by proving that T *v*1;:::;T *v*n is a basis of range T.

D

D C

Let *v* 2 V. Because u1;:::; um; *v*1;:::; *v*n spans V, we can write

*v* D a1u1 C ... C amum C b1*v*1 C ... C bn*v*n;

where the a’s and b’s are in **F**. Applying T to both sides of this equation, we get

T *v* D b1T *v*1 C ... C bnT *v*n;

where the terms of the form T uj disappeared because each uj is in null T. The last equation implies that T *v*1;:::;T *v*n spans range T. In particular, range T is ﬁnite-dimensional.

To show T *v*1;:::;T *v*n is linearly independent, suppose c1;:::; cn **F**

2

and

Then Hence

c1T *v*1 C ... C cnT *v*n D 0:

T .c1*v*1 C ... C cn*v*n/ D 0: c1*v*1 C ... C cn*v*n 2 null T:

Because u1;:::; um spans null T, we can write

c1*v*1 C ... C cn*v*n D d1u1 C ... C dmum;

where the d’s are in **F**. This equation implies that all the c’s (and d’s) are 0 (because u1;:::; um; *v*1;:::; *v*n is linearly independent). Thus T *v*1;:::;T *v*n is linearly independent and hence is a basis of range T, as desired.

**64** CHAPTER 3 Linear Maps

Now we can show that no linear map from a ﬁnite-dimensional vector space to a “smaller” vector space can be injective, where “smaller” is measured by dimension.

3.23 A map to a smaller dimensional space is not injective

Suppose V and W are ﬁnite-dimensional vector spaces such that dim V > dim W. Then no linear map from V to W is injective.

Proof Let T 2 *L*.V; W /. Then

dim null T D dim V — dim range T

乏 dim V — dim W

> 0;

where the equality above comes from the Fundamental Theorem of Linear Maps (3.22). The inequality above states that dim null T > 0. This means that null T contains vectors other than 0. Thus T is not injective (by 3.16).

The next result shows that no linear map from a ﬁnite-dimensional vector space to a “bigger” vector space can be surjective, where “bigger” is measured by dimension.

3.24 A map to a larger dimensional space is not surjective

Suppose V and W are ﬁnite-dimensional vector spaces such that dim V < dim W. Then no linear map from V to W is surjective.

Proof Let T 2 *L*.V; W /. Then

dim range T D dim V — dim null T

三 dim V

< dim W;

where the equality above comes from the Fundamental Theorem of Linear Maps (3.22). The inequality above states that dim range T < dim W. This means that range T cannot equal W. Thus T is not surjective.

As we will now see, 3.23 and 3.24 have important consequences in the theory of linear equations. The idea here is to express questions about systems of linear equations in terms of linear maps.

SECTION 3.B Null Spaces and Ranges **65**

3.25 **Example** Rephrase in terms of a linear map the question of whether a homogeneous system of linear equations has a nonzero solution.

Solution

Fix positive integers m and n, and let Aj;k **F** for j 1; :::;m and k 1; :::; n. Consider the homoge-

***Homogeneous****, in this context, means that the constant term on the*

*right side of each equation below is* 0*.*

D

2 D

neous system of linear equations

kD1

Xn

Xn

A1;kxk D 0

:

#### :

kD1

Am;kxk D 0:

Obviously x1 xn 0 is a solution of the system of equations above; the question here is whether any other solutions exist.

D ... D D

Deﬁne T W **F**n ! **F**m by

X

T .x1;:::; xn/ D

（Xn

n

A1;kxk;:::;

Am;kxk）:

kD1 kD1

The equation T .x1;:::; xn/ 0 (the 0 here is the additive identity in **F**m, namely, the list of length m of all 0’s) is the same as the homogeneous system of linear equations above.

D

Thus we want to know if null T is strictly bigger than 0 . In other words,

f g

we can rephrase our question about nonzero solutions as follows (by 3.16): What condition ensures that T is not injective?

3.26 Homogeneous system of linear equations

A homogeneous system of linear equations with more variables than equations has nonzero solutions.

Proof Use the notation and result from the example above. Thus T is a linear map from **F**n to **F**m, and we have a homogeneous system of m linear equations with n variables x1;:::; xn. From 3.23 we see that T is not injective if n> m.

Example of the result above: a homogeneous system of four linear equa- tions with ﬁve variables has nonzero solutions.

**66** CHAPTER 3 Linear Maps

3.27 **Example** Consider the question of whether an inhomogeneous sys- tem of linear equations has no solutions for some choice of the constant terms. Rephrase this question in terms of a linear map.

Solution Fix positive integers m and n, and let Aj;k 2 **F** for j D 1; :::;m

and k D 1; :::; n. For c1;:::; cm 2 **F**, consider the system of linear equations

Xn

#### 3.28

kD1

Xn

A1;kxk D c1

:

#### :

kD1

Am;kxk D cm:

The question here is whether there is some choice of c1;:::; cm **F** such that no solution exists to the system above.

2

Deﬁne T W **F**n ! **F**m by

X

T .x1;:::; xn/ D

（Xn

n

A1;kxk;:::;

Am;kxk）:

kD1 kD1

The equation T .x1;:::; xn/ .c1;:::; cm/ is the same as the system of equa- tions 3.28. Thus we want to know if range T **F**m. Hence we can rephrase our question about not having a solution for some choice of c1;:::; cm **F** as follows: What condition ensures that T is not surjective?

3.29 Inhomogeneous system of linear equations

An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of the constant terms.

2

¤

D

Proof Use the notation and result from the example above. Thus T is a lin- ear map from **F**n to **F**m, and we have a system of m equations with n variables x1;:::; xn. From 3.24 we see that T is not surjective if n< m.

Example of the result above: an inhomogeneous system of ﬁve linear

*Our results about homogeneous systems with more variables than equations and inhomogeneous sys- tems with more equations than vari- ables (3.26 and 3.29) are often proved using Gaussian elimination. The abstract approach taken here leads to cleaner proofs.*

equations with four variables has no solution for some choice of the con- stant terms.

EXERCISES 3.B

SECTION 3.B Null Spaces and Ranges **67**

1. Give an example of a linear map T such that dim null T D 3 and dim range T D 2.
2. Suppose V is a vector space and S; T 2 *L*.V; V / are such that

range S c null T:

Prove that .ST /2 D 0.

1. Suppose *v*1;:::; *v*m is a list of vectors in V. Deﬁne T 2 *L*.**F**m;V / by

T .z1;:::; zm/ D z1*v*1 C ... C zm*v*m:

* 1. What property of T corresponds to *v*1;:::; *v*m spanning V ?
  2. What property of T corresponds to *v*1;:::; *v*m being linearly independent?

1. Show that

fT 2 *L*.**R**5; **R**4/ W dim null T > 2g

is not a subspace of *L*.**R**5; **R**4/.

1. Give an example of a linear map T W **R**4 ! **R**4 such that

range T D null T:

1. Prove that there does not exist a linear map T W **R**5 ! **R**5 such that

range T D null T:

1. Suppose V and W are ﬁnite-dimensional with 2 三 dim V 三 dim W. Show that fT 2 *L*.V; W / W T is not injectiveg is not a subspace of *L*.V; W /.
2. Suppose V and W are ﬁnite-dimensional with dim V 乏 dim W 乏 2. Show that fT 2 *L*.V; W / W T is not surjectiveg is not a subspace of *L*.V; W /.
3. Suppose T .V; W / is injective and *v*1;:::; *v*n is linearly independent in V. Prove that T *v*1;:::;T *v*n is linearly independent in W.

2 *L*

**68** CHAPTER 3 Linear Maps

1. Suppose *v*1;:::; *v*n spans V and T .V; W /. Prove that the list

2 *L*

T *v*1;:::;T *v*n spans range T.

1. Suppose S1;:::; Sn are injective linear maps such that S1S2 ... Sn

makes sense. Prove that S1S2 ... Sn is injective.

1. Suppose that V is ﬁnite-dimensional and that T 2 *L*.V; W /. Prove that there exists a subspace U of V such that U \ null T D f0g and range T D fTu W u 2 U g.
2. Suppose T is a linear map from **F**4 to **F**2 such that

null T D f.x1; x2; x3; x4/ 2 **F**4 W x1 D 5x2 and x3 D 7x4g:

Prove that T is surjective.

1. Suppose U is a 3-dimensional subspace of **R**8 and that T is a linear map from **R**8 to **R**5 such that null T D U. Prove that T is surjective.
2. Prove that there does not exist a linear map from **F**5 to **F**2 whose null space equals

f.x1; x2; x3; x4; x5/ 2 **F**5 W x1 D 3x2 and x3 D x4 D x5g:

1. Suppose there exists a linear map on V whose null space and range are both ﬁnite-dimensional. Prove that V is ﬁnite-dimensional.
2. Suppose V and W are both ﬁnite-dimensional. Prove that there exists an injective linear map from V to W if and only if dim V 三 dim W.
3. Suppose V and W are both ﬁnite-dimensional. Prove that there exists a surjective linear map from V onto W if and only if dim V 乏 dim W.
4. Suppose V and W are ﬁnite-dimensional and that U is a subspace of V. Prove that there exists T 2 *L*.V; W / such that null T D U if and only if dim U 乏 dim V — dim W.
5. Suppose W is ﬁnite-dimensional and T .V; W /. Prove that T is injective if and only if there exists S .W; V / such that ST is the identity map on V.

2 *L*

2 *L*

1. Suppose V is ﬁnite-dimensional and T .V; W /. Prove that T is surjective if and only if there exists S .W; V / such that TS is the identity map on W.

2 *L*

2 *L*

SECTION 3.B Null Spaces and Ranges **69**

1. Suppose U and V are ﬁnite-dimensional vector spaces and S 2 *L*.V; W /

and T 2 *L*.U; V /. Prove that

dim null ST 三 dim null S C dim null T:

1. Suppose U and V are ﬁnite-dimensional vector spaces and S 2 *L*.V; W /

and T 2 *L*.U; V /. Prove that

dim range ST 三 minfdim range S; dim range T g:

1. Suppose W is ﬁnite-dimensional and T1; T2 2 *L*.V; W /. Prove that null T1 c null T2 if and only if there exists S 2 *L*.W; W / such that T2 D S T1.
2. Suppose V is ﬁnite-dimensional and T1; T2 2 *L*.V; W /. Prove that range T1 c range T2 if and only if there exists S 2 *L*.V; V / such that T1 D T2S.

（ ）

1. Suppose D .**R**/; .**R**/ is such that deg Dp .deg p/ 1 for every nonconstant polynomial p .**R**/. Prove that D is surjective. [*The notation* D *is used above to remind you of the differentiation map that sends a polynomial* p *to* p0*. Without knowing the formula for the derivative of a polynomial (except that it reduces the degree by* 1*), you can use the exercise above to show that for every polynomial* q 2 *P*.**R**/*, there exists a polynomial* p 2 *P*.**R**/ *such that* p0 D q*.*]

2 *P*

2 *L P P* D —

1. Suppose p .**R**/. Prove that there exists a polynomial q .**R**/ such that 5q00 3q0 p.

C D

2 *P* 2 *P*

[*This exercise can be done without linear algebra, but it’s more fun to do it using linear algebra.*]

1. Suppose T 2 *L*.V; W /, and *w*1;:::; *w*m is a basis of range T. Prove that there exist '1;:::; 'm 2 *L*.V; **F**/ such that

T *v* D '1.*v*/*w*1 C C 'm.*v*/*w*m

for every *v* 2 V.

1. Suppose ' 2 *L*.V; **F**/. Suppose u 2 V is not in null '. Prove that

V D null ' ˚ fau W a 2 **F**g:

1. Suppose '1 and '2 are linear maps from V to **F** that have the same null space. Show that there exists a constant c 2 **F** such that '1 D c'2.
2. Give an example of two linear maps T1 and T2 from **R**5 to **R**2 that have the same null space but are such that T1 is not a scalar multiple of T2.

**70** CHAPTER 3 Linear Maps

## *Matrices*

3.C

### Representing a Linear Map by a Matrix

We know that if *v*1;:::; *v*n is a basis of V and T V W is linear, then the values of T *v*1;:::;T *v*n determine the values of T on arbitrary vectors in V

W !

(see 3.5). As we will soon see, matrices are used as an efﬁcient method of recording the values of the T *v*j ’s in terms of a basis of W.

3.30 **Deﬁnition *matrix,*** Aj;k

Let m and n denote positive integers. An m-by-n ***matrix*** A is a rectangular array of elements of **F** with m rows and n columns:

A D

B

A

1;1

::: A

1;n

@

:

Am;1

C

A

:

:::

:

Am;n

The notation Aj;k denotes the entry in row j , column k of A. In other

words, the ﬁrst index refers to the row number and the second index refers

to the column number.

Thus A2;3 refers to the entry in the second row, third column of a matrix A.

3.31 **Example** If A 8 4 5 — 3i

D

1 9 7

, then A2;3 D 7.

Now we come to the key deﬁnition in this section.

The matrix .T / of a linear map T .V; W / depends on the basis

3.32 **Deﬁnition *matrix of a linear map,*** *M*.T /

a basis of W. The ***matrix of*** T with respect to these bases is the m-by-n

matrix *M*.T / whose entries Aj;k are deﬁned by

T *v*k D A1;k*w*1 C ... C Am;k*w*m:

Suppose T 2 *L*.V; W / and *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is

If the bases are not clear from the context, then the notation

*M* T; .

（

*v* ;:::;

1

*v* /; .*w* ;:::;

n 1

*w* / is used.

m

）

*M* 2 *L*

*v*1;:::; *v*n of V and the basis *w*1;:::; *w*m of W, as well as on T. However, the

bases should be clear from the context, and thus they are often not included in

the notation.

SECTION 3.C Matrices **71**

To remember how .T / is constructed from T, you might write across the top of the matrix the basis vectors *v*1;:::; *v*n for the domain and along the left the basis vectors *w*1;:::; *w*m for the vector space into which T maps, as

*M*

follows:

*w*1

B

0

*M*.T / D :

@

*w*

m

*v*1 ::: *v*k ::: *v*n

A1;k

:

Am;k

*The* k*th column of M*.T / *con-*

*sists of the scalars needed to write*

T *v*k *as a linear combination of*

.*w*1;:::; *w*m/*:*

X

m

T *v*k D

Aj;k *w*j *.*

j D1

1CA :

In the matrix above only the kth col- umn is shown. Thus the second index

of each displayed entry of the matrix above is k. The picture above should

remind you that T *v* can be computed

k

*M*.T / by multiplying each entry

in the kth column by the correspond- ing *w*j from the left column, and then adding up the resulting vectors.

from

If T is a linear map from **F**n to **F**m,

*If* T *maps an* n*-dimensional vector space to an* m*-dimensional vector*

*space, then M*.T / *is an* m*-by-*n

*matrix.*

then unless stated otherwise, assume the

bases in question are the standard ones (where the k basis vector is 1 in the kth slot and 0 in all the other slots). If you think of elements of **F**m as columns of m numbers, then you can think of the kth column of .T / as T applied to the kth standard basis vector.

*M*

th

* 1. **Example** Suppose T 2 *L*.**F**2; **F**3/ is deﬁned by

T .x; y/ D .x C 3y; 2x C 5y; 7x C 9y/:

Find the matrix of T with respect to the standard bases of **F**2 and **F**3.

Solution Because T .1; 0/ .1; 2; 7/ and T .0; 1/ .3; 5; 9/, the matrix of

D D

T with respect to the standard bases is the 3-by-2 matrix below:

0 1 3 1

*M*.T / D @ 2 5 A :

7 9

**72** CHAPTER 3 Linear Maps

When working with m.**F**/, use the standard basis 1; x; x2;:::; xm unless the context indicates otherwise.

*P*

（ ）

* 1. **Example** Suppose D 2 *L P*3.**R**/; *P*2.**R**/ is the differentiation map deﬁned by Dp D p0. Find the matrix of D with respect to the standard bases of *P*3.**R**/ and *P*2.**R**/.

Solution Because .xn/0 nxn 1, the matrix of T with respect to the standard bases is the 3-by-4 matrix below:

D

0@ 0 0 1A

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1 | 0 |  |
| *M*.D/ D 0 | 0 | 2 | 0 : |
| 0 | 0 | 0 | 3 |

### Addition and Scalar Multiplication of Matrices

For the rest of this section, assume that V and W are ﬁnite-dimensional and that a basis has been chosen for each of these vector spaces. Thus for each linear map from V to W, we can talk about its matrix (with respect to the chosen bases, of course). Is the matrix of the sum of two linear maps equal to

the sum of the matrices of the two maps?

Right now this question does not make sense, because although we have deﬁned the sum of two linear maps, we have not deﬁned the sum of two matrices. Fortunately, the obvious deﬁnition of the sum of two matrices has the right properties. Speciﬁcally, we make the following deﬁnition.

3.35 **Deﬁnition *matrix addition***

The ***sum of two matrices of the same size*** is the matrix obtained by adding corresponding entries in the matrices:

B

A

1;1

::: A

1;n

:

Am;1

1

C

1;1

:::

C

1;n

@

:::

:

Am;n

C B

A @

C

:

Cm;1

:::

:

Cm;n

0B

C

A1;1 C C1;1

A

:::

D @

A1;n C C1;n

:

Am;1 C Cm;1

:::

:

Am;n C Cm;n

1C

In other words, .A C C /j;k D Aj;k C Cj;k.

A :

SECTION 3.C Matrices **73**

In the following result, the assumption is that the same bases are used for all three linear maps S C T, S, and T.

3.36 The matrix of the sum of linear maps

Suppose S; T 2 *L*.V; W /. Then *M*.S C T/ D *M*.S/ C *M*.T /.

The veriﬁcation of the result above is left to the reader.

Still assuming that we have some bases in mind, is the matrix of a scalar times a linear map equal to the scalar times the matrix of the linear map? Again the question does not make sense, because we have not deﬁned scalar multiplication on matrices. Fortunately, the obvious deﬁnition again has the right properties.

3.37 **Deﬁnition *scalar multiplication of a matrix***

The product of a scalar and a matrix is the matrix obtained by multiplying

each entry in the matrix by the scalar:

入

B

A

1;1

:::

A

1;n

@

:

Am;1

1

入A

1;1

::: 入A

1;n

:::

:

Am;n

C B

A @

D

:

入Am;1

C

A

:

:::

:

入Am;n

In other words, .入A/j;k D 入Aj;k.

In the following result, the assumption is that the same bases are used for both linear maps 入T and T.

3.38 The matrix of a scalar times a linear map

Suppose 入 2 **F** and T 2 *L*.V; W /. Then *M*.入T / D 入*M*.T /.

The veriﬁcation of the result above is also left to the reader.

Because addition and scalar multiplication have now been deﬁned for matrices, you should not be surprised that a vector space is about to appear. We need only a bit of notation so that this new vector space has a name.

3.39 **Notation**

**F**m;n

For m and n positive integers, the set of all m-by-n matrices with entries

in **F** is denoted by **F**m;n.

**74** CHAPTER 3 Linear Maps

3.40 dim **F**m;n D mn

Suppose m and n are positive integers. With addition and scalar multipli- cation deﬁned as above, **F**m;n is a vector space with dimension mn.

Proof The veriﬁcation that **F**m;n is a vector space is left to the reader. Note that the additive identity of **F**m;n is the m-by-n matrix whose entries all equal 0.

The reader should also verify that the list of m-by-n matrices that have 0 in all entries except for a 1 in one entry is a basis of **F**m;n. There are mn such matrices, so the dimension of **F**m;n equals mn.

### Matrix Multiplication

Suppose, as previously, that *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is a basis of W. Suppose also that we have another vector space U and that u1;:::; up is a basis of U.

Consider linear maps T U V and S V W. The composition

W ! W !

ST is a linear map from U to W. Does .ST / equal .S/ .T /? This

*M M M*

question does not yet make sense, because we have not deﬁned the product of

two matrices. We will choose a deﬁnition of matrix multiplication that forces this question to have a positive answer. Let’s see how to do this.

Suppose *M*.S/ D A and *M*.T / D C . For 1 三 k 三 p, we have

（X

.ST /uk D S

n

r D1

Cr;k *v*r ）

n

X

D

r 1 n

X

D

D

r D1

Cr;k S *v*r

m

X

Cr;k

j D1

Aj;r *w*j

m n

X（X

D

j D1 r D1

Aj;r Cr;k

）*w*j :

Thus *M*.ST / is the m-by-p matrix whose entry in row j , column k, equals

Xn

Aj;r Cr;k :

r D1

SECTION 3.C Matrices **75**

Now we see how to deﬁne matrix multiplication so that the desired equation

*M*.ST / D *M*.S/*M*.T / holds.

3.41 **Deﬁnition *matrix multiplication***

Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then AC is deﬁned to be the m-by-p matrix whose entry in row j , column k, is given by the following equation:

.AC/j;k D

X

n

Aj;r Cr;k :

r D1

In other words, the entry in row j , column k, of AC is computed by taking row j of A and column k of C , multiplying together corresponding

entries, and then summing.

Note that we deﬁne the product of two matrices only when the number of columns of the ﬁrst matrix equals the number of rows of the second matrix.

3.42 **Example** Here we multiply together a 3-by-2 matrix and a 2-by-4

matrix, obtaining a 3-by-4 matrix:

*You may have learned this deﬁni- tion of matrix multiplication in an earlier course, although you may not have seen the motivation for it.*

0 1 2 1 6 5 4 3

@ A

@ A

3 4

5 6

2 1 0 —1

D 26 19 12 5 :

42 31 20 9

0 10 7 4 1 1

Matrix multiplication is not commutative. In other words, AC is not necessarily equal to CA even if both products are deﬁned (see Exercise 12). Matrix multiplication is distributive and associative (see Exercises 13 and 14).

In the following result, the assumption is that the same basis of V is used in considering T 2 *L*.U; V / and S 2 *L*.V; W /, the same basis of W is used in considering S 2 *L*.V; W / and ST 2 *L*.U; W /, and the same basis of U is used in considering T 2 *L*.U; V / and ST 2 *L*.U; W /.

3.43 The matrix of the product of linear maps

If T 2 *L*.U; V / and S 2 *L*.V; W /, then *M*.ST / D *M*.S/*M*.T /.

The proof of the result above is the calculation that was done as motivation before the deﬁnition of matrix multiplication.

**76** CHAPTER 3 Linear Maps

In the next piece of notation, note that as usual the ﬁrst index refers to a row and the second index refers to a column, with a vertically centered dot used as a placeholder.

3.44 **Notation**

Aj; ***,*** A ;k

Suppose A is an m-by-n matrix.

* If 1 三 j 三 m, then Aj; denotes the 1-by-n matrix consisting of

row j of A.

* If 1 三 k 三 n, then A ;k denotes the m-by-1 matrix consisting of

column k of A.

3.45 **Example** If A 8 4 5 , then A is row 2 of A and A is

1 9 7

D 2; ;2

column 2 of A. In other words,

（ ）D D2; ;2

A 1 9 7 and A 4 :

9

The product of a 1-by-n matrix and an n-by-1 matrix is a 1-by-1 matrix.

However, we will frequently identify a 1-by-1 matrix with its entry.

3.46 **Example** （ 3 4 ） 6 D （ 26 ） because 3 . 6 C 4 . 2 D 26.

However, we can identify （ 26 ） with 26, writing （ 3 4 ） 6 D 26.

2

2

Our next result gives another way to think of matrix multiplication: the entry in row j , column k, of AC equals (row j of A) times (column k of C ).

3.47 Entry of matrix product equals row times column

Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then

.AC/j;k D Aj; C ;k

for 1 三 j 三 m and 1 三 k 三 p.

The proof of the result above follows immediately from the deﬁnitions.

3.48 **Example** The result above and Example 3.46 show why the entry in row 2, column 1, of the product in Example 3.42 equals 26.

SECTION 3.C Matrices **77**

The next result gives yet another way to think of matrix multiplication. It states that column k of AC equals A times column k of C .

3.49 Column of matrix product equals matrix times column

Suppose A is an m-by-n matrix and C is an n-by-p matrix. Then

.AC/ ;k D AC ;k

for 1 三 k 三 p.

Again, the proof of the result above follows immediately from the deﬁni- tions and is left to the reader.

* 1. **Example** From the result above and the equation

0 1 2 1 5 0 7 1

1

@ A

@ A

3 4

5 6

D 19 ;

31

we see why column 2 in the matrix product in Example 3.42 is the right side of the equation above.

We give one more way of thinking about the product of an m-by-n matrix and an n-by-1 matrix. The following example illustrates this approach.

* 1. **Example** In the example above, the product of a 3-by-2 matrix and a 2-by-1 matrix is a linear combination of the columns of the 3-by-2 matrix,

with the scalars that multiply the columns coming from the 2-by-1 matrix.

Speciﬁcally, 0 7 1

0 1 1

0 2 1

@ 19 A D 5 @ 3 A C 1 @ 4 A :

31

5

6

The next result generalizes the example above. Again, the proof follows easily from the deﬁnitions and is left to the reader.

3.52 Linear combination of columns

Suppose A is an m-by-

n matrix and c D

0

c

1

@

:

cn

1

Then

CA is an n-by-1 matrix.

Ac D c1A ;1 C ... C cnA ;n:

In other words, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.

**78** CHAPTER 3 Linear Maps

Two more ways to think about matrix multiplication are given by Exercises 10 and 11.

EXERCISES 3.C

2 *L*

1. Suppose V and W are ﬁnite-dimensional and T .V; W /. Show that with respect to each choice of bases of V and W, the matrix of T has at least dim range T nonzero entries.
2. Suppose D 3.**R**/; 2.**R**/ is the differentiation map deﬁned by Dp p0. Find a basis of 3.**R**/ and a basis of 2.**R**/ such that the matrix of D with respect to these bases is

（ ）

D *P P*

2 *L P P*

@0 1A :

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 |

[*Compare the exercise above to Example 3.34. The next exercise generalizes the exercise above.*]

1. Suppose V and W are ﬁnite-dimensional and T .V; W /. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of *M*.T / are 0 except that the entries in row j , column j , equal 1 for 1 三 j 三 dim range T.

2 *L*

1. Suppose *v*1;:::; *v*m is a basis of V and W is ﬁnite-dimensional. Suppose T .V; W /. Prove that there exists a basis *w*1;:::; *w*n of W such that all the entries in the ﬁrst column of .T / (with respect to the bases *v*1;:::; *v*m and *w*1;:::; *w*n) are 0 except for possibly a 1 in the ﬁrst row,

*M*

2 *L*

ﬁrst column.

[*In this exercise, unlike Exercise 3, you are given the basis of* V *instead of being able to choose a basis of* V*.*]

1. Suppose *w*1;:::; *w*n is a basis of W and V is ﬁnite-dimensional. Suppose T .V; W /. Prove that there exists a basis *v*1;:::; *v*m of V such that all the entries in the ﬁrst row of .T / (with respect to the bases *v*1;:::; *v*m and *w*1;:::; *w*n) are 0 except for possibly a 1 in the ﬁrst row,

*M*

2 *L*

ﬁrst column.

[*In this exercise, unlike Exercise 3, you are given the basis of* W *instead of being able to choose a basis of* W*.*]

SECTION 3.C Matrices **79**

1. Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Prove that dim range T D 1 if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of *M*.T / equal 1.
2. Verify 3.36.
3. Verify 3.38.
4. Prove 3.52.
5. Suppose A is an m-by-n matrix and C is an n-by-p matrix. Prove that

.AC/j; D Aj; C

for 1 j m. In other words, show that row j of AC equals (row j of A) times C .

三 三

（ ）...D

1. Suppose a a1 an is a 1-by-n matrix and C is an n-by-p

matrix. Prove that

aC D a1C1; C ... C anCn; :

In other words, show that aC is a linear combination of the rows of C , with the scalars that multiply the rows coming from a.

1. Give an example with 2-by-2 matrices to show that matrix multiplication is not commutative. In other words, ﬁnd 2-by-2 matrices A and C such that AC ¤ CA.
2. Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, C , D, E, and F are matrices whose sizes are such that A.B C C/ and .D C E/F make sense. Prove that AB C AC and DF C EF both make sense and that

A.B C C/ D AB C AC and .D C E/F D DF C EF .

1. Prove that matrix multiplication is associative. In other words, suppose A, B, and C are matrices whose sizes are such that .AB/C makes sense. Prove that A.BC / makes sense and that .AB/C D A.BC /.
2. Suppose A is an n-by-n matrix and 1 j; k n. Show that the entry in row j , column k, of A3 (which is deﬁned to mean AAA) is

三 三

Xn Xn

Aj;pAp;r Ar;k :

pD1 r D1

**80** CHAPTER 3 Linear Maps

## *Invertibility and Isomorphic Vector*Spaces

3.D

### Invertible Linear Maps

We begin this section by deﬁning the notions of invertible and inverse in the context of linear maps.

3.53 **Deﬁnition *invertible, inverse***

* A linear map T 2 *L*.V; W / is called ***invertible*** if there exists a

linear map S 2 *L*.W; V / such that ST equals the identity map on

V and TS equals the identity map on W.

* A linear map S 2 *L*.W; V / satisfying ST D I and TS D I is

called an ***inverse*** of T (note that the ﬁrst I is the identity map on V

and the second I is the identity map on W ).

3.54 Inverse is unique

An invertible linear map has a unique inverse.

Proof Suppose T .V; W / is invertible and S1 and S2 are inverses of T. Then

2 *L*

S1 D S1I D S1.TS2/ D .S1T /S2 D IS2 D S2:

Thus S1 D S2.

Now that we know that the inverse is unique, we can give it a notation.

3.55 **Notation**

T 1

If T is invertible, then its inverse is denoted by T 1. In other words, if

such that T 1T D I and TT 1 D I.

T 2 *L*.V; W / is invertible, then T 1 is the unique element of *L*.W; V /

The following result characterizes the invertible linear maps.

3.56 Invertibility is equivalent to injectivity and surjectivity

A linear map is invertible if and only if it is injective and surjective.

SECTION 3.D Invertibility and Isomorphic Vector Spaces **81**

Proof Suppose T .V; W /. We need to show that T is invertible if and only if it is injective and surjective.

2 *L*

First suppose T is invertible. To show that T is injective, suppose u; *v* 2 V

and Tu D T *v*. Then

u D T 1.T u/ D T 1.T *v*/ D *v*;

so u *v*. Hence T is injective.

D

We are still assuming that T is invertible. Now we want to prove that T is surjective. To do this, let *w* W. Then *w* T .T 1*w*/, which shows that *w* is in the range of T. Thus range T W. Hence T is surjective, completing this

D

2 D

direction of the proof.

Now suppose T is injective and surjective. We want to prove that T is invertible. For each *w* W, deﬁne S *w* to be the unique element of V such that T .S *w*/ *w* (the existence and uniqueness of such an element follow from the surjectivity and injectivity of T ). Clearly T S equals the identity map on W.

ı

D

2

To prove that S ı T equals the identity map on V, let *v* 2 V. Then

（ ）

T .S ı T /*v* D .T ı S/.T *v*/ D I.T *v*/ D T *v*:

This equation implies that .S T /*v v* (because T is injective). Thus S T

* D ı

equals the identity map on V.

To complete the proof, we need to show that S is linear. To do this, suppose

*w*1, *w*2 2 W. Then

T .S *w*1 C S *w*2/ D T .S *w*1/ C T .S *w*2/ D *w*1 C *w*2:

Thus S *w*1 S *w*2 is the unique element of V that T maps to *w*1 *w*2. By the deﬁnition of S, this implies that S.*w*1 *w*2/ S *w*1 S *w*2. Hence S satisﬁes the additive property required for linearity.

C D C

C C

The proof of homogeneity is similar. Speciﬁcally, if *w* W and 入 **F**,

2 2

then

T .入S *w*/ D 入T .S *w*/ D 入*w*:

Thus 入S *w* is the unique element of V that T maps to 入*w*. By the deﬁnition of

S, this implies that S.入*w*/ D 入S *w*. Hence S is linear, as desired.

3.57 **Example *linear maps that are not invertible***

The multiplication by x2 linear map from .**R**/ to .**R**/ (see 3.4) is not invertible because it is not surjective (1 is not in the range).

* *P P*

The backward shift linear map from **F**1 to **F**1 (see 3.4) is not invertible because it is not injective [.1; 0; 0; 0; : : : / is in the null space].

•

**82** CHAPTER 3 Linear Maps

### Isomorphic Vector Spaces

The next deﬁnition captures the idea of two vector spaces that are essentially the same, except for the names of the elements of the vector spaces.

3.58 **Deﬁnition *isomorphism, isomorphic***

* An ***isomorphism*** is an invertible linear map.
* Two vector spaces are called ***isomorphic*** if there is an isomorphism

from one vector space onto the other one.

Think of an isomorphism T V W as relabeling *v* V as T *v* W. This viewpoint explains why two isomorphic vector spaces have the same vector

space properties. The terms “isomorphism” and “invertible linear map” mean

W ! 2 2

the same thing. Use “isomorphism" when you want to emphasize that the two spaces are essentially the same.

*The Greek word* ***isos*** *means equal; the Greek word* ***morph*** *means shape. Thus* ***isomorphic*** *literally means equal shape.*

3.59 Dimension shows whether vector spaces are isomorphic

Two ﬁnite-dimensional vector spaces over **F** are isomorphic if and only if they have the same dimension.

Proof First suppose V and W are isomorphic ﬁnite-dimensional vector spaces. Thus there exists an isomorphism T from V onto W. Because T is invertible, we have null T D f0g and range T D W. Thus dim null T D 0 and dim range T D dim W. The formula

dim V D dim null T C dim range T

(the Fundamental Theorem of Linear Maps, which is 3.22) thus becomes the equation dim V dim W, completing the proof in one direction.

D

To prove the other direction, suppose V and W are ﬁnite-dimensional

vector spaces with the same dimension. Let *v*1;:::; *v*n be a basis of V and

*w*1;:::; *w*n be a basis of W. Let T 2 *L*.V; W / be deﬁned by

T .c1*v*1 C ... C cn*v*n/ D c1*w*1 C ... C cn*w*n:

Then T is a well-deﬁned linear map because *v*1;:::; *v*n is a basis of V (see 3.5). Also, T is surjective because *w*1;:::; *w*n spans W. Furthermore, null T 0 because *w*1;:::; *w*n is linearly independent; thus T is injective. Because T is injective and surjective, it is an isomorphism (see 3.56). Hence V and W are isomorphic, as desired.

D f g

SECTION 3.D Invertibility and Isomorphic Vector Spaces **83**

The previous result implies that each ﬁnite-dimensional vector space V is iso- morphic to **F**n, where n dim V.

*Because every ﬁnite-dimensional vector space is isomorphic to some* **F**n*, why not just study* **F**n *instead of more general vector spaces? To an- swer this question, note that an in- vestigation of* **F**n *would soon lead to other vector spaces. For exam- ple, we would encounter the null space and range of linear maps. Al- though each of these vector spaces is isomorphic to some* **F**n*, thinking of them that way often adds com- plexity but no new insight.*

D

If *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is a basis of W, then for each T .V; W /, we have a matrix

2 *L*

m;n

*M*.T / 2 **F** . In other words, once

bases have been ﬁxed for V and W, becomes a function from .V; W /

*M L*

to **F**m;n. Notice that 3.36 and 3.38 show

that is a linear map. This linear map is actually invertible, as we now show.

*M*

3.60 *L*.V; W / and **F**m;n are isomorphic

Suppose *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is a basis of W. Then *M* is an isomorphism between *L*.V; W / and **F**m;n.

Proof We already noted that is linear. We need to prove that is injec- tive and surjective. Both are easy. We begin with injectivity. If T .V; W / and .T / 0, then T *v*k 0 for k 1; :::; n. Because *v*1;:::; *v*n is a basis of V, this implies T 0. Thus is injective (by 3.16).

To prove that is surjective, suppose A **F**m;n. Let T be the linear map

*M* 2

D *M*

*M* D D D

2 *L*

*M M*

from V to W such that

X

T *v*k D

m

j D1

Aj;k *w*j

for k D 1; :::;n (see 3.5). Obviously *M*.T / equals A, and thus the range of

*M* equals **F**m;n, as desired.

Now we can determine the dimension of the vector space of linear maps from one ﬁnite-dimensional vector space to another.

3.61 dim *L*.V; W / D .dim V /.dim W/

Suppose V and W are ﬁnite-dimensional. Then *L*.V; W / is ﬁnite-

dimensional and

dim *L*.V; W / D .dim V /.dim W /:

Proof This follows from 3.60, 3.59, and 3.40.

**84** CHAPTER 3 Linear Maps

### Linear Maps Thought of as Matrix Multiplication

Previously we deﬁned the matrix of a linear map. Now we deﬁne the matrix of a vector.

3.62 **Deﬁnition *matrix of a vector,*** *M*.*v*/

Suppose *v* 2 V and *v*1;:::; *v*n is a basis of V. The ***matrix of*** *v* with

respect to this basis is the n-by-1 matrix

*M*.*v*/ D @ A ;

c

1

:

cn

where c1;:::; cn are the scalars such that

*v* D c1*v*1 C ... C cn*v*n:

The matrix .*v*/ of a vector *v* V depends on the basis *v*1;:::; *v*n of V, as well as on *v*. However, the basis should be clear from the context and thus it is not included in the notation.

3.63 **Example *matrix of a vector***

* The matrix of 2 — 7x C 5x3 with respect to the standard basis of *P*3.**R**/

*M* 2

is

0 2 1

B —7 C :

@

0

5

A

The matrix of a vector x **F**n with respect to the standard basis is obtained by writing the coordinates of x as the entries in an n-by-1

* 2

matrix. In other words, if x D .x1;:::; xn/ 2 **F**n, then

A

0

x1

B

*M*.x/ D : xn

@

1C :

Occasionally we want to think of elements of V as relabeled to be n-by-1 matrices. Once a basis *v*1;:::; *v*n is chosen, the function *M* that takes *v* 2 V to *M*.*v*/ is an isomorphism of V onto **F**n;1 that implements this relabeling.

SECTION 3.D Invertibility and Isomorphic Vector Spaces **85**

Recall that if A is an m-by-n matrix, then A ;k denotes the kth column of A, thought of as an m-by-1 matrix. In the next result, .*v*k/ is computed with respect to the basis *w*1;:::; *w*m of W.

*M*

3.64 *M*.T / ;k D *M*.*v*k/.

a basis of W. Let 1 三 k 三 n. Then the kth column of *M*.T /, which is

Suppose T 2 *L*.V; W / and *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is

denoted by *M*.T / ;k, equals *M*.*v*k/.

Proof The desired result follows immediately from the deﬁnitions of *M*.T /

and *M*.*v*k/.

The next result shows how the notions of the matrix of a linear map, the matrix of a vector, and matrix multiplication ﬁt together.

3.65 Linear maps act like matrix multiplication

Suppose T 2 *L*.V; W / and *v* 2 V. Suppose *v*1;:::; *v*n is a basis of V and

*w*1;:::; *w*m is a basis of W. Then

*M*.T *v*/ D *M*.T /*M*.*v*/:

Proof Suppose *v* D c1*v*1 C ... C cn*v*n, where c1;:::; cn 2 **F**. Thus

**3.66** T *v* D c1T *v*1 C ... C cnT *v*n:

Hence

*M*.T *v*/ D c1*M*.T *v*1/ C ... C cn*M*.T *v*n/ D c1*M*.T / ;1 C ... C cn*M*.T / ;n D *M*.T /*M*.*v*/;

where the ﬁrst equality follows from 3.66 and the linearity of , the second equality comes from 3.64, and the last equality comes from 3.52.

*M*

Each m-by-n matrix A induces a linear map from **F**n;1 to **F**m;1, namely the matrix multiplication function that takes x **F**n;1 to Ax **F**m;1. The result

2 2

above can be used to think of every linear map (from one ﬁnite-dimensional vector space to another ﬁnite-dimensional vector space) as a matrix multi- plication map after suitable relabeling via the isomorphisms given by *M*.

Speciﬁcally, if T 2 *L*.V; W / and we identify *v* 2 V with *M*.*v*/ 2 **F**n;1, then the result above says that we can identify T *v* with *M*.T /*M*.*v*/.

**86** CHAPTER 3 Linear Maps

Because the result above allows us to think (via isomorphisms) of each linear map as multiplication on **F**n;1 by some matrix A, keep in mind that the speciﬁc matrix A depends not only on the linear map but also on the choice

of bases. One of the themes of many of the most important results in later chapters will be the choice of a basis that makes the matrix A as simple as possible.

In this book, we concentrate on linear maps rather than on matrices. How- ever, sometimes thinking of linear maps as matrices (or thinking of matrices as linear maps) gives important insights that we will ﬁnd useful.

### Operators

Linear maps from a vector space to itself are so important that they get a special name and special notation.

3.67 **Deﬁnition *operator,*** *L*.V /

* A linear map from a vector space to itself is called an ***operator***.
* The notation *L*.V / denotes the set of all operators on V. In other

words, *L*.V / D *L*.V; V /.

A linear map is invertible if it is injective and surjective. For an op- erator, you might wonder whether in- jectivity alone, or surjectivity alone, is enough to imply invertibility. On

inﬁnite-dimensional vector spaces, neither condition alone implies invert- ibility, as illustrated by the next example, which uses two familiar operators from Example 3.4.

*The deepest and most important parts of linear algebra, as well as most of the rest of this book, deal with operators.*

3.68 **Example *neither injectivity nor surjectivity implies invertibility***

* The multiplication by x2 operator on *P*.**R**/ is injective but not surjective.
* The backward shift operator on **F**1 is surjective but not injective.

In view of the example above, the next result is remarkable—it states that for operators on a ﬁnite-dimensional vector space, either injectivity or surjectivity alone implies the other condition. Often it is easier to check that an operator on a ﬁnite-dimensional vector space is injective, and then we get surjectivity for free.

SECTION 3.D Invertibility and Isomorphic Vector Spaces **87**

3.69 Injectivity is equivalent to surjectivity in ﬁnite dimensions

Suppose V is ﬁnite-dimensional and T 2 *L*.V /. Then the following are

equivalent:

1. T is invertible;
2. T is injective;
3. T is surjective.

Proof Clearly (a) implies (b).

Now suppose (b) holds, so that T is injective. Thus null T 0 (by 3.16).

D f g

From the Fundamental Theorem of Linear Maps (3.22) we have

dim range T D dim V — dim null T

D dim V:

Thus range T equals V. Thus T is surjective. Hence (b) implies (c).

Now suppose (c) holds, so that T is surjective. Thus range T V. From

D

the Fundamental Theorem of Linear Maps (3.22) we have

dim null T D dim V — dim range T

D 0:

Thus null T equals 0 . Thus T is injective (by 3.16), and so T is invertible (we already knew that T was surjective). Hence (c) implies (a), completing the proof.

f g

The next example illustrates the power of the previous result. Although it is possible to prove the result in the example below without using linear algebra, the proof using linear algebra is cleaner and easier.

polynomial p 2 *P*.**R**/ with

.x2 C 5x C 7/p 00 D q.

3.70 **Example** Show tha（t for each polyno）mial q 2 *P*.**R**/, there exists a

Solution Example 3.68 shows that the magic of 3.69 does not apply to the inﬁnite-dimensional vector space .**R**/. However, each nonzero polynomial q has some degree m. By restricting attention to m.**R**/, we can work with a

*P*

*P*

ﬁnite-dimensional vector space.

Suppose q 2 *P*m.**R**/. Deﬁne T W *P*m.**R**/ ! *P*m.**R**/ by

Tp D （.x2 C 5x C 7/p）00:

**88** CHAPTER 3 Linear Maps

Multiplying a nonzero polynomial by .x2 5x 7/ increases the degree by 2, and then differentiating twice reduces the degree by 2. Thus T is indeed an operator on m.**R**/.

C C

*P*

Every polynomial whose second derivative equals 0 is of the form ax b, where a; b **R**. Thus null T 0 . Hence T is injective.

2 D f g

C

Now 3.69 implies that T is surjective. Thus there exists a polynomial

（ ）

p 2 *P*m.**R**/ such that .x2 C 5x C 7/p 00 D q, as desired.

Exercise 30 in Section 6.A gives a similar but more spectacular application of 3.69. The result in that exercise is quite difﬁcult to prove without using linear algebra.

EXERCISES 3.D

1. Suppose T 2 *L*.U; V / and S 2 *L*.V; W / are both invertible linear maps. Prove that ST 2 *L*.U; W / is invertible and that .ST / 1 D T 1S 1.
2. Suppose V is ﬁnite-dimensional and dim V > 1. Prove that the set of noninvertible operators on V is not a subspace of *L*.V /.
3. Suppose V is ﬁnite-dimensional, U is a subspace of V, and S 2 *L*.U; V /. Prove there exists an invertible operator T 2 *L*.V / such that Tu D Su for every u 2 U if and only if S is injective.
4. Suppose W is ﬁnite-dimensional and T1; T2 2 *L*.V; W /. Prove that null T1 D null T2 if and only if there exists an invertible operator S 2 *L*.W / such that T1 D S T2.
5. Suppose V is ﬁnite-dimensional and T1; T2 2 *L*.V; W /. Prove that range T1 D range T2 if and only if there exists an invertible operator S 2 *L*.V / such that T1 D T2S.
6. Suppose V and W are ﬁnite-dimensional and T1; T2 2 *L*.V; W /. Prove that there exist invertible operators R 2 *L*.V / and S 2 *L*.W / such that T1 D S T2R if and only if dim null T1 D dim null T2.
7. Suppose V and W are ﬁnite-dimensional. Let *v* 2 V. Let

E D fT 2 *L*.V; W / W T *v* D 0g:

* 1. Show that E is a subspace of *L*.V; W /.
  2. Suppose *v* ¤ 0. What is dim E?

SECTION 3.D Invertibility and Isomorphic Vector Spaces **89**

1. Suppose V is ﬁnite-dimensional and T V W is a surjective linear map of V onto W. Prove that there is a subspace U of V such that T jU is an isomorphism of U onto W. (Here T jU means the function T restricted to U. In other words, T jU is the function whose domain is U, with T jU deﬁned by T jU .u/ D Tu for every u 2 U.)

W !

1. Suppose V is ﬁnite-dimensional and S; T .V /. Prove that ST is invertible if and only if both S and T are invertible.

2 *L*

1. Suppose V is ﬁnite-dimensional and S; T 2 *L*.V /. Prove that ST D I

if and only if TS D I.

1. Suppose V is ﬁnite-dimensional and S; T; U 2 *L*.V / and ST U D I. Show that T is invertible and that T 1 D US.
2. Show that the result in the previous exercise can fail without the hypoth- esis that V is ﬁnite-dimensional.
3. Suppose V is a ﬁnite-dimensional vector space and R; S; T .V / are such that RST is surjective. Prove that S is injective.

2 *L*

1. Suppose *v*1;:::; *v*n is a basis of V. Prove that the map T V **F**n;1

W !

deﬁned by

T *v* D *M*.*v*/

is an isomorphism of V onto **F**n;1; here .*v*/ is the matrix of *v* V

*M* 2

with respect to the basis *v*1;:::; *v*n.

1. Prove that every linear map from **F**n;1 to **F**m;1 is given by a matrix multiplication. In other words, prove that if T 2 *L*.**F**n;1; **F**m;1/, then there exists an m-by-n matrix A such that Tx D Ax for every x 2 **F**n;1.
2. Suppose V is ﬁnite-dimensional and T 2 *L*.V /. Prove that T is a scalar multiple of the identity if and only if ST D TS for every S 2 *L*.V /.
3. Suppose V is ﬁnite-dimensional and *E* is a subspace of *L*.V / such that ST 2 *E* and TS 2 *E* for all S 2 *L*.V / and all T 2 *E* . Prove that *E* D f0g or *E* D *L*.V /.
4. Show that V and *L*.**F**;V / are isomorphic vector spaces.

（ ）

1. Suppose T 2 *L P*.**R**/ is such that T is injective and deg Tp 三 deg p for every nonzero polynomial p 2 *P*.**R**/.
   1. Prove that T is surjective.
   2. Prove that deg Tp D deg p for every nonzero p 2 *P*.**R**/.

**90** CHAPTER 3 Linear Maps

**20** Suppose n is a positive integer and Ai;j **F** for i; j 1; :::; n. Prove that the following are equivalent (note that in both parts below, the number of equations equals the number of variables):

2 D

1. The trivial solution x1 xn 0 is the only solution to the homogeneous system of equations

D ... D D

kD1

Xn

n kD1

X

A1;kxk D 0

:

An;kxk D 0:

1. For every c1;:::; cn **F**, there exists a solution to the system of equations

2

n kD1

X

n kD1

X

A1;kxk D c1

:

An;kxk D cn:

SECTION 3.E Products and Quotients of Vector Spaces **91**

## *Products and Quotients of Vector Spaces*

3.E

### Products of Vector Spaces

As usual when dealing with more than one vector space, all the vector spaces in use should be over the same ﬁeld.

3.71 **Deﬁnition *product of vector spaces***

Suppose V1;:::; Vm are vector spaces over **F**.

* The ***product*** V1 x ... x Vm is deﬁned by

V1 x ... x Vm D f.*v*1;:::; *v*m/ W *v*1 2 V1;:::; *v*m 2 Vmg:

* Addition on V1 x ... x Vm is deﬁned by

.u1;:::; um/ C .*v*1;:::; *v*m/ D .u1 C *v*1;:::; um C *v*m/:

* Scalar multiplication on V1 x ... x Vm is deﬁned by

入.*v*1;:::; *v*m/ D .入*v*1;:::; 入*v*m/:

3.72 **Example** Elements of 2.**R**/ **R**3 are lists of length 2, with the ﬁrst item in the list an element of 2.**R**/ and the second item in the list an element of **R**3.

For example, 5 — 6x C 4x2; .3; 8; 7/ 2 *P*2.**R**/ x **R**3.

（ ）

*P*

*P* x

The next result should be interpreted to mean that the product of vector spaces is a vector space with the operations of addition and scalar multiplica- tion as deﬁned above.

3.73 Product of vector spaces is a vector space

vector space over **F**.

Suppose V1;:::; Vm are vector spaces over **F**. Then V1 x. . .x Vm is a

The proof of the result above is left to the reader. Note that the additive identity of V1 x ... x Vm is .0; : : : ; 0/, where the 0 in the j th slot is the additive identity of Vj . The additive inverse of .*v*1;:::; *v*m/ 2 V1 x ... x Vm is .—*v*1;:::; —*v*m/.

**92** CHAPTER 3 Linear Maps

* 1. **Example** Is **R**2 x **R**3 equal to **R**5? Is **R**2 x **R**3 isomorphic to **R**5?

Solution Elements of **R**2 x **R**3 are lists （.x1; x2/; .x3; x4; x5/）, where

x ; x ; x ; x ; x 2 **R**.1 2 3 4 5

Elements of **R**5 are lists .x1; x2; x3; x4; x5/, where x1; x2; x3; x4; x5 **R**. Although these look almost the same, they are not the same kind of object.

2

Elements of **R**2 **R**3 are lists of length 2 (with the ﬁrst item itself a list of length 2 and the second item a list of length 3), and elements of **R**5 are lists

x

of length 5. Thus **R**2 x **R**3 does not eq（ual **R**5. ）

The linear map that takes a vector

.x1; x2/; .x3; x4; x5/

2 **R**2 x **R**3 to

2 x

.x1; x2; x3; x4; x5/ **R**5 is clearly an isomorphism of **R**2 **R**3 onto **R**5.

Thus these two vector spaces are isomorphic.

In this case, the isomorphism is so natural that we should think of it as a relabeling. Some people would even informally say that **R**2 **R**3 equals **R**5, which is not technically correct but which captures the spirit of identiﬁcation via relabeling.

x

The next example illustrates the idea of the proof of 3.76.

* 1. **Example** Find a basis of *P*2.**R**/ x **R**2.

Solution Consider this list of length 5 of elements of *P*2.**R**/ x **R**2:

（1; .0; 0/）; （x; .0; 0/）; （x2; .0; 0/）; （0; .1; 0/）; （0; .0; 1/）:

The list above is linearly independent and it spans *P*2.**R**/ x **R**2. Thus it is a basis of *P*2.**R**/ x **R**2.

3.76 Dimension of a product is the sum of dimensions

Suppose V1;:::; Vm are ﬁnite-dimensional vector spaces. Then

V1 x ... x Vm is ﬁnite-dimensional and

dim.V1 x ... x Vm/ D dim V1 C ... C dim Vm:

Proof Choose a basis of each Vj . For each basis vector of each Vj , consider the element of V1 Vm that equals the basis vector in the j th slot and 0 in the other slots. The list of all such vectors is linearly independent and spans V1 x. . . x Vm. Thus it is a basis of V1 x. . . x Vm. The length of this basis is dim V1 C ... C dim Vm, as desired.

x. . . x

SECTION 3.E Products and Quotients of Vector Spaces **93**

### Products and Direct Sums

In the next result, the map r is surjective by the deﬁnition of U1 Um. Thus the last word in the result below could be changed from “injective” to “invertible”.

C. . .C

3.77 Products and direct sums

Suppose that U1;:::; Um are subspaces of V. Deﬁne a linear map

r W U1 x ... x Um ! U1 C ... C Um by

r.u1;:::; um/ D u1 C ... C um:

Then U1 C ... C Um is a direct sum if and only if r is injective.

Proof The linear map r is injective if and only if the only way to write 0 as a sum u1 um, where each uj is in Uj , is by taking each uj equal to 0. Thus 1.44 shows that r is injective if and only if U1 Um is a direct

sum, as desired.

C ... C

C. . . C

3.78 A sum is a direct sum if and only if dimensions add up

Suppose V is ﬁnite-dimensional and U1;:::; Um are subspaces of V. Then

U1 C ... C Um is a direct sum if and only if

dim.U1 C ... C Um/ D dim U1 C ... C dim Um:

Proof The map r in 3.77 is surjective. Thus by the Fundamental Theorem of Linear Maps (3.22), r is injective if and only if

dim.U1 C ... C Um/ D dim.U1 x ... x Um/:

Combining 3.77 and 3.76 now shows that U1 Um is a direct sum if and only if

C ... C

as desired.

dim.U1 C ... C Um/ D dim U1 C ... C dim Um;

In the special case m 2, an alternative proof that U1 U2 is a direct sum if and only if dim.U1 U2/ dim U1 dim U2 can be obtained by combining 1.45 and 2.43.

C D C

D C

**94** CHAPTER 3 Linear Maps

### Quotients of Vector Spaces

We begin our approach to quotient spaces by deﬁning the sum of a vector and a subspace.

3.79 **Deﬁnition** *v* C U

deﬁned by

Suppose *v* 2 V and U is a subspace of V. Then *v* C U is the subset of V

*v* C U D f*v* C u W u 2 U g:

3.80 **Example** Suppose

U D f.x; 2x/ 2 **R**2 W x 2 **R**g:

20

*U*

*(*17, 20*)* + *U*

10

17

Then U is the line in **R**2 through the origin with slope 2. Thus

.17; 20/ C U

is the line in **R**2 that contains the point

.17; 20/ and has slope 2.

*(*10, 20*)*

*(*17, 20*)*

3.81 **Deﬁnition *afﬁne subset, parallel***

* An ***afﬁne subset*** of V is a subset of V of the form *v* C U for some

*v* 2 V and some subspace U of V.

* For *v* 2 V and U a subspace of V, the afﬁne subset *v* C U is said to

be ***parallel*** to U.

3.82 **Example *parallel afﬁne subsets***

* In Example 3.80 above, all the lines in **R**2 with slope 2 are parallel to U. If U .x; y; 0/ **R**3 x; y **R** , then the afﬁne subsets of **R**3
* D f 2 W 2 g

parallel to U are the planes in **R**3 that are parallel to the xy-plane U in

the usual sense.

**Important:** With the deﬁnition of ***parallel*** given in 3.81, no line in **R**3 is considered to be an afﬁne subset that is parallel to the plane U.

SECTION 3.E Products and Quotients of Vector Spaces **95**

3.83 **Deﬁnition *quotient space,*** V=U

Suppose U is a subspace of V. Then the ***quotient space*** V=U is the set of all afﬁne subsets of V parallel to U. In other words,

V=U D f*v* C U W *v* 2 V g:

3.84 **Example *quotient spaces***

If U .x; 2x/ **R**2 x **R** , then **R**2=U is the set of all lines in

* D f 2 W 2 g

**R**2 that have slope 2.

If U is a line in **R**3 containing the origin, then **R**3=U is the set of all lines in **R**3 parallel to U.

•

If U is a plane in **R**3 containing the origin, then **R**3=U is the set of all planes in **R**3 parallel to U.

•

Our next goal is to make V=U into a vector space. To do this, we will need the following result.

3.85 Two afﬁne subsets parallel to U are equal or disjoint

Suppose U is a subspace of V and *v*; *w* 2 V. Then the following are

equivalent:

1. *v* — *w* 2 U ;
2. *v* C U D *w* C U ;
3. .*v* C U/ \ .*w* C U/ ¤ ¿.

Proof First suppose (a) holds, so *v* — *w* 2 U. If u 2 U, then

*v* C u D *w* C .*v* — *w*/ C u 2 *w* C U:

（ ）

Thus *v* U *w* U. Similarly, *w* U *v* U. Thus *v* U *w* U, completing the proof that (a) implies (b).

C c C C c C C D C

Obviously (b) implies (c).

Now suppose (c) holds, so .*v* C U/ \ .*w* C U/ ¤ ¿. Thus there exist

u1; u2 2 U such that

*v* C u1 D *w* C u2:

Thus *v w* u2 u1. Hence *v w* U, showing that (c) implies (a) and completing the proof.

— D — — 2

**96** CHAPTER 3 Linear Maps

Now we can deﬁne addition and scalar multiplication on V=U.

3.86 **Deﬁnition *addition and scalar multiplication on*** V=U

Suppose U is a subspace of V. Then ***addition*** and ***scalar multiplication***

are deﬁned on V=U by

.*v* C U/ C .*w* C U/ D .*v* C *w*/ C U 入.*v* C U/ D .入*v*/ C U

for *v*; *w* 2 V and 入 2 **F**.

As part of the proof of the next result, we will show that the deﬁnitions above make sense.

3.87 Quotient space is a vector space

Suppose U is a subspace of V. Then V=U, with the operations of addition and scalar multiplication as deﬁned above, is a vector space.

Proof The potential problem with the deﬁnitions above of addition and scalar multiplication on V=U is that the representation of an afﬁne subset parallel to U is not unique. Speciﬁcally, suppose *v*; *w* 2 V. Suppose also that *v*O; *w*O 2 V

are such that *v* C U D *v*O C U and *w* C U D *w*O C U. To show that the

deﬁnition of addition on V=U given above makes sense, we must show that

.*v w*/ U .*v w*/ U.

C C D O C O C

By 3.85, we have

*v* — *v*O 2 U and *w* — *w*O 2 U:

Because U is a subspace of V and thus is closed under addition, this implies that .*v* — *v*O/ C .*w* — *w*O / 2 U. Thus .*v* C *w*/ — .*v*O C *w*O / 2 U. Using 3.85 again,

we see that

.*v* C *w*/ C U D .*v*O C *w*O / C U;

as desired. Thus the deﬁnition of addition on V=U makes sense.

Similarly, suppose 入 **F**. Because U is a subspace of V and thus is closed under scalar multiplication, we have 入.*v v*/ U. Thus 入*v* 入*v* U. Hence 3.85 implies that .入*v*/ U .入*v*/ U. Thus the deﬁnition of scalar multiplication on V=U makes sense.

C D O C

— O 2 — O 2

2

Now that addition and scalar multiplication have been deﬁned on V=U, the veriﬁcation that these operations make V=U into a vector space is straightfor- ward and is left to the reader. Note that the additive identity of V=U is 0 C U (which equals U ) and that the additive inverse of *v* C U is .—*v*/ C U.

SECTION 3.E Products and Quotients of Vector Spaces **97**

The next concept will give us an easy way to compute the dimension of V=U.

3.88 **Deﬁnition *quotient map,*** 兀

Suppose U is a subspace of V. The ***quotient map*** 兀 is the linear map

兀W V ! V=U deﬁned by

兀.*v*/ D *v* C U

for *v* 2 V.

The reader should verify that 兀 is indeed a linear map. Although 兀depends on U as well as V, these spaces are left out of the notation because they should be clear from the context.

3.89 Dimension of a quotient space

Suppose V is ﬁnite-dimensional and U is a subspace of V. Then

dim V=U D dim V — dim U:

Proof Let 兀 be the quotient map from V to V=U. From 3.85, we see that null 兀 U. Clearly range 兀 V=U. The Fundamental Theorem of Linear Maps (3.22) thus tells us that

D D

dim V D dim U C dim V=U;

which gives the desired result.

Each linear map T on V induces a linear map TQ on V=.null T /, which we now deﬁne.

3.90 **Deﬁnition**

TQ

Suppose T 2 *L*.V; W /. Deﬁne TQ W V=.null T/ ! W by

TQ .*v* C null T/ D T *v*:

To show that the deﬁnition of TQ makes sense, suppose u; *v* 2 V are such that u C null T D *v* C null T. By 3.85, we have u — *v* 2 null T. Thus

T .u *v*/ 0. Hence Tu T *v*. Thus the deﬁnition of TQ

— D D

sense.

indeed makes

**98** CHAPTER 3 Linear Maps

3.91 Null space and range of TQ

Suppose T 2 *L*.V; W /. Then

1. TQ is a linear map from V=.null T/ to W ;
2. TQ is injective;
3. range TQ D range T ;
4. V=.null T/ is isomorphic to range T.

Proof

1. The routine veriﬁcation that TQ is linear is left to the reader.
2. Suppose *v* 2 V and TQ .*v* Cnull T/ D 0. Then T *v* D 0. Thus *v* 2 null T. Hence 3.85 implies that *v* C null T D 0 C null T. This implies that null TQ D 0, and hence TQ is injective, as desired.
3. The deﬁnition of TQ shows that range TQ D range T.
4. Parts (b) and (c) imply that if we think of TQ as mapping into range T, then TQ is an isomorphism from V=.null T/ onto range T.

EXERCISES 3.E

* 1. Suppose T is a function from V to W. The ***graph*** of T is the subset of

V x W deﬁned by

graph of T D f.*v*;T *v*/ 2 V x W W *v* 2 V g:

Prove that T is a linear map if and only if the graph of T is a subspace of V W.

x

[*Formally, a function* T *from* V *to* W *is a subset* T *of* V W *such that for each v* V*, there exists exactly one element* .*v*; *w*/ T*. In other*

2 2

x

*words, formally a function is what is called above its graph. We do not usually think of functions in this formal manner. However, if we do become formal, then the exercise above could be rephrased as follows:*

*Prove that a function* T *from* V *to* W *is a linear map if and only if* T *is a subspace of* V x W*.*]

SECTION 3.E Products and Quotients of Vector Spaces **99**

* 1. Suppose V1;:::; Vm are vector spaces such that V1 x. x Vm is ﬁnite-

dimensional. Prove that Vj is ﬁnite-dimensional for each j D 1; :::; m.

* 1. Give an example of a vector space V and subspaces U1; U2 of V such that U1 x U2 is isomorphic to U1 C U2 but U1 C U2 is not a direct sum.

*Hint:* The vector space V must be inﬁnite-dimensional.

* 1. Suppose V1;:::; Vm are vector spaces. Prove that *L*.V1 x x Vm;W /

and *L*.V1;W / x x *L*.Vm;W / are isomorphic vector spaces.

* 1. Suppose W1;:::; Wm are vector spaces. Prove that *L*.V; W1 x x Wm/

and *L*.V; W1/ x x *L*.V; Wm/ are isomorphic vector spaces.

* 1. For n a positive integer, deﬁne V n by

V n D V x . . . x V :

„ ƒ‚ …

n times

Prove that V n and *L*.**F**n;V / are isomorphic vector spaces.

* 1. Suppose *v*;x are vectors in V and U; W are subspaces of V such that

*v* C U D x C W. Prove that U D W.

* 1. Prove that a nonempty subset A of V is an afﬁne subset of V if and only if 入*v* C .1 — 入/*w* 2 A for all *v*; *w* 2 A and all 入 2 **F**.
  2. Suppose A1 and A2 are afﬁne subsets of V. Prove that the intersection

A1 \ A2 is either an afﬁne subset of V or the empty set.

* 1. Prove that the intersection of every collection of afﬁne subsets of V is either an afﬁne subset of V or the empty set.
  2. Suppose *v*1;:::; *v*m 2 V. Let

A D f入1*v*1 C ... C 入m*v*m W 入1;:::; 入m 2 **F** and 入1 C C 入m D 1g:

* + 1. Prove that A is an afﬁne subset of V.
    2. Prove that every afﬁne subset of V that contains *v*1;:::; *v*m also contains A.
    3. Prove that A D *v* C U for some *v* 2 V and some subspace U of

V with dim U 三 m — 1.

* 1. Suppose U is a subspace of V such that V=U is ﬁnite-dimensional. Prove that V is isomorphic to U x .V=U /.

**100** CHAPTER 3 Linear Maps

1. Suppose U is a subspace of V and *v*1 U; :::; *v*m U is a basis of V=U and u1;:::; un is a basis of U. Prove that *v*1;:::; *v*m; u1;:::; un is a basis of V.

C C

1. Suppose U D f.x1; x2;:::/ 2 **F**1 W xj ¤ 0 for only ﬁnitely many jg.
   1. Show that U is a subspace of **F**1.
   2. Prove that **F**1=U is inﬁnite-dimensional.
2. Suppose ' 2 *L*.V; **F**/ and ' ¤ 0. Prove that dim V=.null '/ D 1.
3. Suppose U is a subspace of V such that dim V=U D 1. Prove that there exists ' 2 *L*.V; **F**/ such that null ' D U.
4. Suppose U is a subspace of V such that V=U is ﬁnite-dimensional. Prove that there exists a subspace W of V such that dim W D dim V=U and V D U ˚ W.
5. Suppose T 2 *L*.V; W / and U is a subspace of V. Let 兀 denote the quotient map from V onto V=U. Prove that there exists S 2 *L*.V=U; W / such that T D S ı 兀 if and only if U c null T.
6. Find a correct statement analogous to 3.78 that is applicable to ﬁnite sets, with unions analogous to sums of subspaces and disjoint unions analogous to direct sums.
7. Suppose U is a subspace of V. Deﬁne r W *L*.V=U; W / ! *L*.V; W / by

r.S/ D S ı 兀:

1. Show that r is a linear map.
2. Show that r is injective.
3. Show that range r D fT 2 *L*.V; W / W Tu D 0 for every u 2 U g.

SECTION 3.F Duality **101**

## *Duality*

3.F

### The Dual Space and the Dual Map

Linear maps into the scalar ﬁeld **F** play a special role in linear algebra, and thus they get a special name:

3.92 **Deﬁnition *linear functional***

A ***linear functional*** on V is a linear map from V to **F**. In other words, a linear functional is an element of *L*.V; **F**/.

3.93 **Example *linear functionals***

Deﬁne ' **R**3 **R** by '.x; y; z/ 4x 5y 2z. Then ' is a linear functional on **R**3.

* W ! D — C
* Fix .c1;:::; cn/ 2 **F**n. Deﬁne ' W **F**n ! **F** by

'.x1;:::; xn/ D c1x1 C ... C cnxn:

Then ' is a linear functional on **F**n.

* Deﬁne ' W *P*.**R**/ ! **R** by '.p/ D 3p00.5/ C 7p.4/. Then ' is a linear functional on *P*.**R**/.
* Deﬁne ' W *P*.**R**/ ! **R** by '.p/ D R 1 p.x/ dx. Then ' is a linear

0

functional on *P*.**R**/.

The vector space *L*.V; **F**/ also gets a special name and special notation:

3.94 **Deﬁnition *dual space,*** V 0

The ***dual space*** of V, denoted V 0, is the vector space of all linear functionals on V. In other words, V 0 D *L*.V; **F**/.

3.95 dim V 0 D dim V

Suppose V is ﬁnite-dimensional. Then V 0 is also ﬁnite-dimensional and dim V 0 D dim V.

Proof This result follows from 3.61.

**102** CHAPTER 3 Linear Maps

In the following deﬁnition, 3.5 implies that each 'j is well deﬁned.

3.96 **Deﬁnition *dual basis***

If *v*1;:::; *v*n is a basis of V, then the ***dual basis*** of *v*1;:::; *v*n is the list '1;:::; 'n of elements of V 0, where each 'j is the linear functional on V such that

'j .*v*k/ D

(

1 if k D j;

0 if k ¤ j:

3.97 **Example** What is the dual basis of the standard basis e1;:::; en

of **F**n?

Solution For 1 j n, deﬁne 'j to be the linear functional on **F**n that selects the j th coordinate of a vector in **F**n. In other words,

三 三

'j .x1;:::; xn/ D xj

for .x1;:::; xn/ 2 **F**n. Clearly

(D

'j .ek

/ 1 if k D j;

0 if k ¤ j:

Thus '1;:::; 'n is the dual basis of the standard basis e1;:::; en of **F**n.

The next result shows that the dual basis is indeed a basis. Thus the terminology “dual basis” is justiﬁed.

3.98 Dual basis is a basis of the dual space

Suppose V is ﬁnite-dimensional. Then the dual basis of a basis of V is a basis of V 0.

Proof Suppose *v*1;:::; *v*n is a basis of V. Let '1;:::; 'n denote the dual basis.

To show that '1;:::; 'n is a linearly independent list of elements of V 0, suppose a1;:::; an 2 F are such that

a1'1 C ... C an'n D 0:

Now .a1'1 an'n/.*v*j / aj for j 1; :::; n. The equation above thus shows that a1 an 0. Hence '1;:::; 'n is linearly independent.

D ... D D

C . . . C D D

Now 2.39 and 3.95 imply that '1;:::; 'n is a basis of V 0.

SECTION 3.F Duality **103**

In the deﬁnition below, note that if T is a linear map from V to W then T 0

is a linear map from W 0 to V 0.

3.99 **Deﬁnition *dual map,*** T 0

deﬁned by T 0.'/ D ' ı T for ' 2 W 0.

If T 2 *L*.V; W /, then the ***dual map*** of T is the linear map T 0 2 *L*.W 0;V 0/

If T .V; W / and ' W 0, then T 0.'/ is deﬁned above to be the composition of the linear maps ' and T. Thus T 0.'/ is indeed a linear map from V to **F**; in other words, T 0.'/ V 0.

The veriﬁcation that T 0 is a linear map from W 0 to V 0 is easy:

2

2 *L* 2

* If '; 2 W 0, then

T 0.' C / D .' C / ı T D ' ı T C ı T D T 0.'/ C T 0. /:

* If 入 2 **F** and ' 2 W 0, then

T 0.入'/ D .入'/ ı T D 入.' ı T/ D 入T 0.'/:

In the next example, the prime notation is used with two unrelated mean- ings: D0 denotes the dual of a linear map D, and p0 denotes the derivative of a polynomial p.

3.100 **Example** Deﬁne D W *P*.**R**/ ! *P*.**R**/ by Dp D p0.

* Suppose ' is the linear functional on *P*.**R**/ deﬁned by '.p/ D p.3/. Then D0.'/ is the linear functional on *P*.**R**/ given by

（ ）

D0.'/ .p/ D .' ı D/.p/ D '.Dp/ D '.p0/ D p0.3/:

In other words, D0.'/ is the linear functional on .**R**/ that takes p to

*P*

p0.3/.

* Suppose ' is the linear functional on *P*.**R**/ deﬁned by '.p/ D R 1 p.

0

Then D0.'/ is the linear functional on *P*.**R**/ given by

D .'/ .p/ D .'ıD/.p/ D '.Dp/ D '.p / D

p

D p.1/—p.0/:

（ 0 ）

0 Z 1 0

0

In other words, D0.'/ is the linear functional on *P*.**R**/ that takes p to

p.1/ — p.0/.

**104** CHAPTER 3 Linear Maps

The ﬁrst two bullet points in the result below imply that the function that takes T to T 0 is a linear map from .V; W / to .W 0;V 0/.

*L L*

In the third bullet point below, note the reversal of order from ST on the

left to T 0S 0 on the right (here we assume that U is a vector space over **F**).

3.101 Algebraic properties of dual maps

* .S C T /0 D S 0 C T 0 for all S; T 2 *L*.V; W /.
* .入T /0 D 入T 0 for all 入 2 **F** and all T 2 *L*.V; W /.
* .ST /0 D T 0S 0 for all T 2 *L*.U; V / and all S 2 *L*.V; W /.

Proof The proofs of the ﬁrst two bullet points above are left to the reader.

To prove the third bullet point, suppose ' 2 W 0. Then

（ ）

.ST /0.'/ D ' ı.ST / D .' ıS/ıT D T 0.' ıS/ D T 0 S 0.'/ D .T 0S 0/.'/;

where the ﬁrst, third, and fourth equal- ities above hold because of the deﬁni- tion of the dual map, the second equality holds because composition of functions is associative, and the last equality fol- lows from the deﬁnition of composition.

*Some books use the notation* V \* *and* T \* *for duality instead of* V 0 *and* T 0*. However, here we reserve the notation* T \* *for the adjoint,*

*which will be introduced when we study linear maps on inner product spaces in Chapter 7.*

The equality of the ﬁrst and last terms above for all ' 2 W 0 means that

.ST /0 D T 0S 0.

### The Null Space and Range of the Dual of a Linear Map

Our goal in this subsection is to describe null T 0 and range T 0 in terms of range T and null T. To do this, we will need the following deﬁnition.

3.102 **Deﬁnition *annihilator,*** U 0

For U c V, the ***annihilator*** of U, denoted U 0, is deﬁned by

U 0 D f' 2 V 0 W '.u/ D 0 for all u 2 U g:

* 1. **Example** Suppose U is the subspace of *P*.**R**/ consisting of all polynomial multiples of x2. If ' is the linear functional on *P*.**R**/ deﬁned by '.p/ D p0.0/, then ' 2 U 0.

SECTION 3.F Duality **105**

For U V, the annihilator U 0 is a subset of the dual space V 0. Thus U 0 depends on the vector space containing U, so a notation such as U 0 would be more precise. However, the containing vector space will always be clear from

c

V

the context, so we will use the simpler notation U 0.

* 1. **Example** Let e1; e2; e3; e4; e5 denote the standard basis of **R**5, and let '1; '2; '3; '4; '5 denote the dual basis of .**R**5/0. Suppose

U D span.e1; e2/ D f.x1; x2; 0; 0; 0/ 2 **R**5 W x1; x2 2 **R**g:

Show that U 0 D span.'3; '4; '5/.

Solution Recall (see 3.97) that 'j is the linear functional on **R**5 that selects that j th coordinate: 'j .x1; x2; x3; x4; x5/ D xj .

First suppose ' 2 span.'3; '4; '5/. Then there exist c3; c4; c5 2 **R** such that ' D c3'3 C c4'4 C c5'5. If .x1; x2; 0; 0; 0/ 2 U, then

'.x1; x2; 0; 0; 0/ D .c3'3 C c4'4 C c5'5/.x1; x2; 0; 0; 0/ D 0:

Thus ' U 0. In other words, we have shown that span.'3; '4; '5/ U 0.

2 c

To show the inclusion in the other direction, suppose ' U 0. Because the dual basis is a basis of .**R**5/0, there exist c1; c2; c3; c4; c5 2 **R** such that ' D c1'1 C c2'2 C c3'3 C c4'4 C c5'5. Because e1 2 U and ' 2 U 0, we

2

have

0 D '.e1/ D .c1'1 C c2'2 C c3'3 C c4'4 C c5'5/.e1/ D c1:

Similarly, e2 2 U and thus c2 D 0. Hence ' D c3'3 C c4'4 C c5'5. Thus

' 2 span.'3; '4; '5/, which shows that U 0 c span.'3; '4; '5/.

3.105 The annihilator is a subspace

Suppose U c V. Then U 0 is a subspace of V 0.

Proof Clearly 0 U 0 (here 0 is the zero linear functional on V ), because the zero linear functional applied to every vector in U is 0.

2

Suppose '; 2 U 0. Thus '; 2 V 0 and '.u/ D .u/ D 0 for every u 2 U. If u 2 U, then .' C /.u/ D '.u/ C .u/ D 0 C 0 D 0. Thus ' U

C 2

0.

Similarly, U 0 is closed under scalar multiplication. Thus 1.34 implies that

U 0 is a subspace of V 0.

**106** CHAPTER 3 Linear Maps

The next result shows that dim U 0 is the difference of dim V and dim U. For example, this shows that if U is a 2-dimensional subspace of **R**5, then U 0 is a 3-dimensional subspace of .**R**5/0, as in Example 3.104.

The next result can be proved following the pattern of Example 3.104: choose a basis u1;:::; um of U, extend to a basis u1;:::; um;:::; un of V, let '1;:::; 'm;:::; 'n be the dual basis of V 0, and then show 'mC1;:::; 'n is a basis of U 0, which implies the desired result.

You should construct the proof outlined in the paragraph above, even though a slicker proof is presented here.

3.106 Dimension of the annihilator

Suppose V is ﬁnite-dimensional and U is a subspace of V. Then

dim U C dim U 0 D dim V:

Proof Let i .U; V / be the inclusion map deﬁned by i.u/ u for u U. Thus i is a linear map from V 0 to U 0. The Fundamental Theorem of Linear

Maps (3.22) applied to i0 shows that

0 2 *L* D 2

dim range i0 C dim null i0 D dim V 0:

However, null i0 D U 0 (as can be seen by thinking about the deﬁnitions) and dim V 0 D dim V (by 3.95), so we can rewrite the equation above as

dim range i0 C dim U 0 D dim V:

If ' U 0, then ' can be extended to a linear functional on V (see, for example, Exercise 11 in Section 3.A). The deﬁnition of i0 shows that i0. / '. Thus ' range i0, which implies that range i0 U 0. Hence dim range i0 dim U 0 dim U, and the displayed equation above becomes

2

D D

D 2 D

the desired result.

The proof of part (a) of the result below does not use the hypothesis that

V and W are ﬁnite-dimensional.

3.107 The null space of T 0

Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Then

1. null T 0 D .range T /0;
2. dim null T 0 D dim null T C dim W — dim V.

SECTION 3.F Duality **107**

Proof

1. First suppose ' 2 null T 0. Thus 0 D T 0.'/ D ' ı T. Hence

0 D .' ı T /.*v*/ D '.T *v*/ for every *v* 2 V:

Thus ' 2 .range T /0. This implies that null T 0 c .range T /0.

To prove the inclusion in the opposite direction, now suppose that ' 2 .range T /0. Thus '.T *v*/ D 0 for every vector *v* 2 V. Hence 0 D ' ı T D T 0.'/. In other words, ' 2 null T 0, which shows that

.range T /0 c null T 0, completing the proof of (a).

1. We have

dim null T 0 D dim.range T /0

D dim W — dim range T

D dim W — .dim V — dim null T/

D dim null T C dim W — dim V;

where the ﬁrst equality comes from (a), the second equality comes from 3.106, and the third equality comes from the Fundamental Theorem of Linear Maps (3.22).

The next result can be useful because sometimes it is easier to verify that

T 0 is injective than to show directly that T is surjective.

3.108 T surjective is equivalent to T 0 injective

Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Then T is

surjective if and only if T 0 is injective.

Proof The map T 2 *L*.V; W / is surjective if and only if range T D W, which happens if and only if .range T /0 D f0g, which happens if and only if null T 0 D f0g [by 3.107(a)], which happens if and only if T 0 is injective.

3.109 The range of T 0

Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Then

1. dim range T 0 D dim range T ;
2. range T 0 D .null T /0.

**108** CHAPTER 3 Linear Maps

Proof

1. We have

dim range T 0 D dim W 0 — dim null T 0

0

D dim W — dim.range T/

D dim range T;

where the ﬁrst equality comes from the Fundamental Theorem of Linear Maps (3.22), the second equality comes from 3.95 and 3.107(a), and the third equality comes from 3.106.

1. First suppose ' 2 range T 0. Thus there exists 2 W 0 such that

' D T 0. /. If *v* 2 null T, then

'.*v*/ D （T 0. /）*v* D . ı T /.*v*/ D .T *v*/ D .0/ D 0:

Hence ' 2 .null T /0. This implies that range T 0 c .null T /0.

We will complete the proof by showing that range T 0 and .null T /0

have the same dimension. To do this, note that

dim range T 0 D dim range T

D dim V — dim null T

D dim.null T/ ;

0

where the ﬁrst equality comes from (a), the second equality comes from the Fundamental Theorem of Linear Maps (3.22), and the third equality comes from 3.106.

The next result should be compared to 3.108.

3.110 T injective is equivalent to T 0 surjective

Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Then T is

injective if and only if T 0 is surjective.

Proof The map T 2 *L*.V; W / is injective if and only if null T D f0g, which happens if and only if .null T /0 D V 0, which happens if and only if range T 0 D V 0 [by 3.109(b)], which happens if and only if T 0 is surjective.

SECTION 3.F Duality **109**

### The Matrix of the Dual of a Linear Map

We now deﬁne the transpose of a matrix.

0@ 5 —7 1A

3.111 **Deﬁnition *transpose,*** At

The ***transpose*** of a matrix A, denoted At, is the matrix obtained from A by interchanging the rows and columns. More speciﬁcally, if A is an m-by-n matrix, then At is the n-by-m matrix whose entries are given by

the equation

.At/k;j D Aj;k:

3.112 **Example** If A D

3 8

—4 2

, then A

t

t 5 3 —4

D

—7 8 2

.

Note that here A is a 3-by-2 matrix and A is a 2-by-3 matrix.

The transpose has nice algebraic properties: .A C /t At C t and

.入A/t 入At for all m-by-n matrices A; C and all 入 **F** (see Exercise 33).

D 2

C D C

The next result shows that the transpose of the product of two matrices is

the product of the transposes in the opposite order.

3.113 The transpose of the product of matrices

If A is an m-by-n matrix and C is an n-by-p matrix, then

.AC/t D C tAt:

Proof Suppose 1 三 k 三 p and 1 三 j 三 m. Then

（.AC/t）

k;j

D .AC/j;k

Xn

D

D

r 1

X

n

D

r D1

Aj;r Cr;k

.Ct/k;r .At/r;j

D .CtAt/k;j :

Thus .AC/t D C tAt, as desired.

**110** CHAPTER 3 Linear Maps

The setting for the next result is the assumption that we have a basis *v*1;:::; *v*n of V, along with its dual basis '1;:::; 'n of V 0. We also have a basis *w*1;:::; *w*m of W, along with its dual basis 1;:::; m of W 0. Thus

.T / is computed with respect to the bases just mentioned of V and W, and .T 0/ is computed with respect to the dual bases just mentioned of W 0 and V .

*M*0

*M*

3.114 The matrix of T 0 is the transpose of the matrix of T

Suppose T 2 *L*.V; W /. Then *M*.T 0/ D （*M*.T /）t.

Proof Let A D *M*.T / and C D *M*.T 0/. Suppose 1 三 j 三 m and

1 三 k 三 n. 0

From the deﬁnition of *M*.T / we have

T 0. j / D

n

r D1

X

Cr;j 'r :

The left side of the equation above equals j T. Thus applying both sides of the equation above to *v*k gives

ı

. j ı T /.*v*k/ D

n

r D1

X

Cr;j 'r .*v*k/

We also have

D Ck;j :

. j ı T /.*v*k/ D j .T *v*k/

Xm

D j

Xm

D

Ar;k j .*w*r /

r D1

Ar;k *w*r

D Aj;k:

r D1

Comparing the last line of the last two sets o（ f equat）ions, we have Ck;j D Aj;k.

t

Thus C D At. In other words, *M*.T 0/ D

*M*.T / , as desired.

### The Rank of a Matrix

SECTION 3.F Duality **111**

We begin by deﬁning two nonnegative integers that are associated with each matrix.

3.115 **Deﬁnition *row rank, column rank***

Suppose A is an m-by-n matrix with entries in **F**.

* The ***row rank*** of A is the dimension of the span of the rows of A in

**F**1;n.

* The ***column rank*** of A is the dimension of the span of the columns

of A in **F**m;1.

3.116 **Example** Suppose A 4 7 1 8 . Find the row rank of A

3 5 2 9

D

and the column rank of A.

Solution The row rank of A is the dimension of

span （ 4 7 1 8 ） ; （ 3 5 2 9 ）

in **F**1;4. Neither of the two vectors listed above in **F**1;4 is a scalar multiple of the other. Thus the span of this list of length 2 has dimension 2. In other words, the row rank of A is 2.

The column rank of A is the dimension of

span 4 ; 7 ; 1 ; 8 !

3

5

2

9

in **F**2;1. Neither of the ﬁrst two vectors listed above in **F**2;1 is a scalar multiple of the other. Thus the span of this list of length 4 has dimension at least 2. The span of this list of vectors in **F**2;1 cannot have dimension larger than 2 because dim **F**2;1 2. Thus the span of this list has dimension 2. In other words, the column rank of A is 2.

D

*M*

*M*

Notice that no bases are in sight in the statement of the next result. Al- though .T / in the next result depends on a choice of bases of V and W, the next result shows that the column rank of .T / is the same for all such choices (because range T does not depend on a choice of basis).

**112** CHAPTER 3 Linear Maps

3.117 Dimension of range T equals column rank of *M*.T /

Suppose V and W are ﬁnite-dimensional and T 2 *L*.V; W /. Then

dim range T equals the column rank of *M*.T /.

Proof Suppose *v*1;:::; *v*n is a basis of V and *w*1;:::; *w*m is a basis of W. The function that takes *w* span.T *v*1;:::;T *v*n/ to .*w*/ is easily seen to be an isomorphism from span.T *v*1;:::;T *v*n/ onto span .T *v*1/; :::; .T *v*n/ . Thus dim span.T *v*1;:::;T *v*n/ dim span .T *v*1/; :::; .T *v*n/ , where the last dimension equals the column rank of .T /.

It is easy to see that range T span.T *v*1;:::;T *v*n/. Thus we have dim range T dim span.T *v*1;:::;T *v*n/ the column rank of .T /, as

（ ）

（ ）

D D *M*

D

*M*

D *M M*

*M M*

2 *M*

desired.

In Example 3.116, the row rank and column rank turned out to equal each other. The next result shows that this always happens.

3.118 Row rank equals column rank

Suppose A 2 **F**m;n. Then the row rank of A equals the column rank of A.

Proof Deﬁne T W **F**n;1 ! **F**m;1 by Tx D Ax. Thus *M*.T / D A, where

*M*.T / is computed with respect to the standard bases of **F**n;1 and **F**m;1. Now

column rank of A D column rank of *M*.T /

D dim range T

D dim range T 0

D column rank of *M*.T 0/

D column rank of At

D row rank of A;

where the second equality above comes from 3.117, the third equality comes from 3.109(a), the fourth equality comes from 3.117 (where .T 0/ is com- puted with respect to the dual bases of the standard bases), the ﬁfth equality

*M*

comes from 3.114, and the last equality follows easily from the deﬁnitions.

The last result allows us to dispense with the terms “row rank” and “column rank” and just use the simpler term “rank”.

3.119 **Deﬁnition *rank***

The ***rank*** of a matrix A 2 **F**m;n is the column rank of A.

EXERCISES 3.F

SECTION 3.F Duality **113**

1. Explain why every linear functional is either surjective or the zero map.
2. Give three distinct examples of linear functionals on **R**Œ0;1].
3. Suppose V is ﬁnite-dimensional and *v* 2 V with *v* ¤ 0. Prove that there exists ' 2 V 0 such that '.*v*/ D 1.
4. Suppose V is ﬁnite-dimensional and U is a subspace of V such that U ¤ V. Prove that there exists ' 2 V 0 such that '.u/ D 0 for every u 2 U but ' ¤ 0.
5. Suppose V1;:::; Vm are vector spaces. Prove that .V1 x ... x Vm/0 and

V10 x ... x Vm0 are isomorphic vector spaces.

1. Suppose V is ﬁnite-dimensional and *v*1;:::; *v*m 2 V. Deﬁne a linear map r W V 0 ! **F**m by

（ ）

r.'/ D '.*v*1/; :::; '.*v*m/ :

* 1. Prove that *v*1;:::; *v*m spans V if and only if r is injective.
  2. Prove that *v*1;:::; *v*m is linearly independent if and only if r is surjective.

1. Suppose m is a positive integer. Show that the dual basis of the basis

*.j/*

1; x; :::; xm of *P* .**R**/ is ' ;' ;:::;' , where ' .p/ D p .0/ . Here

m

0

1

m

j

jŠ

p.j / denotes the j th derivative of p, with the understanding that the 0th

derivative of p is p.

1. Suppose m is a positive integer.
   1. Show that 1; x — 5; : : : ; .x — 5/m is a basis of *P*m.**R**/.
   2. What is the dual basis of the basis in part (a)?
2. Suppose *v*1;:::; *v*n is a basis of V and '1;:::; 'n is the corresponding dual basis of V 0. Suppose 2 V 0. Prove that

D .*v*1/'1 C ... C .*v*n/'n:

1. Prove the ﬁrst two bullet points in 3.101.

**114** CHAPTER 3 Linear Maps

1. Suppose A is an m-by-n matrix with A ¤ 0. Prove that the rank of A is 1 if and only if there exist .c1;:::; cm/ 2 **F**m and .d1;:::; dn/ 2 **F**n such that Aj;k D cj dk for every j D 1; :::;m and every k D 1; :::; n.
2. Show that the dual map of the identity map on V is the identity map on V 0.
3. Deﬁne T **R**3 **R**2 by T .x; y; z/ .4x 5y 6z; 7x 8y 9z/. Suppose '1; '2 denotes the dual basis of the standard basis of **R**2 and

W ! D C C C C

1; 2; 3 denotes the dual basis of the standard basis of **R**3.

* 1. Describe the linear functionals T 0.'1/ and T 0.'2/.
  2. Write T 0.'1/ and T 0.'2/ as linear combinations of 1; 2; 3.

1. Deﬁne T W *P*.**R**/ ! *P*.**R**/ by .Tp/.x/ D x2p.x/ C p00.x/ for x 2 **R**.
   1. Suppose ' 2 *P*.**R**/0 is deﬁned by '.p/ D p0.4/. Describe the linear functional T 0.'/ on *P*.**R**/.

R 1

（T 0.'/）.x3/. 0

(b) Suppose ' 2 *P*.**R**/0 is deﬁned by '.p/ D p.x/ dx. Evaluate

1. Suppose W is ﬁnite-dimensional and T 2 *L*.V; W /. Prove that T 0 D 0

if and only if T D 0.

1. Suppose V and W are ﬁnite-dimensional. Prove that the map that takes

T 2 *L*.V; W / to T 0 2 *L*.W 0;V 0/ is an isomorphism of *L*.V; W / onto

*L*.W 0;V 0/.

1. Suppose U c V. Explain why U 0 D f' 2 V 0 W U c null 'g.
2. Suppose V is ﬁnite-dimensional and U c V. Show that U D f0g if and only if U 0 D V 0.
3. Suppose V is ﬁnite-dimensional and U is a subspace of V. Show that

U D V if and only if U 0 D f0g.

1. Suppose U and W are subsets of V with U c W. Prove that W 0 c U 0.
2. Suppose V is ﬁnite-dimensional and U and W are subspaces of V with

W 0 c U 0. Prove that U c W.

1. Suppose U; W are subspaces of V. Show that .U C W /0 D U 0 \ W 0.

SECTION 3.F Duality **115**

1. Suppose V is ﬁnite-dimensional and U and W are subspaces of V. Prove that .U \ W /0 D U 0 C W 0.
2. Prove 3.106 using the ideas sketched in the discussion before the state- ment of 3.106.
3. Suppose V is ﬁnite-dimensional and U is a subspace of V. Show that

U D f*v* 2 V W '.*v*/ D 0 for every ' 2 U 0g:

1. Suppose V is ﬁnite-dimensional and r is a subspace of V 0. Show that

r D f*v* 2 V W '.*v*/ D 0 for every ' 2 rg0:

（ ）

1. Suppose T 2 *L P*5.**R**/; *P*5.**R**/ and null T 0 D span.'/, where ' is the linear functional on *P*5.**R**/ deﬁned by '.p/ D p.8/. Prove that range T D fp 2 *P*5.**R**/ W p.8/ D 0g.
2. Suppose V and W are ﬁnite-dimensional, T 2 *L*.V; W /, and there exists

' 2 W 0 such that null T 0 D span.'/. Prove that range T D null '.

1. Suppose V and W are ﬁnite-dimensional, T 2 *L*.V; W /, and there exists

' 2 V 0 such that range T 0 D span.'/. Prove that null T D null '.

1. Suppose V is ﬁnite-dimensional and '1;:::; 'm is a linearly independent list in V 0. Prove that

dim（.null '1/ \ \ .null 'm/） D .dim V/ — m:

1. Suppose V is ﬁnite-dimensional and '1;:::; 'n is a basis of V 0. Show that there exists a basis of V whose dual basis is '1;:::; 'n.
2. Suppose T .V /, and u1;:::; un and *v*1;:::; *v*n are bases of V. Prove that the following are equivalent:

2 *L*

* 1. T is invertible.
  2. The columns of *M*.T / are linearly independent in **F**n;1.
  3. The columns of *M*.T / span **F**n;1.
  4. The rows of *M*.T / are linearly independent in **F**1;n.
  5. The rows of *M*.T / span **F**1;n.

Here *M*.T / means *M*（T; .u1;:::; un/; .*v*1;:::; un/）.

**116** CHAPTER 3 Linear Maps

1. Suppose m and n are positive integers. Prove that the function that takes A to At is a linear map from **F**m;n to **F**n;m. Furthermore, prove that this linear map is invertible.
2. The ***double dual space*** of V, denoted V 00, is deﬁned to be the dual space of V 0. In other words, V 00 D .V 0/0. Deﬁne ƒ W V ! V 00 by

.ƒ*v*/.'/ D '.*v*/

for *v* 2 V and ' 2 V 0.

* 1. Show that ƒ is a linear map from V to V 00.
  2. Show that if T 2 *L*.V /, then T 00 ı ƒ D ƒ ı T, where T 00 D .T 0/0.
  3. Show that if V is ﬁnite-dimensional, then ƒ is an isomorphism from V onto V 00.

[*Suppose* V *is ﬁnite-dimensional. Then* V *and* V 0 *are isomorphic, but ﬁnding an isomorphism from* V *onto* V 0 *generally requires choosing a basis of* V*. In contrast, the isomorphism* ƒ *from* V *onto* V 00 *does not*

*require a choice of basis and thus is considered more natural.*]

（ ）

1. Show that *P*.**R**/ 0 and **R**1 are isomorphic.
2. Suppose U is a subspace of V. Let i W U ! V be the inclusion map deﬁned by i.u/ D u. Thus i0 2 *L*.V 0;U 0/.
   1. Show that null i0 D U 0.
   2. Prove that if V is ﬁnite-dimensional, then range i0 D U 0.

e

* 1. Prove that if V is ﬁnite-dimensional, then i0 is an isomorphism from V 0=U 0 onto U 0.

[*The isomorphism in part (c) is natural in that it does not depend on a choice of basis in either vector space.*]

1. Suppose U is a subspace of V. Let 兀W V ! V=U be the usual quotient map. Thus 兀0 2 *L* .V=U /0;V 0 .

（ ）

* 1. Show that 兀0 is injective.
  2. Show that range 兀0 D U 0.
  3. Conclude that 兀0 is an isomorphism from .V=U /0 onto U 0.

[*The isomorphism in part (c) is natural in that it does not depend on a choice of basis in either vector space. In fact, there is no assumption here that any of these vector spaces are ﬁnite-dimensional.*]

# *Polynomials*



*Statue of Persian mathematician and poet Omar Khayyám* (*1048–1131*)*, whose*

*algebra book written in 1070 contained the ﬁrst serious study of cubic polynomials.*

4

CHAPTER

This short chapter contains material on polynomials that we will need to understand operators. Many of the results in this chapter will already be familiar to you from other courses; they are included here for completeness. Because this chapter is not about linear algebra, your instructor may go through it rapidly. You may not be asked to scrutinize all the proofs. Make sure, however, that you at least read and understand the statements of all the

results in this chapter—they will be used in later chapters.

The standing assumption we need for this chapter is as follows:

4.1 **Notation F F** denotes **R** or **C**.

LEARNING OBJECTIVES FOR THIS CHAPTER

Division Algorithm for Polynomials factorization of polynomials over **C** factorization of polynomials over **R**

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

117

DOI 10.1007/978-3-319-11080-6 4

**118** CHAPTER 4 Polynomials

### Complex Conjugate and Absolute Value

Before discussing polynomials with complex or real coefﬁcients, we need to learn a bit more about the complex numbers.

* 1. **Deﬁnition** Re z***,*** Im z

Suppose z D a C bi, where a and b are real numbers.

* + - The ***real part*** of z, denoted Re z, is deﬁned by Re z D a.
    - The ***imaginary part*** of z, denoted Im z, is deﬁned by Im z D b.

Thus for every complex number z, we have

z D Re z C .Im z/i:

* 1. **Deﬁnition *complex conjugate,*** zN***, absolute value,*** jzj

Suppose z 2 **C**.

* + - The ***complex conjugate*** of z 2 **C**, denoted zN, is deﬁned by

zN D Re z — .Im z/i:

* The ***absolute value*** of a complex number z, denoted jzj, is deﬁned

by

jzjD .Re z/ C .Im z/ :

q

2

2

* 1. **Example** Suppose z D 3 C 2i. Then
     + Re z D 3 and Im z D 2;
     + zN D 3 — 2i;
     + jzjD p32 C 22 D p13.

Note that jzj is a nonnegative number for every z 2 **C**.

The real and imaginary parts, com-

*You should verify that* z D zN *if and*

*only if* z *is a real number.*

plex conjugate, and absolute value have the following properties:

CHAPTER 4 Polynomials **119**

4.5 Properties of complex numbers

Suppose *w*;z 2 **C**. Then

**sum of** z **and** zN

z C zN D 2 Re z;

**difference of** z **and** zN

z — zN D 2.Im z/i;

zzN D jzj2;

**additivity and multiplicativity of complex conjugate**

*w* C z D *w*N C zN and *w*z D *w*N zN;

**conjugate of conjugate**

zN D z;

**product of** z **and** zN

**real and imaginary parts are bounded by** jzj

j Re zj三 jzj and j Im zj三 jzj

**absolute value of the complex conjugate**

jzNjD jzj;

**multiplicativity of absolute value**

j*w*zjD j*w*j jzj;

**Triangle Inequality**

j*w* C zj三 j*w*jC jzj.

*z*

*w*

*w* + *z*

Proof Except for the last item, the routine veriﬁcations of the assertions above are left to the reader. To verify the last item, we have

2

j*w* C zj

D .*w* C z/.*w*N C zN/

D *ww*N C zzN C *w*zN C z*w*N

D j*w*j D j*w*j 三 j*w*j D j*w*j

2

2

2

2

2

C jzj C jzj C jzj C jzj

C *w*zN C *w*zN C 2 Re.*w*zN/ C 2j*w*zNj

C 2j*w*j jzj

D .j*w*jC jzj/2:

2

2

2

Taking the square root of both sides of the inequality *w* z 2 . *w* z /2

j C j 三 j jC j j

now gives the desired inequality.

**120** CHAPTER 4 Polynomials

### Uniqueness of Coefﬁcients for Polynomials

Recall that a function p W **F** ! **F** is called a polynomial with coefﬁcients in **F**if there exist a0;:::; am 2 **F** such that

**4.6** p.z/ D a0 C a1z C a2z2 C ... C amzm

for all z 2 **F**.

4.7 If a polynomial is the zero function, then all coefﬁcients are 0

Suppose a0;:::; am 2 **F**. If

a0 C a1z C ... C amz D 0

for every z 2 **F**, then a0 D . . .D am D 0.

m

Proof We will prove the contrapositive. If not all the coefﬁcients are 0, then by changing m we can assume am ¤ 0. Let

z ja0jC ja1jC. . .C jam 1j 1:

D C

jamj

Note that z 1, and thus zj zm 1 for j 0; 1; :::;m 1. Using the Triangle Inequality, we have

乏 三 D —

ja0 C a1z C ... C am 1zm 1j三 .ja0jC ja1jC. . .C jam 1j/zm 1

< jamzmj:

Thus a0 C a1z C . . .C am 1zm 1 ¤ —amzm. Hence we conclude that

a0 C a1z C ... C am 1zm 1 C amzm ¤ 0.

The result above implies that the coefﬁcients of a polynomial are uniquely determined (because if a polynomial had two different sets of coefﬁcients, then subtracting the two representations of the polynomial would give a contradiction to the result above).

Recall that if a polynomial p can be written in the form 4.6 with am ¤ 0, then we say that p has degree m and we write deg p D m.

The degree of the 0 polynomial is deﬁned to be —1. When necessary, use the obvious arithmetic with —1. For example, —1 < m and —1 C m D

*The* 0 *polynomial is declared to*

*tions are not needed for various reasonable results. For example,*

*have degree* —1 *so that excep-*

deg.pq/ D deg p C deg q *even if*

p D 0*.*

—1 for every integer m.

CHAPTER 4 Polynomials **121**

### The Division Algorithm for Polynomials

If p and s are nonnegative integers, with s 0, then there exist nonnegative integers q and r such that

¤

p D sq C r

and r < s. Think of dividing p by s, getting quotient q with remainder r. Our next task is to prove an analogous result for polynomials.

The result below is often called the Division Algorithm for Polynomials, al- though as stated here it is not really an algorithm, just a useful result.

*Think of the Division Algorithm for*

*Polynomials as giving the remain- der* r *when* p *is divided by* s*.*

Recall that .**F**/ denotes the vector space of all polynomials with co- efﬁcients in **F** and that m.**F**/ is the subspace of .**F**/ consisting of the polynomials with coefﬁcients in **F** and degree at most m.

*P P*

*P*

The next result can be proved without linear algebra, but the proof given here using linear algebra is appropriate for a linear algebra textbook.

4.8 Division Algorithm for Polynomials

Suppose that p; s 2 *P*.**F**/, with s ¤ 0. Then there exist unique

polynomials q; r 2 *P*.**F**/ such that

p D sq C r

and deg r < deg s.

Proof Let n D deg p and m D deg s. If n< m, then take q D 0 and r D p

to get the desired result. Thus we can assume that n 乏 m.

Deﬁne T W *P*n m.**F**/ x *P*m 1.**F**/ ! *P*n.**F**/ by

T .q; r/ D sq C r:

The reader can easily verify that T is a linear map. If .q; r/ null T , then sq r 0, which implies that q 0 and r 0 [because otherwise deg sq m and thus sq cannot equal r]. Thus dim null T 0 (proving the

乏 — D

C D D D

2

“unique” part of the result).

From 3.76 we have

（ ）

dim Pn m.**F**/ x *P*m 1.**F**/ D .n — m C 1/ C .m — 1 C 1/ D n C 1:

The Fundamental Theorem of Linear Maps (3.22) and the equation displayed above now imply that dim range T D n C 1, which equals dim *P*n.**F**/. Thus range T D *P*n.**F**/, and hence there exist q 2 *P*n m.**F**/ and r 2 *P*m 1.**F**/ such that p D T .q; r/ D sq C r.

**122** CHAPTER 4 Polynomials

### Zeros of Polynomials

The solutions to the equation p.z/ D 0 play a crucial role in the study of a polynomial p 2 *P*.**F**/. Thus these solutions have a special name.

4.9 **Deﬁnition *zero of a polynomial***

A number 入 2 **F** is called a ***zero*** (or ***root***) of a polynomial p 2 *P*.**F**/ if

p.入/ D 0:

4.10 **Deﬁnition *factor***

A polynomial s 2 *P*.**F**/ is called a ***factor*** of p 2 *P*.**F**/ if there exists a

polynomial q 2 *P*.**F**/ such that p D sq.

We begin by showing that 入 is a zero of a polynomial p 2 *P*.**F**/ if and only if z — 入is a factor of p.

4.11 Each zero of a polynomial corresponds to a degree-1 factor

Suppose p 2 *P*.**F**/ and 入 2 **F**. Then p.入/ D 0 if and only if there is a

polynomial q 2 *P*.**F**/ such that

p.z/ D .z — 入/q.z/

for every z 2 **F**.

Proof One direction is obvious. Namely, suppose there is a polynomial

q 2 *P*.**F**/ such that p.z/ D .z — 入/q.z/ for all z 2 **F**. Then

p.入/ D .入— 入/q.入/ D 0;

as desired.

To prove the other direction, suppose p.入/ 0. The polynomial z 入

D —

has degree 1. Because a polynomial with degree less than 1 is a constant

function, the Division Algorithm for Polynomials (4.8) implies that there exist a polynomial q 2 *P*.**F**/ and a number r 2 **F** such that

p.z/ D .z — 入/q.z/ C r

for every z 2 **F**. The equation above and the equation p.入/ D 0 imply that

r D 0. Thus p.z/ D .z — 入/q.z/ for every z 2 **F**.

Now we can prove that polynomials do not have too many zeros.

CHAPTER 4 Polynomials **123**

4.12 A polynomial has at most as many zeros as its degree

Suppose p 2 *P*.**F**/ is a polynomial with degree m 乏 0. Then p has at

most m distinct zeros in **F**.

Proof If m 0, then p.z/ a0 0 and so p has no zeros.

If m 1, then p.z/ a0 a1z, with a1 0, and thus p has exactly one zero, namely, a0=a1.

—

D D C ¤

D D ¤

Now suppose m > 1. We use induction on m, assuming that every polynomial with degree m 1 has at most m 1 distinct zeros. If p has no zeros in **F**, then we are done. If p has a zero 入 **F**, then by 4.11 there is a polynomial q such that

2

— —

p.z/ D .z — 入/q.z/

for all z **F**. Clearly deg q m 1. The equation above shows that if p.z/ 0, then either z 入 or q.z/ 0. In other words, the zeros of p consist of 入 and the zeros of q. By our induction hypothesis, q has at most m — 1 distinct zeros in **F**. Thus p has at most m distinct zeros in **F**.

D D D

2 D —

### Factorization of Polynomials over C

So far we have been handling polynomials with complex coefﬁcients and polynomials with real coefﬁcients simultaneously through our convention that **F** denotes **R** or **C**. Now we will see some differences between these two cases. First we treat polynomials with complex coefﬁcients. Then we will use our results about polynomials with complex coefﬁcients to prove corresponding results for polynomials with real coefﬁcients.

The next result, although called the Fundamental Theorem of Algebra, uses analysis its proof. The short proof pre- sented here uses tools from complex analysis. If you have not had a course in

*The Fundamental Theorem of Al- gebra is an existence theorem. Its proof does not lead to a method for ﬁnding zeros. The quadratic for-*

*mula gives the zeros explicitly for polynomials of degree* 2*. Similar but more complicated formulas ex- ist for polynomials of degree* 3 *and*

4*. No such formulas exist for poly-*

*nomials of degree* 5 *and above.*

complex analysis, this proof will almost

certainly be meaningless to you. In that case, just accept the Fundamental The- orem of Algebra as something that we need to use but whose proof requires more advanced tools that you may learn in later courses.

**124** CHAPTER 4 Polynomials

4.13 Fundamental Theorem of Algebra

Every nonconstant polynomial with complex coefﬁcients has a zero.

Proof Let p be a nonconstant polynomial with complex coefﬁcients. Sup- pose p has no zeros. Then 1=p is an analytic function on **C**. Furthermore, p.z/ as z , which implies that 1=p 0 as z . Thus 1=p is a bounded analytic function on **C**. By Liouville’s theorem, every such function is constant. But if 1=p is constant, then p is constant, contradicting our assumption that p is nonconstant.

Although the proof given above is probably the shortest proof of the Fundamental Theorem of Algebra, a web search can lead you to several other proofs that use different techniques. All proofs of the Fundamental Theorem of Algebra need to use some analysis, because the result is not true if **C** is

j j!1 j j! 1 ! j j! 1

replaced, for example, with the set of numbers of the form c di where c; d

C

are rational numbers.

Remarkably, mathematicians have

*The cubic formula, which was discovered in the 16th century, is presented below for your amusement only. Do not memorize it.*

*Suppose*

p.x/ D ax C bx C cx C d;

*where* a ¤ 0*. Set*

3

2

u D

9abc — 2b3 — 27a2d

54a3

*and then set*

*v* D u C

2

3ac — b

9a2

2

3

:

*Suppose v* 乏 0*. Then*

— C u C *v* C u — *v*

b

q

*3*

p

q

*3*

3a

p

*is a zero of* p*.*

proved that no formula exists for the ze- ros of polynomials of degree 5 or higher. But computers and calculators can use

clever numerical methods to ﬁnd good approximations to the zeros of any poly- nomial, even when exact zeros cannot be found.

For example, no one will ever be

able to give an exact formula for a zero of the polynomial p deﬁned by

p.x/ D x5—5x4—6x3C17x2C4x—7: However, a computer or symbolic cal- culator can ﬁnd approximate zeros of

this polynomial.

The Fundamental Theorem of Alge-

bra leads to the following factorization

result for polynomials with complex co- efﬁcients. Note that in this factorization,

the numbers 入1;:::; 入m are precisely the zeros of p, for these are the only values of z for which the right side of the equation in the next result equals 0.

CHAPTER 4 Polynomials **125**

4.14 Factorization of a polynomial over **C**

If p 2 *P*.**C**/ is a nonconstant polynomial, then p has a unique factoriza-

tion (except for the order of the factors) of the form

p.z/ D c.z — 入1/ z — 入m/;

where c; 入1;:::; 入m 2 **C**.

Proof Let p .**C**/ and let m deg p. We will use induction on m. If m 1, then clearly the desired factorization exists and is unique. So assume that m > 1 and that the desired factorization exists and is unique for all polynomials of degree m 1.

First we will show that the desired factorization of p exists. By the Fundamental Theorem of Algebra (4.13), p has a zero 入. By 4.11, there is a polynomial q such that

D

—

2 *P* D

p.z/ D .z — 入/q.z/

for all z **C**. Because deg q m 1, our induction hypothesis implies that q has the desired factorization, which when plugged into the equation above gives the desired factorization of p.

2 D —

Now we turn to the question of uniqueness. Clearly c is uniquely deter- mined as the coefﬁcient of zm in p. So we need only show that except for the order, there is only one way to choose 入1;:::; 入m. If

.z — 入1/ ... .z — 入m/ D .z — r1/ z — rm/

for all z 2 **C**, then because the left side of the equation above equals 0 when z D 入1, one of the r’s on the right side equals 入1. Relabeling, we can assume that r1 D 入 1. Now for z ¤ 入 1, we can divide both sides of the equation above by z — 入1, getting

.z — 入2/ ... .z — 入m/ D .z — r2/ z — rm/

for all z **C** except possibly z 入 1. Actually the equation above holds for all z **C**, because otherwise by subtracting the right side from the left side we would get a nonzero polynomial that has inﬁnitely many zeros. The

2

2 D

equation above and our induction hypothesis imply that except for the order, the 入’s are the same as the r’s, completing the proof of uniqueness.

**126** CHAPTER 4 Polynomials

### Factorization of Polynomials over R

A polynomial with real coefﬁcients may

*The failure of the Fundamental Theorem of Algebra for* **R** *accounts for the differences between oper- ators on real and complex vector spaces, as we will see in later chapters.*

have no real zeros. For example, the polynomial 1 x2 has no real zeros.

C

To obtain a factorization theorem

over **R**, we will use our factorization theorem over **C**. We begin with the fol- lowing result.

4.15 Polynomials with real coefﬁcients have zeros in pairs

Suppose p 2 *P*.**C**/ is a polynomial with real coefﬁcients. If 入 2 **C** is a

zero of p, then so is 入N .

Proof Let

p.z/ D a0 C a1z C ... C amzm;

where a0;:::; am are real numbers. Suppose 入 2 **C** is a zero of p. Then

a0 C a1入C ... C am入m D 0:

Take the complex conjugate of both sides of this equation, obtaining

a0 C a1入N C ... C am入N m D 0;

where we have used basic properties of complex conjugation (see 4.5). The equation above shows that 入N is a zero of p.

We want a factorization theorem for polynomials with real coefﬁcients. First

*Think about the connection be- tween the quadratic formula and 4.16.*

we need to characterize the polynomi- als of degree 2 with real coefﬁcients that can be written as the product of two polynomials of degree 1 with real coefﬁcients.

4.16 Factorization of a quadratic polynomial

Suppose b; c 2 **R**. Then there is a polynomial factorization of the form

x2 C bx C c D .x — 入1/.x — 入2/

with 入1; 入2 2 **R** if and only if b2 乏 4c.

CHAPTER 4 Polynomials **127**

Proof Notice that

2

x

b 2

b2

4

First suppose b2 < 4c. Then clearly the right side of the equation above is positive for every x **R**. Hence the

*The equation above is the basis of the technique called* ***completing the square****.*

C C

2

C bx C c D

x C 2

C

c —

:

polynomial x2 bx c has no real

zeros and thus cannot be factored in the form .x — 入1/.x — 入2/ with 入1; 入2 2 **R**.

Conversely, now suppose b2 乏 4c. Then there is a real number d such

that d 2 D b*2* — c. From the displayed equation above, we have

4

x2 C bx C c D x C b 2 — d 2

2

D x

b

C 2 C

d x

b

C 2 —

d ;

which gives the desired factorization.

The next result gives a factorization of a polynomial over **R**. The idea of the proof is to use the factorization 4.14 of p as a polynomial with complex coefﬁcients. Complex but nonreal zeros of p come in pairs; see 4.15. Thus if the factorization of p as an element of *P*.**C**/ includes terms of the form

.x 入/ with 入a nonreal complex number, then .x 入N / is also a term in the

— —

factorization. Multiplying together these two terms, we get

（ ）

x2 — 2.Re 入/x C j入j2 ;

which is a quadratic term of the required form.

The idea sketched in the paragraph above almost provides a proof of the existence of our desired factorization. However, we need to be careful about

one point. Suppose 入is a nonreal complex number and .x — 入/ is a term in the factorization of p as an element of *P*.**C**/. We are guaranteed by 4.15 that

.x 入N / also appears as a term in the factorization, but 4.15 does not state that

—

these two factors appear the same number of times, as needed to make the idea above work. However, the proof works around this point.

In the next result, either m or M may equal 0. The numbers 入1;:::; 入m are precisely the real zeros of p, for these are the only real values of x for which the right side of the equation in the next result equals 0.

**128** CHAPTER 4 Polynomials

4.17 Factorization of a polynomial over **R**

Suppose p 2 *P*.**R**/ is a nonconstant polynomial. Then p has a unique

factorization (except for the order of the factors) of the form

p.x/ D c.x — 入1/ ... .x — 入m/.x C b1x C c1/ ... .x C bM x C cM /;

2

2

where c; 入1;:::; 入m; b1;:::; bM ; c1;:::; cM 2 **R**, with bj 2 < 4cj for

each j .

Proof Think of p as an element of *P*.**C**/. If all the (complex) zeros of p are real, then we are done by 4.14. Thus suppose p has a zero 入 2 **C** with 入 … **R**.

By 4.15, 入N is a zero of p. Thus we can write

p.x/ D .x — 入/.x — 入N /q.x/

（ ）

D x2 — 2.Re 入/x C j入j2 q.x/

for some polynomial q .**C**/ with degree two less than the degree of p. If we can prove that q has real coefﬁcients, then by using induction on the degree of p, we can conclude that .x — 入/ appears in the factorization of p exactly as many times as .x 入N /.

—

2 *P*

To prove that q has real coefﬁcients, we solve the equation above for q,

getting

p.x/

q.x/ D x2 — 2.Re 入/x C j入j2

for all x **R**. The equation above implies that q.x/ **R** for all x **R**. Writing

2 2 2

q.x/ D a0 C a1x C ... C an 2xn 2;

where n D deg p and a0;:::; an 2 2 **C**, we thus have

0 D Im q.x/ D .Im a0/ C .Im a1/x C ... C .Im an 2/xn 2

for all x **R**. This implies that Im a0;:::; Im an 2 all equal 0 (by 4.7). Thus all the coefﬁcients of q are real, as desired. Hence the desired factorization exists.

2

Now we turn to the question of uniqueness of our factorization. A factor of p of the form x2 C bj x C cj with bj 2 < 4cj can be uniquely written as .x — 入j /.x — 入j / with 入j 2 **C**. A moment’s thought shows that two different factorizations of p as an element of *P*.**R**/ would lead to two different factorizations of p as an element of *P*.**C**/, contradicting 4.14.

EXERCISES 4

CHAPTER 4 Polynomials **129**

1. Verify all the assertions in 4.5 except the last one.
2. Suppose m is a positive integer. Is the set

f0g[ fp 2 *P*.**F**/ W deg p D mg

a subspace of *P*.**F**/?

1. Is the set

f0g[ fp 2 *P*.**F**/ W deg p is eveng

a subspace of *P*.**F**/?

1. Suppose m and n are positive integers with m n, and suppose 入 1;:::; 入 m **F**. Prove that there exists a polynomial p .**F**/ with deg p n such that 0 p.入1/ p.入m/ and such that p has no

D D D . . .D

2 2 *P*

三

other zeros.

1. Suppose m is a nonnegative integer, z1;:::; zmC1 are distinct elements of **F**, and *w*1;:::; *w*mC1 2 **F**. Prove that there exists a unique polynomial p 2 *P*m.**F**/ such that

p.zj / D *w*j

for j D 1; :::;m C 1.

[*This result can be proved without using linear algebra. However, try to ﬁnd the clearer, shorter proof that uses some linear algebra.*]

1. Suppose p .**C**/ has degree m. Prove that p has m distinct zeros if and only if p and its derivative p0 have no zeros in common.

2 *P*

1. Prove that every polynomial of odd degree with real coefﬁcients has a real zero.
2. Deﬁne T W *P*.**R**/ ! **RR** by

8ˆ< p — p.3/

if x ¤ 3;

Tp D ˆ:

x — 3

p0.3/ if x D 3:

Show that Tp .**R**/ for every polynomial p .**R**/ and that T is a linear map.

2 *P* 2 *P*

**130** CHAPTER 4 Polynomials

1. Suppose p 2 *P*.**C**/. Deﬁne q W **C** ! **C** by

q.z/ D p.z/p.zN/:

Prove that q is a polynomial with real coefﬁcients.

1. Suppose m is a nonnegative integer and p 2 *P*m.**C**/ is such that there exist distinct real numbers x0; x1;:::; xm such that p.xj / 2 **R** for j D 0; 1; :::; m. Prove that all the coefﬁcients of p are real.
2. Suppose p 2 *P*.**F**/ with p ¤ 0. Let U D fpq W q 2 *P*.**F**/g.
   1. Show that dim *P*.**F**/=U D deg p.
   2. Find a basis of dim *P*.**F**/=U.

# *Eigenvalues, Eigenvectors, and* Invariant Subspaces



*Statue of Italian mathematician Leonardo of Pisa* (*1170–1250, approximate dates*)*, also known as Fibonacci. Exercise 16 in Section 5.C shows how linear algebra can be used to ﬁnd an explicit formula for the Fibonacci sequence.*

5

CHAPTER

Linear maps from one vector space to another vector space were the objects of study in Chapter 3. Now we begin our investigation of linear maps from a ﬁnite-dimensional vector space to itself. Their study constitutes the most important part of linear algebra.

Our standing assumptions are as follows:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a vector space over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

invariant subspaces

eigenvalues, eigenvectors, and eigenspaces

each operator on a ﬁnite-dimensional complex vector space has an eigenvalue and an upper-triangular matrix with respect to some basis

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

131

DOI 10.1007/978-3-319-11080-6 5

**132** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

## *Invariant Subspaces*

5.A

In this chapter we develop the tools that will help us understand the structure of operators. Recall that an operator is a linear map from a vector space to

itself. Recall also that we denote the set of operators on V by .V /; in other words, .V / .V; V /.

*L* D *L*

*L*

Let’s see how we might better understand what an operator looks like.

Suppose T 2 *L*.V /. If we have a direct sum decomposition

V D U1 ˚ ... ˚ Um;

where each Uj is a proper subspace of V, then to understand the behavior of T, we need only understand the behavior of each T jU*j* ; here T jU*j* denotes the restriction of T to the smaller domain Uj . Dealing with T U*j* should be easier than dealing with T because Uj is a smaller vector space than V.

j

However, if we intend to apply tools useful in the study of operators (such as taking powers), then we have a problem: T jU*j* may not map Uj into itself; in other words, T U*j* may not be an operator on Uj . Thus we are led to consider only decompositions of V of the form above where T maps each Uj

j

into itself.

The notion of a subspace that gets mapped into itself is sufﬁciently impor- tant to deserve a name.

5.2 **Deﬁnition *invariant subspace***

Suppose T 2 *L*.V /. A subspace U of V is called ***invariant*** under T if

u 2 U implies Tu 2 U.

In other words, U is invariant under T if T jU is an operator on U.

2 *L*

* 1. **Example** Suppose T .V /. Show that each of the following subspaces of V is invariant under T :

1. f0g;

*The most famous unsolved problem in functional analysis is called the* ***invariant subspace problem****. It deals with invariant subspaces of operators on inﬁnite-dimensional vector spaces.*

1. V ;
2. null T ;
3. range T.

SECTION 5.A Invariant Subspaces **133**

Solution

1. If u 2 f0g, then u D 0 and hence Tu D 0 2 f0g. Thus f0g is invariant

T

under .

1. If u 2 V, then Tu 2 V. Thus V is invariant under T.
2. If u null T, then Tu 0, and hence Tu null T. Thus null T is invariant under T.

2 D 2

1. If u 2 range T, then Tu 2 range T. Thus range T is invariant under T.

Must an operator T .V / have any invariant subspaces other than 0 and V ? Later we will see that this question has an afﬁrmative answer if V is ﬁnite-dimensional and dim V > 1 (for **F C**) or dim V > 2 (for **F R**/;

D D

2 *L* f g

see 5.21 and 9.8.

Although null T and range T are invariant under T, they do not necessarily provide easy answers to the question about the existence of invariant subspaces other than 0 and V , because null T may equal 0 and range T may equal

f g f g

V (this happens when T is invertible).

（ ）

* 1. **Example** Suppose that T 2 *L P*.**R**/ is deﬁned by Tp D p0. Then *P*4.**R**/, which is a subspace of *P*.**R**/, is invariant under T because if p 2 *P*.**R**/ has degree at most 4, then p0 also has degree at most 4.

### Eigenvalues and Eigenvectors

We will return later to a deeper study of invariant subspaces. Now we turn to an investigation of the simplest possible nontrivial invariant subspaces—invariant

subspaces with dimension 1.

Take any *v* V with *v* 0 and let U equal the set of all scalar multiples

2 ¤

of *v*:

U D f入*v* W 入 2 **F**gD span.*v*/:

Then U is a 1-dimensional subspace of V (and every 1-dimensional subspace of V is of this form for an appropriate choice of *v*). If U is invariant under an operator T 2 *L*.V /, then T *v* 2 U, and hence there is a scalar 入 2 **F** such that

T *v* D 入*v*:

Conversely, if T *v* 入*v* for some 入 **F**, then span.*v*/ is a 1-dimensional subspace of V invariant under T.

D 2

**134** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

The equation

T *v* D 入*v*;

which we have just seen is intimately connected with 1-dimensional invariant subspaces, is important enough that the vectors *v* and scalars 入satisfying it are given special names.

5.5 **Deﬁnition *eigenvalue***

Suppose T 2 *L*.V /. A number 入 2 **F** is called an ***eigenvalue*** of T if

there exists *v* 2 V such that *v* ¤ 0 and T *v* D 入*v*.

The comments above show that T has a 1-dimensional invariant subspace if and only if T has an eigenvalue.

*The word* ***eigenvalue*** *is half- German, half-English. The Ger- man adjective* ***eigen*** *means “own” in the sense of characterizing an in- trinsic property. Some mathemati- cians use the term* ***characteristic value*** *instead of eigenvalue.*

In the deﬁnition above, we require that *v* ¤ 0 because every scalar 入 2 **F** satisﬁes T0 D 入0.

5.6 Equivalent conditions to be an eigenvalue

Suppose V is ﬁnite-dimensional, T 2 *L*.V /, and 入 2 F . Then the

following are equivalent:

(a)

(b)

(c)

(d)

入is an eigenvalue of T ;

T — 入I is not injective; T — 入I is not surjective; T — 入I is not invertible.

*Recall that* I 2 *L*.V / *is the iden-*

*tity operator deﬁned by* I *v* D *v for*

*all v* 2 V*.*

Proof Conditions (a) and (b) are equivalent because the equation T *v* 入 *v* is equivalent to the equation .T 入 I /*v* 0. Conditions (b), (c), and (d) are equivalent by 3.69.

— D

D

5.7 **Deﬁnition *eigenvector***

Suppose T 2 *L*.V / and 入 2 **F** is an eigenvalue of T. A vector *v* 2 V is

called an ***eigenvector*** of T corresponding to 入 if *v* ¤ 0 and T *v* D 入*v*.

SECTION 5.A Invariant Subspaces **135**

Because T *v* D 入*v* if and only if .T — 入I /*v* D 0, a vector *v* 2 V with *v* ¤ 0

is an eigenvector of T corresponding to 入 if and only if *v* 2 null.T — 入I /.

* 1. **Example** Suppose T 2 *L*.**F**2/ is deﬁned by

T .*w*; z/ D .—z; *w*/:

1. Find the eigenvalues and eigenvectors of T if **F** D **R**.
2. Find the eigenvalues and eigenvectors of T if **F** D **C**.

Solution

1. If **F R**, then T is a counterclockwise rotation by 90ı about the origin in **R**2. An operator has an eigenvalue if and only if there exists a

D

nonzero vector in its domain that gets sent by the operator to a scalar multiple of itself. A 90ı counterclockwise rotation of a nonzero vector in **R**2 obviously never equals a scalar multiple of itself. Conclusion: if **F** D **R**, then T has no eigenvalues (and thus has no eigenvectors).

1. To ﬁnd eigenvalues of T, we must ﬁnd the scalars 入such that

T .*w*; z/ D 入.*w*; z/

has some solution other than *w* z 0. The equation above is equivalent to the simultaneous equations

D D

* 1. —z D 入*w*; *w* D 入z:

Substituting the value for *w* given by the second equation into the ﬁrst equation gives

2

—z D 入 z:

Now z cannot equal 0 [otherwise 5.9 implies that *w* 0; we are looking for solutions to 5.9 where .*w*; z/ is not the 0 vector], so the equation above leads to the equation

D

—1 D 入2:

The solutions to this equation are 入 i and 入 i. You should be able to verify easily that i and i are eigenvalues of T. Indeed, the eigenvectors corresponding to the eigenvalue i are the vectors of the form .*w*; —*w*i/, with *w* 2 **C** and *w* ¤ 0, and the eigenvectors corresponding to the eigenvalue —i are the vectors of the form .*w*; *w*i/,

—

D D —

with *w* 2 **C** and *w* ¤ 0.

**136** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

Now we show that eigenvectors corresponding to distinct eigenvalues are linearly independent.

5.10 Linearly independent eigenvectors

Let T 2 *L*.V /. Suppose 入1;:::; 入m are distinct eigenvalues of T and

*v*1;:::; *v*m are corresponding eigenvectors. Then *v*1;:::; *v*m is linearly

independent.

Proof Suppose *v*1;:::; *v*m is linearly dependent. Let k be the smallest posi- tive integer such that

**5.11** *v*k 2 span.*v*1;:::; *v*k 1/I

the existence of k with this property follows from the Linear Dependence Lemma (2.21). Thus there exist a1;:::; ak 1 2 **F** such that

**5.12** *v*k D a1*v*1 C ... C ak 1*v*k 1:

Apply T to both sides of this equation, getting

入k*v*k D a1入1*v*1 C ... C ak 1入k 1*v*k 1:

Multiply both sides of 5.12 by 入k and then subtract the equation above, getting

0 D a1.入k — 入1/*v*1 C ... C ak 1.入k — 入k 1/*v*k 1:

Because we chose k to be the smallest positive integer satisfying 5.11, *v*1;:::; *v*k 1 is linearly independent. Thus the equation above implies that all the a’s are 0 (recall that 入k is not equal to any of 入1;:::; 入k 1). However, this means that *v*k equals 0 (see 5.12), contradicting our hypothesis that *v*k is an eigenvector. Therefore our assumption that *v*1;:::; *v*m is linearly dependent

was false.

The corollary below states that an operator cannot have more distinct eigenvalues than the dimension of the vector space on which it acts.

2 *L*

5.13 Number of eigenvalues

Suppose V is ﬁnite-dimensional. Then each operator on V has at most dim V distinct eigenvalues.

Proof Let T .V /. Suppose 入 1;:::; 入 m are distinct eigenvalues of T. Let *v*1;:::; *v*m be corresponding eigenvectors. Then 5.10 implies that the list *v*1;:::; *v*m is linearly independent. Thus m 三 dim V (see 2.23), as desired.

SECTION 5.A Invariant Subspaces **137**

### Restriction and Quotient Operators

If T .V / and U is a subspace of V invariant under T, then U determines two other operators T U .U / and T=U .V=U / in a natural way, as deﬁned below.

j 2 *L* 2 *L*

2 *L*

5.14 **Deﬁnition** T jU ***and*** T=U

Suppose T 2 *L*.V / and U is a subspace of V invariant under T.

* The ***restriction operator*** T jU 2 *L*.U / is deﬁned by

T jU .u/ D Tu

for u 2 U.

* The ***quotient operator*** T=U 2 *L*.V=U / is deﬁned by

.T=U /.*v* C U/ D T *v* C U

for *v* 2 V.

For both the operators deﬁned above, it is worthwhile to pay attention to their domains and to spend a moment thinking about why they are well deﬁned as operators on their domains. First consider the restriction operator

T U .U /, which is T with its domain restricted to U, thought of as mapping into U instead of into V. The condition that U is invariant under T is what allows us to think of T U as an operator on U, meaning a linear map

j

j 2 *L*

into the same space as the domain, rather than as simply a linear map from one vector space to another vector space.

To show that the deﬁnition above of the quotient operator makes sense, we need to verify that if *v* U *w* U, then T *v* U T *w* U. Hence suppose *v* U *w* U. Thus *v w* U (see 3.85). Because U is invariant under T, we also have T .*v w*/ U, which implies that T *v* T *w* U, which implies that T *v* U T *w* U, as desired.

C D C

— 2 — 2

C D C — 2

C D C C D C

Suppose T is an operator on a ﬁnite-dimensional vector space V and U is a subspace of V invariant under T, with U 0 and U V. In some sense, we can learn about T by studying the operators T U and T=U, each of which is an operator on a vector space with smaller dimension than V. For example, proof 2 of 5.27 makes nice use of T=U.

j

¤ f g ¤

However, sometimes T U and T=U do not provide enough information about T. In the next example, both T U and T=U are 0 even though T is not the 0 operator.

j

j

**138** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.15 **Example** Deﬁne an operator T 2 *L*.**F**2/ by T .x; y/ D .y; 0/. Let

U D f.x; 0/ W x 2 **F**g. Show that

1. U is invariant under T and T jU is the 0 operator on U ;
2. there does not exist a subspace W of **F**2 that is invariant under T and such that **F**2 D U ˚ W ;
3. T=U is the 0 operator on **F**2=U.

Solution

1. For .x; 0/ 2 U, we have T .x; 0/ D .0; 0/ 2 U. Thus U is invariant under T and T jU is the 0 operator on U.
2. Suppose W is a subspace of V such that **F**2 U W. Because dim **F**2 2 and dim U 1, we have dim W 1. If W were invariant under T, then each nonzero vector in W would be an eigenvector of T. However, it is easy to see that 0 is the only eigenvalue of T and that all eigenvectors of T are in U. Thus W is not invariant under T.

D D D

D ˚

1. For .x; y/ 2 **F**2, we have

（ ）

.T=U / .x; y/ C U D T .x; y/ C U

D .y; 0/ C U

D 0 C U;

where the last equality holds because .y; 0/ U. The equation above shows that T=U is the 0 operator.

2

EXERCISES 5.A

* 1. Suppose T 2 *L*.V / and U is a subspace of V.
     1. Prove that if U c null T, then U is invariant under T.
     2. Prove that if range T c U, then U is invariant under T.
  2. Suppose S; T .V / are such that ST TS. Prove that null S is invariant under T.

2 *L* D

SECTION 5.A Invariant Subspaces **139**

* 1. Suppose S; T .V / are such that ST TS. Prove that range S is invariant under T.

2 *L* D

* 1. Suppose that T 2 *L*.V / and U1;:::; Um are subspaces of V invariant under T. Prove that U1 C C Um is invariant under T.
  2. Suppose T .V /. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

2 *L*

* 1. Prove or give a counterexample: if V is ﬁnite-dimensional and U is a subspace of V that is invariant under every operator on V, then U D f0g or U D V.
  2. Suppose T .**R**2/ is deﬁned by T .x; y/ . 3y; x/. Find the eigenvalues of T.

2 *L* D —

* 1. Deﬁne T 2 *L*.**F**2/ by

T .*w*; z/ D .z; *w*/:

Find all eigenvalues and eigenvectors of T.

* 1. Deﬁne T 2 *L*.**F**3/ by

T .z1; z2; z3/ D .2z2; 0; 5z3/:

Find all eigenvalues and eigenvectors of T.

* 1. Deﬁne T 2 *L*.**F**n/ by

T .x1; x2; x3;:::; xn/ D .x1; 2x2; 3x3;:::; nxn/:

* + 1. Find all eigenvalues and eigenvectors of T.
    2. Find all invariant subspaces of T.
  1. Deﬁne T .**R**/ .**R**/ by Tp p0. Find all eigenvalues and eigenvectors of T.

W *P* ! *P* D

* 1. Deﬁne T 2 *L*（*P*4.**R**/） by

.Tp/.x/ D xp0.x/

for all x 2 **R**. Find all eigenvalues and eigenvectors of T.

* 1. Suppose V is ﬁnite-dimensional, T 2 *L*.V /, and 入 2 **F**. Prove that there

1000

exists ˛ 2 **F** such that j˛ — 入j < 1 and T — ˛I is invertible.

**140** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

1. Suppose V U W, where U and W are nonzero subspaces of V. Deﬁne P .V / by P.u *w*/ u for u U and *w* W. Find all eigenvalues and eigenvectors of P .

2 *L* C D 2 2

D ˚

1. Suppose T 2 *L*.V /. Suppose S 2 *L*.V / is invertible.
   1. Prove that T and S 1TS have the same eigenvalues.
   2. What is the relationship between the eigenvectors of T and the eigenvectors of S 1TS?
2. Suppose V is a complex vector space, T .V /, and the matrix of T with respect to some basis of V contains only real entries. Show that if 入is an eigenvalue of T, then so is 入N .

2 *L*

1. Give an example of an operator T .**R**4/ such that T has no (real) eigenvalues.

2 *L*

1. Show that the operator T 2 *L*.**C**1/ deﬁned by

T .z1; z2;:::/ D .0; z1; z2;:::/

has no eigenvalues.

1. Suppose n is a positive integer and T 2 *L*.**F**n/ is deﬁned by

T .x1;:::; xn/ D .x1 C ... C xn;:::; x1 C ... C xn/I

in other words, T is the operator whose matrix (with respect to the standard basis) consists of all 1’s. Find all eigenvalues and eigenvectors of T.

1. Find all eigenvalues and eigenvectors of the backward shift operator

T 2 *L*.**F**1/ deﬁned by

T .z1; z2; z3;:::/ D .z2; z3;::: /:

1. Suppose T 2 *L*.V / is invertible.
   1. Suppose 入 **F** with 入 0. Prove that 入is an eigenvalue of T if and only if 1 is an eigenvalue of T 1.

入

2 ¤

* 1. Prove that T and T 1 have the same eigenvectors.

SECTION 5.A Invariant Subspaces R:: 100p2

1. Suppose T .V / and there exist nonzero vectors *v* and *w* in V such that

2 *L*

T *v* D 3*w* and T *w* D 3*v*:

Prove that 3 or —3 is an eigenvalue of T.

1. Suppose V is ﬁnite-dimensional and S; T .V /. Prove that ST and

2 *L*

TS have the same eigenvalues.

1. Suppose A is an n-by-n matrix with entries in **F**. Deﬁne T .**F**n/ by Tx Ax, where elements of **F**n are thought of as n-by-1 column vectors.

D

2 *L*

* 1. Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.
  2. Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

1. Suppose T .V / and u; *v* are eigenvectors of T such that u *v* is also an eigenvector of T. Prove that u and *v* are eigenvectors of T corresponding to the same eigenvalue.

2 *L* C

1. Suppose T .V / is such that every nonzero vector in V is an eigen- vector of T. Prove that T is a scalar multiple of the identity operator.

2 *L*

1. Suppose V is ﬁnite-dimensional and T .V / is such that every sub- space of V with dimension dim V 1 is invariant under T. Prove that T is a scalar multiple of the identity operator.

—

2 *L*

1. Suppose V is ﬁnite-dimensional with dim V 3 and T .V / is such that every 2-dimensional subspace of V is invariant under T. Prove that T is a scalar multiple of the identity operator.

乏 2 *L*

1. Suppose T 2 *L*.V / and dim range T D k. Prove that T has at most

k C 1 distinct eigenvalues.

1. Suppose T 2 *L*.**R**3/ and —4, 5, and p7 are eigenvpalues of T. Prove that

there exists x 2 **R**3 such that Tx — 9x D .—4; 5; 7/.

1. Suppose V is ﬁnite-dimensional and *v*1;:::; *v*m is a list of vectors in V. Prove that *v*1;:::; *v*m is linearly independent if and only if there exists T .V / such that *v*1;:::; *v*m are eigenvectors of T corresponding to

2 *L*

distinct eigenvalues.

**142** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

1. Suppose 入1;:::; 入n is a list of distinct real numbers. Prove that the list e入*1*x;:::; e入*n*x is linearly independent in the vector space of real-valued functions on **R**.

（ ）

*Hint:* Let V D span e入*1*x;:::; e入*n*x , and deﬁne an operator T 2 *L*.V /

by Tf D f 0. Find eigenvalues and eigenvectors of T .

1. Suppose T 2 *L*.V /. Prove that T=.range T/ D 0.
2. Suppose T 2 *L*.V /. Prove that T=.null T/ is injective if and only if

.null T/ \ .range T/ D f0g.

1. Suppose V is ﬁnite-dimensional, T .V /, and U is invariant under T. Prove that each eigenvalue of T=U is an eigenvalue of T.

2 *L*

[*The exercise below asks you to verify that the hypothesis that* V *is*

*ﬁnite-dimensional is needed for the exercise above.*]

1. Give an example of a vector space V, an operator T .V /, and a subspace U of V that is invariant under T such that T=U has an eigenvalue that is not an eigenvalue of T.

2 *L*

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **143**

## *Eigenvectors and Upper-Triangular* Matrices

5.B

### Polynomials Applied to Operators

The main reason that a richer theory exists for operators (which map a vector space into itself) than for more general linear maps is that operators can be raised to powers. We begin this section by deﬁning that notion and the key concept of applying a polynomial to an operator.

If T .V /, then TT makes sense and is also in .V /. We usually write

2 *L L*

T 2 instead of T T. More generally, we have the following deﬁnition.

5.16 **Deﬁnition**

T m

Suppose T 2 *L*.V / and m is a positive integer.

* T m is deﬁned by

T D T

m

„ ƒ‚ …

m times

* . . T :
* T 0 is deﬁned to be the identity operator I on V.
* If T is invertible with inverse T 1, then T m is deﬁned by

T m D .T 1/ :

m

You should verify that if T is an operator, then

T mT n D T mCn and .T m/n D T mn;

where m and n are allowed to be arbitrary integers if T is invertible and nonnegative integers if T is not invertible.

5.17 **Deﬁnition** p.T /

Suppose T 2 *L*.V / and p 2 *P*.**F**/ is a polynomial given by

p.z/ D a0 C a1z C a2z C ... C amz

for z 2 **F**. Then p.T / is the operator deﬁned by

2

m

p.T / D a0I C a1T C a2T C ... C amT :

2

m

This is a new use of the symbol p because we are applying it to operators, not just elements of **F**.

**144** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.18 **Example**

deﬁned by Dq D q

Suppose D 2 *L P*.**R**/ is the differentiation operator and p is the polynomial deﬁned by p.x/ D 7 — 3x C5x2.

Then p.D/ D 7I — 3D C 5D2; thus

（ ）0

（ ）

p.D/ q D 7q — 3q0 C 5q00

for every q 2 *P*.**R**/.

If we ﬁx an operator T 2 *L*.V /, then the function from *P*.**F**/ to *L*.V /

given by p 7! p.T / is linear, as you should verify.

5.19 **Deﬁnition *product of polynomials***

If p; q 2 *P*.**F**/, then pq 2 *P*.**F**/ is the polynomial deﬁned by

.pq/.z/ D p.z/q.z/

for z 2 **F**.

Any two polynomials of an operator commute, as shown below.

5.20 Multiplicative properties

Then (a)

(b)

Suppose p; q 2 *P*.**F**/ and T 2 *L*.V /.

.pq/.T / D p.T /q.T /;

p.T /q.T/ D q.T /p.T /.

*Part (a) holds because when ex- panding a product of polynomials using the distributive property, it does not matter whether the sym-*

*bol is* z *or* T*.*

Proof

1. Suppose p.z/ D Pm

aj zj and q.z/ D Pn

bkzk for z 2 **F**.

Then

j D0

m n

X

X

.pq/.z/ D

kD0

aj bkzj Ck:

Thus

j D0 kD0

m n

X

X

.pq/.T / D

aj bk T j Ck

j D0 kD0 m

X

j

D aj T

j D0

n

kD0

X

bkT k

D p.T /q.T /:

1. Part (a) implies p.T /q.T/ D .pq/.T / D .qp/.T / D q.T /p.T /.

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **145**

### Existence of Eigenvalues

Now we come to one of the central results about operators on complex vector spaces.

5.21 Operators on complex vector spaces have an eigenvalue

Every operator on a ﬁnite-dimensional, nonzero, complex vector space has an eigenvalue.

Proof Suppose V is a complex vector space with dimension n > 0 and

T 2 *L*.V /. Choose *v* 2 V with *v* ¤ 0. Then

*v*;T *v*;T 2*v*;:::;T n*v*

is not linearly independent, because V has dimension n and we have n 1

C

vectors. Thus there exist complex numbers a0;:::; an, not all 0, such that

0 D a0*v* C a1T *v* C ... C anT n*v*:

Note that a1;:::; an cannot all be 0, because otherwise the equation above would become 0 a0*v*, which would force a0 also to be 0.

D

Make the a’s the coefﬁcients of a polynomial, which by the Fundamental

Theorem of Algebra (4.14) has a factorization

a0 C a1z C ... C anzn D c.z — 入1/ z — 入m/;

where c is a nonzero complex number, each 入j is in **C**, and the equation holds for all z **C** (here m is not necessarily equal to n, because an may equal 0). We then have

2

0 D a0*v* C a1T *v* C ... C anT n*v*

D .a0I C a1T C ... C anT /*v*

n

D c.T — 入1I/ ... .T — 入mI /*v*:

Thus T 入j I is not injective for at least one j . In other words, T has an eigenvalue.

—

The proof above depends on the Fundamental Theorem of Algebra, which is typical of proofs of this result. See Exercises 16 and 17 for possible ways to rewrite the proof above using the idea of the proof in a slightly different form.

**146** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

### Upper-Triangular Matrices

In Chapter 3 we discussed the matrix of a linear map from one vector space to another vector space. That matrix depended on a choice of a basis of each of the two vector spaces. Now that we are studying operators, which map a vector space to itself, the emphasis is on using only one basis.

Note that the matrices of operators are square arrays, rather than the more general rectangular arrays that we considered earlier for linear maps.

5.22 **Deﬁnition *matrix of an operator,*** *M*.T /

Suppose T 2 *L*.V / and *v*1;:::; *v*n is a basis of V. The ***matrix of*** T with

respect to this basis is the n-by-n matrix

*M*.T / D

B

A

1;1

::: A

1;n

@

:

An;1

:::

:

An;n

C

whose entries Aj;k are deﬁned by

A

T *v*k D A1;k*v*1 C ... C An;k*v*n:

If the basis is not clear from the context, then the notation

*M* T; .

（

*v* ;:::;

1

*v* / is used.

n

）

If T is an operator on **F**n and no

*The* k*th column of the matrix*

*M*.T / *is formed from the coefﬁ-*

*cients used to write* T *v*k *as a linear combination of v*1;:::; *v*n*.*

basis is speciﬁed, assume that the basis

in question is the standard one (where the j th basis vector is 1 in the j th slot and 0 in all the other slots). You can

then think of the j th column of *M*.T / as T applied to the j th basis vector.

Then

0@ 1A

5.23 **Example** Deﬁne T 2 *L*.**F**3/ by T .x; y; z/ D .2xCy; 5y C3z; 8z/.

*M*.T / D

2 1 0

0 5 3 :

0 0 8

A central goal of linear algebra is to show that given an operator T .V /, there exists a basis of V with respect to which T has a reasonably simple matrix. To make this vague formulation a bit more precise, we might try to

2 *L*

choose a basis of V such that *M*.T / has many 0’s.

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **147**

If V is a ﬁnite-dimensional complex vector space, then we already know enough to show that there is a basis of V with respect to which the matrix of T has 0’s everywhere in the ﬁrst column, except possibly the ﬁrst entry. In other words, there is a basis of V with respect to which the matrix of T looks

like

0B 入 1C

0

@B

\*

: CA I

0

\*

here the denotes the entries in all the columns other than the ﬁrst column. To prove this, let 入be an eigenvalue of T (one exists by 5.21) and let *v* be a corresponding eigenvector. Extend *v* to a basis of V. Then the matrix of T

with respect to this basis has the form above.

Soon we will see that we can choose a basis of V with respect to which the matrix of T has even more 0’s.

5.24 **Deﬁnition *diagonal of a matrix***

The ***diagonal*** of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner.

For example, the diagonal of the matrix in 5.23 consists of the entries

2; 5; 8.

5.25 **Deﬁnition *upper-triangular matrix***

A matrix is called ***upper triangular*** if all the entries below the diagonal equal 0.

For example, the matrix in 5.23 is upper triangular.

Typically we represent an upper-triangular matrix in the form

\*

入

C

A

I

n

*We often use* \* *to denote matrix en-*

*tries that we do not know about or that are irrelevant to the questions being discussed.*

0B 入1

@

: : :

0

the 0 in the matrix above indicates that all entries below the diagonal in this n-by-n matrix equal 0. Upper- triangular matrices can be considered reasonably simple—for n large, almost

half its entries in an n-by-n upper-

triangular matrix are 0.

**148** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

The following proposition demonstrates a useful connection between upper-triangular matrices and invariant subspaces.

5.26 Conditions for upper-triangular matrix

Suppose T 2 *L*.V / and *v*1;:::; *v*n is a basis of V. Then the following are

equivalent:

1. the matrix of T with respect to *v*1;:::; *v*n is upper triangular;
2. T *v*j 2 span.*v*1;:::; *v*j / for each j D 1; :::; n;
3. span.*v*1;:::; *v*j / is invariant under T for each j D 1; :::; n.

Proof The equivalence of (a) and (b) follows easily from the deﬁnitions and a moment’s thought. Obviously (c) implies (b). Hence to complete the proof, we need only prove that (b) implies (c).

Thus suppose (b) holds. Fix j 2 f1; :::; ng. From (b), we know that

T *v*1 2 span.*v*1/ c span.*v*1;:::; *v*j /I

T *v*2 2 span.*v*1; *v*2/ c span.*v*1;:::; *v*j /I

:

**:**

T *v*j 2 span.*v*1;:::; *v*j /:

Thus if *v* is a linear combination of *v*1;:::; *v*j , then

T *v* 2 span.*v*1;:::; *v*j /:

In other words, span.*v*1;:::; *v*j / is invariant under T, completing the proof.

Now we can prove that for each operator on a ﬁnite-dimensional com- plex vector space, there is a basis of the vector space with respect to which the

*The next result does not hold on real vector spaces, because the ﬁrst vector in a basis with respect to which an operator has an upper- triangular matrix is an eigenvector of the operator. Thus if an opera- tor on a real vector space has no eigenvalues* [*see 5.8*(*a*) *for an ex- ample*]*, then there is no basis with respect to which the operator has an upper-triangular matrix.*

matrix of the operator has only 0’s be-

low the diagonal. In Chapter 8 we will

improve even this result.

Sometimes more insight comes from seeing more than one proof of a theo- rem. Thus two proofs are presented of the next result. Use whichever appeals more to you.

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **149**

5.27 Over **C**, every operator has an upper-triangular matrix

Then T has an upper-triangular matrix with respect to some basis of V.

Suppose V is a ﬁnite-dimensional complex vector space and T 2 *L*.V /.

Proof 1 We will use induction on the dimension of V. Clearly the desired result holds if dim V 1.

Suppose now that dim V > 1 and the desired result holds for all complex vector spaces whose dimension is less than the dimension of V. Let 入be any eigenvalue of T (5.21 guarantees that T has an eigenvalue). Let

D

U D range.T — 入I /:

Because T — 入I is not surjective (see 3.69), dim U < dim V. Furthermore,

U is invariant under T. To prove this, suppose u 2 U. Then

Tu D .T — 入I /u C 入u:

Obviously .T 入 I /u U (because U equals the range of T 入 I ) and 入 u U. Thus the equation above shows that Tu U. Hence U is invariant under T, as claimed.

2 2

— 2 —

Thus T U is an operator on U. By our induction hypothesis, there is a basis u1;:::; um of U with respect to which T U has an upper-triangular matrix. Thus for each j we have (using 5.26)

j

j

**5.28** T uj D .T jU /.uj / 2 span.u1;:::; uj /:

Extend u1;:::; um to a basis u1;:::; um; *v*1;:::; *v*n of V. For each k, we have

T *v*k D .T — 入I /*v*k C 入*v*k:

The deﬁnition of U shows that .T 入I /*v*k U span.u1;:::; um/. Thus the equation above shows that

— 2 D

**5.29** T *v*k 2 span.u1;:::; um; *v*1;:::; *v*k/:

From 5.28 and 5.29, we conclude (using 5.26) that T has an upper- triangular matrix with respect to the basis u1;:::; um; *v*1;:::; *v*n of V, as desired.

**150** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

Proof 2 We will use induction on the dimension of V. Clearly the desired result holds if dim V 1.

D

Suppose now that dim V n > 1 and the desired result holds for all complex vector spaces whose dimension is n 1. Let *v*1 be any eigenvector of T (5.21 guarantees that T has an eigenvector). Let U span.*v*1/. Then U is an invariant subspace of T and dim U 1.

D

D

—

D

Because dim V=U n 1 (see 3.89), we can apply our induction hy- pothesis to T=U .V=U /. Thus there is a basis *v*2 U; :::; *v*n U of V=U such that T=U has an upper-triangular matrix with respect to this basis.

2 *L* C C

D —

Hence by 5.26,

.T=U /.*v*j C U/ 2 span.*v*2 C U; :::; *v*j C U/

for each j 2; :::; n. Unraveling the meaning of the inclusion above, we see that

D

T *v*j 2 span.*v*1;:::; *v*j /

for each j 1; :::; n. Thus by 5.26, T has an upper-triangular matrix with respect to the basis *v*1;:::; *v*n of V, as desired (it is easy to verify that *v*1;:::; *v*n is a basis of V ; see Exercise 13 in Section 3.E for a more general

D

result).

How does one determine from looking at the matrix of an operator whether the operator is invertible? If we are fortunate enough to have a basis with respect to which the matrix of the operator is upper triangular, then this problem becomes easy, as the following proposition shows.

5.30 Determination of invertibility from upper-triangular matrix

Suppose T 2 *L*.V / has an upper-triangular matrix with respect to some

basis of V. Then T is invertible if and only if all the entries on the diagonal

of that upper-triangular matrix are nonzero.

Proof Suppose *v*1;:::; *v*n is a basis of V with respect to which T has an upper-triangular matrix

0B 入1 \* 1C

**5.31** *M*.T / D B@

入2

: : :

CA :

0 入n

We need to prove that T is invertible if and only if all the 入j ’s are nonzero.

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **151**

First suppose the diagonal entries 入1;:::; 入n are all nonzero. The upper- triangular matrix in 5.31 implies that T *v*1 入1*v*1. Because 入1 0, we have T .*v*1=入1/ *v*1; thus *v*1 range T.

D 2

D ¤

Now

T .*v*2=入2/ D a*v*1 C *v*2

for some a **F**. The left side of the equation above and a*v*1 are both in range T ; thus *v*2 range T.

2

2

Similarly, we see that

T .*v*3=入3/ D b*v*1 C c*v*2 C *v*3

for some b; c **F**. The left side of the equation above and b*v*1; c*v*2 are all in range T ; thus *v*3 range T.

2

2

Continuing in this fashion, we conclude that *v*1;:::; *v*n range T. Be- cause *v*1;:::; *v*n is a basis of V, this implies that range T V. In other words, T is surjective. Hence T is invertible (by 3.69), as desired.

D

2

To prove the other direction, now suppose that T is invertible. This implies that 入1 ¤ 0, because otherwise we would have T *v*1 D 0.

Let 1 < j 三 n, and suppose 入j D 0. Then 5.31 implies that T maps span.*v*1;:::; *v*j / into span.*v*1;:::; *v*j 1/. Because

dim span.*v*1;:::; *v*j / D j and dim span.*v*1;:::; *v*j 1/ D j — 1;

this implies that T restricted to dim span.*v*1;:::; *v*j / is not injective (by 3.23). Thus there exists *v* span.*v*1;:::; *v*j / such that *v* 0 and T *v* 0. Thus T is not injective, which contradicts our hypothesis (for this direction) that T is invertible. This contradiction means that our assumption that 入j D 0 must be false. Hence 入j ¤ 0, as desired.

2 ¤ D

As an example of the result above, we see that the operator in Example 5.23 is invertible.

Unfortunately no method exists for exactly computing the eigenvalues of an operator from its matrix. However, if we are fortunate enough to ﬁnd a ba- sis with respect to which the matrix of the operator is upper triangular, then the problem of computing the eigenvalues becomes trivial, as the following propo- sition shows.

*Powerful numeric techniques exist for ﬁnding good approximations to the eigenvalues of an operator from its matrix.*

**152** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.32 Determination of eigenvalues from upper-triangular matrix

Suppose T 2 *L*.V / has an upper-triangular matrix with respect to some

basis of V. Then the eigenvalues of T are precisely the entries on the

diagonal of that upper-triangular matrix.

Proof Suppose *v*1;:::; *v*n is a basis of V with respect to which T has an upper-triangular matrix

0B 入1 \* 1C

*M*.T / D B@

入2

: : :

CA :

Let 入 2 **F**. Then

0 入n

0B 入1 — 入 \* 1C

*M*.T — 入I / D B@

入2 — 入

: : :

CA :

0 入n — 入

Hence T 入 I is not invertible if and only if 入 equals one of the numbers 入1;:::; 入n (by 5.30). Thus 入is an eigenvalue of T if and only if 入equals one of the numbers 入1;:::; 入n.

—

5.33 **Example** Deﬁne T .**F**3/ by T .x; y; z/ .2x y; 5y 3z; 8z/. What are the eigenvalues of T ?

Solution The matrix of T with respect to the standard basis is

2 *L* D C C

0@ 2 1 0 1A

*M*.T / D

0 5 3

0 0 8

:

Thus .T / is an upper-triangular matrix. Now 5.32 implies that the eigen- values of T are 2, 5, and 8.

*M*

Once the eigenvalues of an operator on **F**n are known, the eigenvectors can be found easily using Gaussian elimination.

SECTION 5.B Eigenvectors and Upper-Triangular Matrices **153**

EXERCISES 5.B

1. Suppose T 2 *L*.V / and there exists a positive integer n such that T n D 0.
   1. Prove that I — T is invertible and that

.I — T / 1 D I C T C C T n 1:

* 1. Explain how you would guess the formula above.

1. Suppose T 2 *L*.V / and .T — 2I/.T — 3I/.T — 4I/ D 0. Suppose 入is an eigenvalue of T. Prove that 入 D 2 or 入 D 3 or 入 D 4.
2. Suppose T 2 *L*.V / and T 2 D I and —1 is not an eigenvalue of T. Prove that T D I.
3. Suppose P 2 *L*.V / and P 2 D P . Prove that V D null P ˚ range P .
4. Suppose S; T .V / and S is invertible. Suppose p .**F**/ is a polynomial. Prove that

2 *L* 2 *P*

p.STS 1/ D Sp.T /S 1:

1. Suppose T 2 *L*.V / and U is a subspace of V invariant under T. Prove that U is invariant under p.T / for every polynomial p 2 *P*.**F**/.
2. Suppose T 2 *L*.V /. Prove that 9 is an eigenvalue of T 2 if and only if 3

or —3 is an eigenvalue of T.

1. Give an example of T 2 *L*.**R**2/ such that T 4 D —1.
2. Suppose V is ﬁnite-dimensional, T .V /, and *v* V with *v* 0.

2 *L* 2 ¤

Let p be a nonzero polynomial of smallest degree such that p.T /*v* 0.

D

Prove that every zero of p is an eigenvalue of T.

1. Suppose T 2 *L*.V / and *v* is an eigenvector of T with eigenvalue 入 . Suppose p 2 *P*.**F**/. Prove that p.T /*v* D p.入/*v*.
2. Suppose **F C**, T .V /, p .**C**/ is a polynomial, and ˛ **C**. Prove that ˛ is an eigenvalue of p.T / if and only if ˛ p.入/ for some eigenvalue 入of T.

D

D 2 *L* 2 *P* 2

1. Show that the result in the previous exercise does not hold if **C** is replaced with **R**.

**154** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

1. Suppose W is a complex vector space and T .W / has no eigenvalues. Prove that every subspace of W invariant under T is either 0 or inﬁnite- dimensional.

f g

2 *L*

1. Give an example of an operator whose matrix with respect to some basis contains only 0’s on the diagonal, but the operator is invertible.

[*The exercise above and the exercise below show that 5.30 fails without*

*the hypothesis that an upper-triangular matrix is under consideration.*]

1. Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.
2. Rewrite the proof of 5.21 using the linear map that sends p 2 *P*n.**C**/ to

（ ）

p.T / *v* 2 V (and use 3.23).

1. Rewrite the proof of 5.21 using the linear map that sends p 2 *P*n*2* .**C**/ to

p.T / 2 *L*.V / (and use 3.23).

1. Suppose V is a ﬁnite-dimensional complex vector space and T 2 *L*.V /.

Deﬁne a function f W **C** ! **R** by

f .入/ D dim range.T — 入I /:

Prove that f is not a continuous function.

1. Suppose V is ﬁnite-dimensional with dim V > 1 and T .V /. Prove that

2 *L*

fp.T / W p 2 *P*.**F**/g¤ *L*.V /:

1. Suppose V is a ﬁnite-dimensional complex vector space and T .V /. Prove that T has an invariant subspace of dimension k for each k 1; :::; dim V.

D

2 *L*

SECTION 5.C Eigenspaces and Diagonal Matrices **155**

## *Eigenspaces and Diagonal Matrices*

5.34 **Deﬁnition *diagonal matrix***

A ***diagonal matrix*** is a square matrix that is 0 everywhere except possibly along the diagonal.

5.C

5.35 **Example**

is a diagonal matrix.

8 0 0

0 5 0

0@ 1A

0 0 5

Obviously every diagonal matrix is upper triangular. In general, a diagonal matrix has many more 0’s than an upper-triangular matrix.

If an operator has a diagonal matrix with respect to some basis, then the

entries along the diagonal are precisely the eigenvalues of the operator; this follows from 5.32 (or ﬁnd an easier proof for diagonal matrices).

5.36 **Deﬁnition *eigenspace,*** E.入; T /

Suppose T 2 *L*.V / and 入 2 **F**. The ***eigenspace*** of T corresponding to 入,

denoted E.入; T /, is deﬁned by

E.入; T / D null.T — 入I /:

In other words, E.入; T / is the set of all eigenvectors of T corresponding to 入, along with the 0 vector.

For T .V / and 入 **F**, the eigenspace E. 入 ; T / is a subspace of V (because the null space of each linear map on V is a subspace of V ). The deﬁnitions imply that 入is an eigenvalue of T if and only if E.入; T / ¤ f0g.

2 *L*

2 *L* 2

5.37 **Example** Suppose the matrix of an operator T .V / with respect to a basis *v*1; *v*2; *v*3 of V is the matrix in Example 5.35 above. Then

E.8; T / D span.*v*1/; E.5; T / D span.*v*2; *v*3/:

If 入 is an eigenvalue of an operator T .V /, then T restricted to

E.入; T / is just the operator of multiplication by 入.

2 *L*

**156** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.38 Sum of eigenspaces is a direct sum

Suppose V is ﬁnite-dimensional and T 2 *L*.V /. Suppose also that

入1;:::; 入m are distinct eigenvalues of T. Then

E.入1;T / C ... C E.入m;T /

is a direct sum. Furthermore,

dim E.入1;T / C ... C dim E.入m;T / 三 dim V:

Proof To show that E.入1;T / C ... C E.入m;T / is a direct sum, suppose

u1 C ... C um D 0;

where each uj is in E.入; T /. Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.10), this implies that each uj equals 0. This implies (using 1.44) that E.入1;T / E.入m;T / is a direct

C. . .C

sum, as desired.

Now

（ ）

dim E.入1;T / C ... C dim E.入m;T / D dim E.入1;T / ˚ ... ˚ E.入m;T /

三 dim V;

where the equality above follows from Exercise 16 in Section 2.C.

5.39 **Deﬁnition *diagonalizable***

An operator T 2 *L*.V / is called ***diagonalizable*** if the operator has a

diagonal matrix with respect to some basis of V.

* 1. **Example** Deﬁne T 2 *L*.**R**2/ by

T .x; y/ D .41x C 7y; —20x C 74y/:

The matrix of T with respect to the standard basis of **R**2 is

41 7

—20 74 ;

which is not a diagonal matrix. However, T is diagonalizable, because the

matrix of T with respect to the basis .1; 4/; .7; 5/ is

69 0

0 46 ;

as you should verify.

SECTION 5.C Eigenspaces and Diagonal Matrices **157**

* 1. Conditions equivalent to diagonalizability

Suppose V is ﬁnite-dimensional and T .V /. Let 入1;:::; 入m denote the distinct eigenvalues of T. Then the following are equivalent:

2 *L*

1. T is diagonalizable;
2. V has a basis consisting of eigenvectors of T ;
3. there exist 1-dimensional subspaces U1;:::; Un of V, each invariant under T, such that

V D U1 ˚ ... ˚ UnI

(d) V D E.入1;T / ˚ ... ˚ E.入m;T /;

(e) dim V D dim E.入1;T / C ... C dim E.入m;T /.

Proof An operator T 2 *L*.V / has a diagonal matrix

0B

1

入1 0

: : C

@

A

:

0 入n

with respect to a basis *v*1;:::; *v*n of V if and only if T *v*j 入j *v*j for each j . Thus (a) and (b) are equivalent.

D

Suppose (b) holds; thus V has a basis *v*1;:::; *v*n consisting of eigenvectors of T. For each j , let Uj span.*v*j /. Obviously each Uj is a 1-dimensional subspace of V that is invariant under T. Because *v*1;:::; *v*n is a basis of V, each vector in V can be written uniquely as a linear combination of *v*1;:::; *v*n. In other words, each vector in V can be written uniquely as a sum u1 un, where each uj is in Uj . Thus V U1 Un. Hence (b) implies (c).

D

D ˚ ... ˚

C. . .C

Suppose now that (c) holds; thus there are 1-dimensional subspaces U1;:::; Un of V, each invariant under T, such that V U1 Un. For each j, let *v*j be a nonzero vector in Uj . Then each *v*j is an eigenvector of T. Because each vector in V can be written uniquely as a sum u1 un, where each uj is in Uj (so each uj is a scalar multiple of *v*j ), we see that *v*1;:::; *v*n is a basis of V. Thus (c) implies (b).

C. . .C

D ˚ . . .˚

At this stage of the proof we know that (a), (b), and (c) are all equivalent. We will ﬁnish the proof by showing that (b) implies (d), that (d) implies (e), and that (e) implies (b).

**158** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

Suppose (b) holds; thus V has a basis consisting of eigenvectors of T. Hence every vector in V is a linear combination of eigenvectors of T, which implies that

V D E.入1;T / C ... C E.入m;T /:

Now 5.38 shows that (d) holds.

That (d) implies (e) follows immediately from Exercise 16 in Section 2.C. Finally, suppose (e) holds; thus

* 1. dim V D dim E.入1;T / C ... C dim E.入m;T /:

Choose a basis of each E.入j ;T /; put all these bases together to form a list *v*1;:::; *v*n of eigenvectors of T, where n dim V (by 5.42). To show that this list is linearly independent, suppose

D

a1*v*1 C ... C an*v*n D 0;

where a1;:::; an 2 **F**. For each j D 1; :::; m, let uj denote the sum of all the terms ak*v*k such that *v*k 2 E.入j ;T /. Thus each uj is in E.入j ;T /, and

u1 C ... C um D 0:

Because eigenvectors corresponding to distinct eigenvalues are linearly inde- pendent (see 5.10), this implies that each uj equals 0. Because each uj is a sum of terms ak*v*k, where the *v*k’s were chosen to be a basis of E.入j ;T /, this implies that all the ak’s equal 0. Thus *v*1;:::; *v*n is linearly independent and hence is a basis of V (by 2.39). Thus (e) implies (b), completing the proof.

Unfortunately not every operator is diagonalizable. This sad state of affairs can arise even on complex vector spaces, as shown by the next example.

* 1. **Example** Show that the operator T 2 *L*.**C**2/ deﬁned by

T .*w*; z/ D .z; 0/

is not diagonalizable.

Solution As you should verify, 0 is the only eigenvalue of T and furthermore

E.0; T / .*w*; 0/ **C**2 *w* **C** .

D f 2 W 2 g

Thus conditions (b), (c), (d), and (e) of 5.41 are easily seen to fail (of

course, because these conditions are equivalent, it is only necessary to check that one of them fails). Thus condition (a) of 5.41 also fails, and hence T is not diagonalizable.

SECTION 5.C Eigenspaces and Diagonal Matrices **159**

The next result shows that if an operator has as many distinct eigenvalues as the dimension of its domain, then the operator is diagonalizable.

5.44 Enough eigenvalues implies diagonalizability

If T 2 *L*.V / has dim V distinct eigenvalues, then T is diagonalizable.

Proof Suppose T .V / has dim V distinct eigenvalues 入1;:::; 入dim V . For each j , let *v*j V be an eigenvector corresponding to the eigenvalue 入j . Because eigenvectors corresponding to distinct eigenvalues are linearly inde-

pendent (see 5.10), *v*1;:::; *v*dim V is linearly independent. A linearly indepen- dent list of dim V vectors in V is a basis of V (see 2.39); thus *v*1;:::; *v*dim V is a basis of V. With respect to this basis consisting of eigenvectors, T has a

2

2 *L*

diagonal matrix.

5.45 **Example** Deﬁne T .**F**3/ by T .x; y; z/ .2x y; 5y 3z; 8z/. Find a basis of **F**3 with respect to which T has a diagonal matrix.

Solution With respect to the standard basis, the matrix of T is

2 *L* D C C

0@ 2 1 0 1A

0 5 3

0 0 8

:

The matrix above is upper triangular. Thus by 5.32, the eigenvalues of T are 2, 5, and 8. Because T is an operator on a vector space with dimension 3 and T has three distinct eigenvalues, 5.44 assures us that there exists a basis of **F**3 with respect to which T has a diagonal matrix.

To ﬁnd this basis, we only have to ﬁnd an eigenvector for each eigenvalue.

In other words, we have to ﬁnd a nonzero solution to the equation

T .x; y; z/ D 入.x; y; z/

for 入 2, then for 入 5, and then for 入 8. These simple equations are easy to solve: for 入 2 we have the eigenvector .1; 0; 0/; for 入 5 we have the eigenvector .1; 3; 0/; for 入 8 we have the eigenvector .1; 6; 6/.

D

D D

D D D

Thus .1; 0; 0/; .1; 3; 0/; .1; 6; 6/ is a basis of **F**3, and with respect to this basis the matrix of T is

0@ 1A

2 0 0

0 5 0 :

0 0 8

**160** CHAPTER 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

The converse of 5.44 is not true. For example, the operator T deﬁned on the three-dimensional space **F**3 by

T .z1; z2; z3/ D .4z1; 4z2; 5z3/

has only two eigenvalues (4 and 5), but this operator has a diagonal matrix with respect to the standard basis.

In later chapters we will ﬁnd additional conditions that imply that certain operators are diagonalizable.

EXERCISES 5.C

1. Suppose T 2 *L*.V / is diagonalizable. Prove that V D null T ˚ range T.
2. Prove the converse of the statement in the exercise above or give a counterexample to the converse.
3. Suppose V is ﬁnite-dimensional and T .V /. Prove that the following are equivalent:

2 *L*

* 1. V D null T ˚ range T.
  2. V D null T C range T.
  3. null T \ range T D f0g.

1. Give an example to show that the exercise above is false without the hypothesis that V is ﬁnite-dimensional.
2. Suppose V is a ﬁnite-dimensional complex vector space and T .V /. Prove that T is diagonalizable if and only if

2 *L*

V D null.T — 入I / ˚ range.T — 入I /

for every 入 2 **C**.

1. Suppose V is ﬁnite-dimensional, T 2 *L*.V / has dim V distinct eigenval- ues, and S 2 *L*.V / has the same eigenvectors as T (not necessarily with the same eigenvalues). Prove that ST D TS.
2. Suppose T .V / has a diagonal matrix A with respect to some basis of V and that 入 **F**. Prove that 入appears on the diagonal of A precisely dim E.入; T / times.

2

2 *L*

1. Suppose T .**F**5/ and dim E.8; T / 4. Prove that T 2I or T 6I

2 *L* D — —

is invertible.

SECTION 5.C Eigenspaces and Diagonal Matrices **161**

1. Suppose T 2 *L*.V / is invertible. Prove that E.入; T / D E. 1 ;T 1/ for

入

every 入 2 **F** with 入 ¤ 0.

1. Suppose that V is ﬁnite-dimensional and T .V /. Let 入1;:::; 入m

2 *L*

denote the distinct nonzero eigenvalues of T. Prove that

dim E.入1;T / C ... C dim E.入m;T / 三 dim range T:

1. Verify the assertion in Example 5.40.
2. Suppose R; T 2 *L*.**F**3/ each have 2, 6, 7 as eigenvalues. Prove that there exists an invertible operator S 2 *L*.**F**3/ such that R D S 1TS.
3. Find R; T .**F**4/ such that R and T each have 2, 6, 7 as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator S 2 *L*.**F**4/ such that R D S 1TS.

2 *L*

1. Find T .**C**3/ such that 6 and 7 are eigenvalues of T and such that T

2 *L*

does not have a diagonal matrix with respect to any basis of **C**3.

1. Suppose T .**C**3/ is such that 6 and 7 are eigenvalues of T. Fur- thermore, suppose T does not have a diagonal matrix with respect

2 *L*

to any basis of **C**3. Propve that there exists .x; y; z/ 2 **F**3 such that

T .x; y; z/ D .17 C 8x; 5 C 8y; 2兀 C 8z/.

1. The ***Fibonacci sequence*** F1; F2;::: is deﬁned by

F1 D 1; F2 D 1; and Fn D Fn 2 C Fn 1 for n 乏 3:

Deﬁne T 2 *L*.**R**2/ by T .x; y/ D .y; x C y/.

* 1. Show that T n.0; 1/ D .Fn; FnC1/ for each positive integer n.
  2. Find the eigenvalues of T.
  3. Find a basis of **R**2 consisting of eigenvectors of T.
  4. Use the solution to part (c) to compute T n.0; 1/. Conclude that

—

1

Fn D p5

1 C p5 n

2

1 — p5 n\_

for each positive integer n.

2

* 1. Use part (d) to conclude that for each positive integer n, the Fibonacci number Fn is the integer that is closest to

1 1 C p5 n

2

p5

:

# *Inner Product Spaces*



*Woman teaching geometry, from a fourteenth-century edition of Euclid’s geometry book.*

6

CHAPTER

In making the deﬁnition of a vector space, we generalized the linear structure (addition and scalar multiplication) of **R**2 and **R**3. We ignored other important features, such as the notions of length and angle. These ideas are embedded in the concept we now investigate, inner products.

Our standing assumptions are as follows:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a vector space over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

Cauchy–Schwarz Inequality Gram–Schmidt Procedure

linear functionals on inner product spaces calculating minimum distance to a subspace

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

163

DOI 10.1007/978-3-319-11080-6 6

**164** CHAPTER 6 Inner Product Spaces

## *Inner Products and Norms*

6.A

### Inner Products

To motivate the concept of inner prod- uct, think of vectors in **R**2 and **R**3 as

*(x*1 *, x*2*)*

*x*

arrows with initial point at the origin. The length of a vector x in **R**2 or **R**3 is called the ***norm*** of x, denoted kxk.

Thus fopr x D .x1; x2/ 2 **R**2, we have

*The length of this vector* x *is*

p

x 2 C x 2*.*

kxkD

x12

C x22.

1 2 Similarlyp, if x D .x1; x2; x3/ 2 **R**3,

then kxkD

x12 C x22 C x32.

Even though we cannot draw pictures in higher dimensions, the gener- alization to **R**n is obvious: we deﬁne the norm of x .x1;:::; xn/ **R**n by

D 2

p

kxkD x12 C ... C xn2:

The norm is not linear on **R**n. To inject linearity into the discussion, we introduce the dot product.

6.2 **Deﬁnition *dot product***

For x; y 2 **R**n, the ***dot product*** of x and y, denoted x . y, is deﬁned by

x . y D x1y1 C ... C xnyn;

where x D .x1;:::; xn/ and y D .y1;:::; yn/.

*If we think of vectors as points in-*

*be interpreted as the distance from the origin to the point* x*.*

*stead of arrows, then* kxk *should*

* x . x 乏 0 for all x 2 **R**n;
* x . x D 0 if and only if x D 0;

Note that the dot product of two vec- tors in **R**n is a number, not a vector. Ob- viously x x x 2 for all x **R**n.

The dot product on **R**n has the follow-

* D k k 2

ing properties:

for y **R**n ﬁxed, the map from **R**n to **R** that sends x **R**n to x y is linear;

* 2 2 .
* x . y D y . x for all x; y 2 **R**n.

SECTION 6.A Inner Products and Norms **165**

An inner product is a generalization of the dot product. At this point you may be tempted to guess that an inner product is deﬁned by abstracting the properties of the dot product discussed in the last paragraph. For real vector spaces, that guess is correct. However, so that we can make a deﬁnition that will be useful for both real and complex vector spaces, we need to examine the complex case before making the deﬁnition.

Recall that if 入 D a C bi, where a; b 2 **R**, then

* the absolute value of 入, denoted j入j, is deﬁned by j入jD pa2 C b2;
* the complex conjugate of 入, denoted 入N , is deﬁned by 入N D a — bi;
* j入j2 D 入入N .

See Chapter 4 for the deﬁnitions and the basic properties of the absolute value and complex conjugate.

For z D .z1;:::; zn/ 2 **C**n, we deﬁne the norm of z by

kzkD qjz1j2 C. C jznj2:

The absolute values are needed because we want z to be a nonnegative number. Note that

k k

kzk D z1z1 C ... C znzn:

2

We want to think of z 2 as the inner product of z with itself, as we did in **R**n. The equation above thus suggests that the inner product of *w* D .*w*1;:::; *w*n/ 2 **C**n with z should equal

k k

*w*1z1 C ... C *w*nzn:

If the roles of the *w* and z were interchanged, the expression above would be replaced with its complex conjugate. In other words, we should expect that the inner product of *w* with z equals the complex conjugate of the inner

product of z with *w*. With that motivation, we are now ready to deﬁne an

inner product on V, which may be a real or a complex vector space.

Two comments about the notation used in the next deﬁnition:

If 入is a complex number, then the notation 入 0 means that 入is real and nonnegative.

* 乏

We use the common notation u; *v* , with angle brackets denoting an inner product. Some people use parentheses instead, but then .u; *v*/

* h i

becomes ambiguous because it could denote either an ordered pair or an inner product.

**166** CHAPTER 6 Inner Product Spaces

* 1. **Deﬁnition *inner product***

An ***inner product*** on V is a function that takes each ordered pair .u; *v*/ of elements of V to a number hu; *v*i2 **F** and has the following properties:

#### positivity

h*v*; *v*i乏 0 for all *v* 2 V ;

#### deﬁniteness

h*v*; *v*iD 0 if and only if *v* D 0;

#### additivity in ﬁrst slot

hu C *v*; *w*iD hu; *w*iC h*v*; *w*i for all u; *v*; *w* 2 V ;

#### homogeneity in ﬁrst slot

h入u; *v*iD 入hu; *v*i for all 入 2 **F** and all u; *v* 2 V ;

#### conjugate symmetry

hu; *v*iD h*v*; ui for all u; *v* 2 V.

* 1. **Example *inner products***

Every real number equals its com- plex conjugate. Thus if we are dealing with a real vector space, then in the last condition above we can dispense with

the complex conjugate and simply state that hu; *v*iD h*v*; ui for all *v*; *w* 2 V.

*Although most mathematicians de- ﬁne an inner product as above, many physicists use a deﬁnition that requires homogeneity in the second slot instead of the ﬁrst slot.*

1. The ***Euclidean inner product*** on **F**nis deﬁned by

h.*w*1;:::; *w*n/; .z1;:::; zn/iD *w*1z1 C ... C *w*nzn:

1. If c1;:::; cn are positive numbers, then an inner product can be deﬁned on **F**n by

h.*w*1;:::; *w*n/; .z1;:::; zn/iD c1*w*1z1 C ... C cn*w*nzn:

1. An inner product can be deﬁned on the vector space of continuous real-valued functions on the inZterval Œ—1; 1] by

1

hf; giD

f .x/g.x/ dx:

1

1. An inner product can be deﬁZned on *P*.**R**/ by

hp; qiD

0

1 p.x/q.x/e x dx:

SECTION 6.A Inner Products and Norms **167**

6.5 **Deﬁnition *inner product space***

An ***inner product space*** is a vector space V along with an inner product on V.

The most important example of an inner product space is **F**n with the Euclidean inner product given by part (a) of the last example. When **F**n is referred to as an inner product space, you should assume that the inner product is the Euclidean inner product unless explicitly told otherwise.

So that we do not have to keep repeating the hypothesis that V is an inner

product space, for the rest of this chapter we make the following assumption:

6.6 **Notation** V

For the rest of this chapter, V denotes an inner product space over **F**.

Note the slight abuse of language here. An inner product space is a vector space along with an inner product on that vector space. When we say that

a vector space V is an inner product space, we are also thinking that an inner product on V is lurking nearby or is obvious from the context (or is the Euclidean inner product if the vector space is **F**n).

6.7 Basic properties of an inner product

(a) For each ﬁxed u 2 V, the function that takes *v* to h*v*; ui is a linear

(b)

(c)

(d)

(e)

map from V to **F**.

h0; uiD 0 for every u 2 V.

hu; 0iD 0 for every u 2 V.

hu; *v* C *w*iD hu; *v*iC hu; *w*i for all u; *v*; *w* 2 V.

hu; 入*v*iD 入N hu; *v*i for all 入 2 **F** and u; *v* 2 V.

Proof

1. Part (a) follows from the conditions of additivity in the ﬁrst slot and homogeneity in the ﬁrst slot in the deﬁnition of an inner product.
2. Part (b) follows from part (a) and the result that every linear map takes

0 to 0.

**168** CHAPTER 6 Inner Product Spaces

1. Part (c) follows from part (a) and the conjugate symmetry property in the deﬁnition of an inner product.
2. Suppose u; *v*; *w* 2 V. Then

hu; *v* C *w*iD h*v* C *w*; ui

D h*v*; ui C h*w*; ui D h*v*; uiC h*w*; ui D hu; *v*iC hu; *w*i:

1. Suppose 入 2 **F** and u; *v* 2 V. Then

hu; 入*v*iD h入*v*; ui

D 入h*v*; ui D 入N h*v*; ui D 入N hu; *v*i;

as desired.

### Norms

Our motivation for deﬁning inner products came initially from the norms of vectors on **R**2 and **R**3. Now we see that each inner product determines a norm.

6.8 **Deﬁnition *norm,*** k*v*k

For *v* 2 V, the ***norm*** of *v*, denoted k*v*k, is deﬁned by

k*v*kD ph*v*; *v*i:

6.9 **Example *norms***

1. If .z1;:::; zn/ 2 **F**n (with the Euclidean inner product), then

q

k.z1;:::; zn/kD jz1j2 C. . .C jznj2:

1. In the vector space of continuous real-valued functions on Œ 1; 1] [with

—

inner product given as in part (c) of 6.4], we have

sZ 1 （ ）2

kf kD

f .x/

dx:

1

SECTION 6.A Inner Products and Norms **169**

6.10 Basic properties of the norm

Suppose *v* 2 V.

1. k*v*kD 0 if and only if *v* D 0.
2. k入*v*kD j入j k*v*k for all 入 2 **F**.

Proof

1. The desired result holds because h*v*; *v*iD 0 if and only if *v* D 0.
2. Suppose 入 2 **F**. Then

k入*v*k2

D h入*v*; 入*v*i D 入h*v*; 入*v*i D 入入N h*v*; *v*i

D j入j k*v*k2:

2

Taking square roots now gives the desired equality.

The proof above of part (b) illustrates a general principle: working with norms squared is usually easier than working directly with norms.

Now we come to a crucial deﬁnition.

6.11 **Deﬁnition *orthogonal***

Two vectors u; *v* 2 V are called ***orthogonal*** if hu; *v*iD 0.

In the deﬁnition above, the order of the vectors does not matter, because u; *v* 0 if and only if *v*; u 0. Instead of saying that u and *v* are orthogonal, sometimes we say that u is orthogonal to *v*.

Exercise 13 asks you to prove that if u; *v* are nonzero vectors in **R**2, then

h i D h i D

hu; *v*iD kukk*v*k cos 0;

where 0 is the angle between u and *v* (thinking of u and *v* as arrows with initial point at the origin). Thus two vectors in **R**2 are orthogonal (with respect to the

usual Euclidean inner product) if and only if the cosine of the angle between them is 0, which happens if and only if the vectors are perpendicular in the usual sense of plane geometry. Thus you can think of the word *orthogonal* as

a fancy word meaning *perpendicular*.

**170** CHAPTER 6 Inner Product Spaces

We begin our study of orthogonality with an easy result.

6.12 Orthogonality and 0

1. 0 is orthogonal to every vector in V.
2. 0 is the only vector in V that is orthogonal to itself.

Proof

1. Part (b) of 6.7 states that h0; uiD 0 for every u 2 V.
2. If *v* 2 V and h*v*; *v*iD 0, then *v* D 0 (by deﬁnition of inner product).

For the special case V **R**2, the next theorem is over 2,500 years old. Of course, the proof below is not the

*The word* ***orthogonal*** *comes from the Greek word* ***orthogonios****, which means right-angled.*

D

original proof.

6.13 Pythagorean Theorem

Suppose u and *v* are orthogonal vectors in V. Then

ku C *v*k2 D kuk2 C k*v*k2:

Proof We have

ku C *v*k

2

D hu C *v*;u C *v*i

D hu; uiC hu; *v*iC h*v*; uiC h*v*; *v*i

2

as desired.

D kuk C k*v*k2;

Suppose u; *v* 2 V, with *v* ¤ 0. We would like to write u as a scalar multiple of *v* plus a vector *w* orthogonal to *v*, as

*The proof given above of the Pythagorean Theorem shows that the conclusion holds if and only*

*if* hu; *v*i C h*v*; ui*, which equals*

2 Rehu; *v*i*, is* 0*. Thus the converse*

*of the Pythagorean Theorem holds in real inner product spaces.*

suggested in the next picture.

SECTION 6.A Inner Products and Norms **171**

*u*

*w*

*cv*

*v*

0

*An orthogonal decomposition.*

To discover how to write u as a scalar multiple of *v* plus a vector orthogonal to *v*, let c 2 **F** denote a scalar. Then

u D c*v* C .u — c*v*/:

Thus we need to choose c so that *v* is orthogonal to .u c*v*/. In other words, we want

—

0 D hu — c*v*; *v*iD hu; *v*i— ck*v*k2:

The equation above shows that we should choose c to be u; *v* = *v* 2. Making this choice of c, we can write

h i k k

u D hu; *v*i *v* C u — hu; *v*i *v* :

k*v*k2 k*v*k2

As you should verify, the equation above writes u as a scalar multiple of *v*

plus a vector orthogonal to *v*. In other words, we have proved the following

result.

*French mathematician Augustin- Louis Cauchy* (*1789–1857*) *proved 6.17*(*a*) *in 1821. German mathe- matician Hermann Schwarz* (*1843– 1921*) *proved 6.17*(*b*) *in 1886.*

6.14 An orthogonal decomposition

Suppose u; *v* 2 V, with *v* ¤ 0. Set c D hu; *v*i and *w* D u — hu; *v*i *v*. Then

k*v*k2

h*w*; *v*iD 0 and u D c*v* C *w*:

k*v*k2

The orthogonal decomposition 6.14 will be used in the proof of the Cauchy– Schwarz Inequality, which is our next result and is one of the most important inequalities in mathematics.

**172** CHAPTER 6 Inner Product Spaces

6.15 Cauchy–Schwarz Inequality

Suppose u; *v* 2 V. Then

jhu; *v*ij三 kuk k*v*k:

This inequality is an equality if and only if one of u; *v* is a scalar multiple of the other.

Proof If *v* D 0, then both sides of the desired inequality equal 0. Thus we can assume that *v* ¤ 0. Consider the orthogonal decomposition

u hu; *v*i *v w*

D C

k*v*k2

given by 6.14, where *w* is orthogonal to *v*. By the Pythagorean Theorem,

kuk2 D 11 hu; *v*i

k*v* 2k

2

11 C k*w*k

*v*

11

2

#### 6.16

u; *v* 2 2

D k*v*k2 C k*w*k

jh ij

jhu; *v*ij2

2 :

乏

k*v*k

Multiplying both sides of this inequality by *v* 2 and then taking square roots gives the desired inequality.

k k

Looking at the proof in the paragraph above, note that the Cauchy–Schwarz Inequality is an equality if and only if 6.16 is an equality. Obviously this

happens if and only if *w* 0. But *w* 0 if and only if u is a multiple of *v*

D D

(see 6.14). Thus the Cauchy–Schwarz Inequality is an equality if and only if u is a scalar multiple of *v* or *v* is a scalar multiple of u (or both; the phrasing has been chosen to cover cases in which either u or *v* equals 0).

6.17 **Example *examples of the Cauchy–Schwarz Inequality***

1. If x1;:::; xn; y1;:::; yn 2 **R**, then

2

2

2

2

2

jx1y1 C ... C xnynj 三 .x1 C ... C xn /.y1 C ... C yn /:

1. If f; g are continuous real-valued functions on Œ—1; 1], then

f .x/g.x/ dxˇ

三

f .x/

dx

g.x/

dx

:

ˇZ 1

ˇ

1

ˇ2 Z 1 （

1

）2 Z 1 （ ）2

1

SECTION 6.A Inner Products and Norms **173**

The next result, called the Triangle Inequality, has the geometric interpreta- tion that the length of each side of a tri- angle is less than the sum of the lengths of the other two sides.

*v*

*u*

*u* + *v*

Note that the Triangle Inequality im- plies that the shortest path between two points is a line segment.

6.18 Triangle Inequality

Suppose u; *v* 2 V. Then

ku C *v*k三 kukC k*v*k:

This inequality is an equality if and only if one of u; *v* is a nonnegative multiple of the other.

Proof We have

ku C *v*k

2

D hu C *v*;u C *v*i

D hu; uiC h*v*; *v*iC hu; *v*iC h*v*; ui

D hu; uiC h*v*; *v*iC hu; *v*iC hu; *v*i

D kuk C k*v*k2 C 2 Rehu; *v*i

2

#### 6.19

2 2

**6.20**

三 kuk C k*v*k C 2jhu; *v*ij 三 kuk C k*v*k C 2kuk k*v*k

2

2 2

D .kukC k*v*k/ ;

where 6.20 follows from the Cauchy–Schwarz Inequality (6.15). Taking square roots of both sides of the inequality above gives the desired inequality. The proof above shows that the Triangle Inequality is an equality if and only if we have equality in 6.19 and 6.20. Thus we have equality in the

Triangle Inequality if and only if

**6.21** hu; *v*iD kukk*v*k:

If one of u; *v* is a nonnegative multiple of the other, then 6.21 holds, as you should verify. Conversely, suppose 6.21 holds. Then the condition for equality in the Cauchy–Schwarz Inequality (6.15) implies that one of u; *v* is a scalar multiple of the other. Clearly 6.21 forces the scalar in question to be

nonnegative, as desired.

**174** CHAPTER 6 Inner Product Spaces

The next result is called the parallelogram equality because of its geometric interpretation: in every parallelogram, the sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides.

*u*



*u* 一 *v*

*v*

*v*

*u* + *v*

*u*

*The parallelogram equality.*

6.22 Parallelogram Equality

Suppose u; *v* 2 V. Then

ku C *v*k2 C ku — *v*k2 D 2.kuk2 C k*v*k2/:

Proof We have

ku C *v*k2

C ku — *v*k

D hu C *v*;u C *v*iC hu — *v*;u — *v*i

D kuk

2

2

C k*v*k2

2

C hu; *v*iC h*v*; ui

2

C kuk

2

C k*v*k

2

— hu; *v*i— h*v*; ui

as desired.

D 2.kuk

C k*v*k /;

Law professor Richard Friedman presenting a case before the U.S. Supreme Court in 2010:

*Mr. Friedman*: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging—

*Chief Justice Roberts*: I’m sorry. Entirely what?

*Mr. Friedman*: Orthogonal. Right angle. Unrelated. Irrelevant.

*Chief Justice Roberts*: Oh.

*Justice Scalia*: What was that adjective? I liked that.

*Mr. Friedman*: Orthogonal.

*Chief Justice Roberts*: Orthogonal.

*Mr. Friedman*: Right, right.

*Justice Scalia*: Orthogonal, ooh. (Laughter.)

*Justice Kennedy*: I knew this case presented us a problem. (Laughter.)

EXERCISES 6.A

SECTION 6.A Inner Products and Norms **175**

1. Show that the function that takes .x1; x2/; .y1; y2/ 2 **R**2 x **R**2 to

（ ）

jx1y1jC jx2y2j is not an inner product on **R**2.

（ ）

1. Show that the function that takes .x1; x2; x3/; .y1; y2; y3/ 2 **R**3 x **R**3

to x1y1 C x3y3 is not an inner product on **R**3.

1. Suppose **F R** and V 0 . Replace the positivity condition (which states that *v*; *v* 0 for all *v* V ) in the deﬁnition of an inner product (6.3) with the condition that *v*; *v* >0 for some *v* V. Show that this change in the deﬁnition does not change the set of functions from V V to **R** that are inner products on V.

x

h i 2

h i乏 2

D ¤ f g

1. Suppose V is a real inner product space.
   1. Show that hu C *v*;u — *v*iD kuk2 — k*v*k2 for every u; *v* 2 V.
   2. Show that if u; *v* 2 V have the same norm, then uC*v* is orthogonal to u — *v*.
   3. Use part (b) to show that the diagonals of a rhombus are perpen- dicular to each other.
2. Supppose T 2 *L*.V / is such that kT *v*k三 k*v*k for every *v* 2 V. Prove that

T — 2I is invertible.

1. Suppose u; *v* 2 V. Prove that hu; *v*iD 0 if and only if

kuk三 ku C a*v*k

for all a 2 **F**.

1. Suppose u; *v* 2 V. Prove that kau C b*v*kD kbu C a*v*k for all a; b 2 **R**

if and only if kukD k*v*k.

1. Suppose u; *v* 2 V and kukD k*v*kD 1 and hu; *v*iD 1. Prove that u D *v*.
2. Suppose u; *v* 2 V and kuk三 1 and k*v*k三 1. Prove that

q1 — kuk2q1 — k*v*k2 三 1 — jhu; *v*ij:

1. Find vectors u; *v* 2 **R**2 such that u is a scalar multiple of .1; 3/, *v* is orthogonal to .1; 3/, and .1; 2/ D u C *v*.

**176** CHAPTER 6 Inner Product Spaces

**11** Prove that

16 三 .a C b C c C d/ 1 1 1 1

a C b C c C d

for all positive numbers a; b; c; d.

1. Prove that

.x1 C ... C xn/2 三 n.x12 C ... C xn2/

for all positive integers n and all real numbers x1;:::; xn.

1. Suppose u; *v* are nonzero vectors in **R**2. Prove that

hu; *v*iD kukk*v*k cos 0;

where 0 is the angle between u and *v* (thinking of u and *v* as arrows with initial point at the origin).

*Hint:* Draw the triangle formed by u, *v*, and u *v*; then use the law of cosines.

—

1. The angle between two vectors (thought of as arrows with initial point at the origin) in **R**2 or **R**3 can be deﬁned geometrically. However, geometry is not as clear in **R**n for n > 3. Thus the angle between two nonzero

vectors x; y 2 **R**n is deﬁned to be

arccos hx; yi ;

kxkkyk

where the motivation for this deﬁnition comes from the previous exercise. Explain why the Cauchy–Schwarz Inequality is needed to show that this deﬁnition makes sense.

1. Prove that

X

X

X

n

j D1

2

aj bj 三

n

j D1

jaj 2

n

j D1

bj 2

j

for all real numbers a1;:::; an and b1;:::; bn.

1. Suppose u; *v* 2 V are such that

kukD 3; ku C *v*kD 4; ku — *v*kD 6:

What number does k*v*k equal?

SECTION 6.A Inner Products and Norms **177**

1. Prove or disprove: there is an inner product on **R**2 such that the associated norm is given by

for all .x; y/ 2 **R**2.

k.x; y/kD maxfx; yg

1. Suppose p> 0. Prove that there is an inner product on **R**2 such that the associated norm is given by

k.x; y/kD .x

p

C yp/1=p

for all .x; y/ 2 **R**2 if and only if p D 2.

1. Suppose V is a real inner product space. Prove that

ku C *v*k2 — ku — *v*k2

hu; *v*iD 4

for all u; *v* 2 V.

1. Suppose V is a complex inner product space. Prove that

ku C *v*k2 — ku — *v*k2 C ku C i*v*k2i — ku — i*v*k2i

hu; *v*iD 4

for all u; *v* 2 V.

1. A norm on a vector space U is a function U Œ0; / such that u 0 if and only if u 0, ˛u ˛ u for all ˛ **F** and all u U, and u *v* u *v* for all u; *v* U. Prove

2 k C k 三 k kC k k 2

k k D D k k D j jk k 2

k k W ! 1

that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if kk is a norm on U satisfying the parallelogram equality, then there is an inner product h ; i on U such that kukD hu; ui1=2 for all u 2 U ).

1. Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if a1;:::; an **R**, then the square of the average of a1;:::; an is less than or equal to the average of a12;:::; an2.

2

1. Suppose V1;:::; Vm are inner product spaces. Show that the equation

h.u1;:::; um/; .*v*1;:::; *v*m/iD hu1; *v*1iC. . .C hum; *v*mi

deﬁnes an inner product on V1 Vm.

x ... x

[*In the expression above on the right,* u1; *v*1 *denotes the inner product on* V1*, ...,* um; *v*m *denotes the inner product on* Vm*. Each of the spaces* V1;:::; Vm *may have a different inner product, even though the same*

h i

h i

*notation is used here.*]

**178** CHAPTER 6 Inner Product Spaces

1. Suppose S 2 *L*.V / is an injective operator on V. Deﬁne h.; .i1 by

hu; *v*i1 D hSu; S *v*i

for u; *v* 2 V. Show that h.; .i1 is an inner product on V.

1. Suppose S 2 *L*.V / is not injective. Deﬁne h.; .i1 as in the exercise above. Explain why h.; .i1 is not an inner product on V.
2. Suppose f; g are differentiable functions from **R** to **R**n.
   1. Show that

hf .t/; g.t/i0 D hf 0.t/; g.t/iC hf .t/; g0.t/i:

* 1. Suppose c > 0 and kf .t/k D c for every t 2 **R**. Show that

hf 0.t/; f .t/iD 0 for every t 2 **R**.

* 1. Interpret the result in part (b) geometrically in terms of the tangent vector to a curve lying on a sphere in **R**n centered at the origin.

[*For the exercise above, a function* f **R R**n *is called differentiable if there exist differentiable functions* f1;:::; fn *from* **R** *to* **R** *such that*

W !

（ ）

f .t/ D f1.t/; : : : ; fn.t/ *for each* t 2 **R***. Furthe*（*rmore, for each* t 2） **R***,*

*the derivative* f 0.t/ 2 **R**n *is deﬁned by* f 0.t/ D

1. Suppose u; *v*; *w* 2 V. Prove that

:

f10.t/; : : : ; fn0.t/ *.*]

k*w* — 2 .u C *v*/k D

1

2

k*w* — uk2 C k*w* — *v*k2

2

ku — *v*k2

4

1. Suppose C is a subset of V with the property that u; *v* C implies  1 .u *v*/ C . Let *w* V. Show that there is at most one point in C that is closest to *w*. In other words, show that there is at most one u C

—

2

2

C 2 2

2

such that

k*w* — uk三 k*w* — *v*k for all *v* 2 C .

*Hint:* Use the previous exercise.

1. For u; *v* 2 V, deﬁne d.u; *v*/ D ku — *v*k.
   1. Show that d is a metric on V.
   2. Show that if V is ﬁnite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).
   3. Show that every ﬁnite-dimensional subspace of V is a closed subset of V (with respect to the metric d ).

SECTION 6.A Inner Products and Norms **179**

1. Fix a positive integer n. The ***Laplacian*** �p of a twice differentiable function p on **R**n is the function on **R**n deﬁned by

@2p @2p

�p D @x2 C ... C @x2 :

1 n

The function p is called ***harmonic*** if �p D 0.

A ***polynomial*** on **R**n is a linear combination of functions of the form x1m*1* ... xnm*n* , where m1;:::; mn are nonnegative integers.

Suppose q is a polynomial on **R**n. Prove that there exists a harmonic polynomial p on **R**n such that p.x/ D q.x/ for every x 2 **R**n with x 1

k kD

.

[*The only fact about harmonic functions that you need for this exercise is that if* p *is a harmonic function on* **R**n *and* p.x/ D 0 *for all* x 2 **R**n *with* kxkD 1*, then* p D 0*.*]

*Hint:* A reasonable guess is that the desired harmonic polynomial p is of the form q .1 x 2/r for some polynomial r. Prove that there is a polynomial r on **R**n such that q .1 x 2/r is harmonic by deﬁning an operator T on a suitable vector space by

（ ）

C — k k

C — k k

Tr D � .1 — kxk2/r

and then showing that T is injective and hence surjective.

1. Use inner products to prove Apollonius’s Identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

a2 C b2 D 1 c2 C 2d 2:

2

*c*

*a*

*d b*

**180** CHAPTER 6 Inner Product Spaces

## *Orthonormal Bases*

6.B

6.23 **Deﬁnition *orthonormal***

* A list of vectors is called ***orthonormal*** if each vector in the list has

norm 1 and is orthogonal to all the other vectors in the list.

* In other words, a list e1;:::; em of vectors in V is orthonormal if

h

e ;e iD

j

k

(

0 if j ¤ k.

1 if j D k,

6.24 **Example *orthonormal lists***

1. The standard basis in **F**n is an orthonormal list.

p

p

6

p

6

2

2

6

1. （ p1

3

2

; 1 ; 1

3 3

p

p

2

）; （— p1

; 1 ; 0） is an orthonormal list in **F**3.

1. （ p1

3

; 1 ; 1

3 3

p

p

）; （— p1

; 1 ; 0）; （ p1 ; 1 ; — p2 ） is an orthonormal list

in **F**3.

Orthonormal lists are particularly easy to work with, as illustrated by the next result.

6.25 The norm of an orthonormal linear combination

If e1;:::; em is an orthonormal list of vectors in V, then

ka e C ... C a

1

1

m m

e k D j

2

a j C. . .C j

2

2

1

a

m

j

for all a1;:::; am 2 **F**.

Proof Because each ej has norm 1, this follows easily from repeated appli- cations of the Pythagorean Theorem (6.13).

The result above has the following important corollary.

6.26 An orthonormal list is linearly independent

Every orthonormal list of vectors is linearly independent.

SECTION 6.B Orthonormal Bases **181**

Proof Suppose e1;:::; em is an orthonormal list of vectors in V and

a1;:::; am 2 **F** are such that

a1e1 C C amem D 0:

Then a1 2 am 2 0 (by 6.25), which means that all the aj ’s are 0. Thus e1;:::; em is linearly independent.

j j C ... C j j D

6.27 **Deﬁnition *orthonormal basis***

An ***orthonormal basis*** of V is an orthonormal list of vectors in V that is also a basis of V.

For example, the standard basis is an orthonormal basis of **F**n.

6.28 An orthonormal list of the right length is an orthonormal basis

Every orthonormal list of vectors in V with length dim V is an orthonormal basis of V.

Proof By 6.26, any such list must be linearly independent; because it has the right length, it is a basis—see 2.39.

6.29 **Example** Show that

（ 1 ; 1 ; 1 ; 1 ）; （ 1 ; 1 ; — 1 ; — 1 ）; （ 1 ; — 1 ; — 1 ; 1 ）; （— 1 ; 1 ; — 1 ; 1 ）

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

is an orthonormal basis of **F**4.

Solution We have

11（ 1 ; 1 ; 1 ; 1 ）11 D q（ 1 ）2 C （ 1 ）2 C （ 1 ）2 C （ 1 ）2 D 1:

2

2

2

2

2

2

2

2

Similarly, the other three vectors in the list above also have norm 1.

We have

˝（ 1 ; 1 ; 1 ; 1 ）; （ 1 ; 1 ; — 1 ; — 1 ）˛ D 1 . 1 C 1 . 1 C 1 . （— 1 ） C 1 . （— 1 ） D 0:

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

Similarly, the inner product of any two distinct vectors in the list above also equals 0.

Thus the list above is orthonormal. Because we have an orthonormal list of length four in the four-dimensional vector space **F**4, this list is an orthonormal basis of **F**4 (by 6.28).

**182** CHAPTER 6 Inner Product Spaces

In general, given a basis e1;:::; en of V and a vector *v* 2 V, we know that there is some choice of scalars a1;:::; an 2 **F** such that

*v* D a1e1 C ... C anen:

Computing the numbers a1;:::; an that satisfy the equation above can be difﬁ- cult for an arbitrary basis of V. The next result shows, however, that this is

*The importance of orthonormal bases stems mainly from the next result.*

easy for an orthonormal basis—just take

aj D h*v*; ej i.

6.30 Writing a vector as linear combination of orthonormal basis

Suppose e1;:::; en is an orthonormal basis of V and *v* 2 V. Then

*v* D h*v*; e1ie1 C. . .C h*v*; enien

and

k

*v*k D jh*v*; e ij C. . .C jh

2

2

*v*

2

1

; e ij :

n

Proof Because e1;:::; en is a basis of V, there exist scalars a1;:::; an such that

*v* D a1e1 C ... C anen:

Because e1;:::; en is orthonormal, taking the inner product of both sides of this equation with ej gives *v*; ej aj . Thus the ﬁrst equation in 6.30 holds.

h iD

The second equation in 6.30 follows immediately from the ﬁrst equation and 6.25.

Now that we understand the usefulness of orthonormal bases, how do we go about ﬁnding them? For example, does m.**R**/, with inner product given by integration on Œ 1; 1] [see 6.4(c)], have an orthonormal basis? The next

—

*P*

result will lead to answers to these questions.

The algorithm used in the next proof is called the ***Gram–Schmidt Procedure***. It gives a method for turning a linearly independent list into an orthonormal list with the same span as the original list.

*Danish mathematician Jørgen Gram (1850–1916) and German mathematician Erhard Schmidt (1876–1959) popularized this algo- rithm that constructs orthonormal lists.*

SECTION 6.B Orthonormal Bases **183**

6.31 Gram–Schmidt Procedure

Suppose *v*1;:::; *v*m is a linearly independent list of vectors in V. Let

e1 D *v*1=k*v*1k. For j D 2; :::; m, deﬁne ej inductively by

ej D k*v*

*v*j — h*v*j ; e1ie1 —. . .— h*v*j ; ej  1iej  1

j j 1 1

— h*v* ; e ie —. . .— h*v* ; e ie

j j 1 j 1

:

k

Then e1;:::; em is an orthonormal list of vectors in V such that

span.*v*1;:::; *v*j / D span.e1;:::; ej /

for j D 1; :::; m.

Proof We will show by induction on j that the desired conclusion holds. To get started with j 1, note that span.*v*1/ span.e1/ because *v*1 is a positive multiple of e1.

Suppose 1<j <m and we have veriﬁed that

D D

**6.32** span.*v*1;:::; *v*j 1/ D span.e1;:::; ej 1/:

Note that *v*j … span.*v*1;:::; *v*j  1/ (because *v*1;:::; *v*m is linearly indepen- dent). Thus *v*j span.e1;:::; ej 1/. Hence we are not dividing by 0 in the deﬁnition of ej given in 6.31. Dividing a vector by its norm produces a new vector with norm 1; thus kej kD 1.

…

Let 1 三 k< j . Then

he ;e iD  *v*j — h*v*j ; e1ie1 —. . .— h*v*j ; ej  1iej  1 ;e

j k

k

k*v*j — h*v*j ; e1ie1 —. . .— h*v*j ; ej  1iej  1k

h*v*j ; eki— h*v*j ; eki

D

k*v*j — h*v*j ; e1ie1 —. . .— h*v*j ; ej  1iej  1k

D 0:

Thus e1;:::; ej is an orthonormal list.

From the deﬁnition of ej given in 6.31, we see that *v*j span.e1;:::; ej /.

2

Combining this information with 6.32 shows that

span.*v*1;:::; *v*j / c span.e1;:::; ej /:

Both lists above are linearly independent (the *v*’s by hypothesis, the e’s by orthonormality and 6.26). Thus both subspaces above have dimension j , and hence they are equal, completing the proof.

**184** CHAPTER 6 Inner Product Spaces

6.33 **Example** Find aRn orthonormal basis of *P*2.**R**/, where the inner prod-

Solution We will apply the Gram–Schmidt Procedure (6.31) to the basis

uct is given by hp; qiD

1

1

p.x/q.x/ dx.

1; x; x2.

To get started, with this inner product we have

k1k D

2

1

— 1 dx D 2:

Z

2

1

p 1

qThus k1kD 2, and hence e D .1

2

Now the numerator in the expression for e2 is

We have

2

x — hx; e1ie1 D x —

1

x 1 dx

Z

2

q

1

q 1 D x:

kxk D

2

1

x2 dx 2 :

Z

3

D

1

Thus kxkD q 2 , and hence e2 D q 3 x.

Now the numerator in the expression for e3 is

3

2

x2 — hx2; e1ie1 — hx2; e2ie2

Z

2

Z

q

q

D x —

2

2

2

1

x2 1 dx

2

1

1

q 1 —

1

x2 3 x dx

2

1

q 3 x

We have

x

— 3 x

C 9

dx D 45 :

D x — 3 :

2 1 2

kx

— 3 k

D

Thus kx2 — 1 kD q 8 , and hence e3 D q 45 （x2 — 1 ）.

Z 1 （ 4 2 2 1 ） 8

1

3

Thus

45 8 3

q 1 ; q 3 x; q 45 （x2 — 1 ）

2

2

8

3

is an orthonormal list of length 3 in *P*2.**R**/. Hence this orthonormal list is an orthonormal basis of *P*2.**R**/ by 6.28.

SECTION 6.B Orthonormal Bases **185**

Now we can answer the question about the existence of orthonormal bases.

6.34 Existence of orthonormal basis

Every ﬁnite-dimensional inner product space has an orthonormal basis.

Proof Suppose V is ﬁnite-dimensional. Choose a basis of V. Apply the Gram–Schmidt Procedure (6.31) to it, producing an orthonormal list with length dim V. By 6.28, this orthonormal list is an orthonormal basis of V.

Sometimes we need to know not only that an orthonormal basis exists, but also that every orthonormal list can be extended to an orthonormal basis. In the next corollary, the Gram–Schmidt Procedure shows that such an extension is always possible.

6.35 Orthonormal list extends to orthonormal basis

Suppose V is ﬁnite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proof Suppose e1;:::; em is an orthonormal list of vectors in V. Then e1;:::; em is linearly independent (by 6.26). Hence this list can be extended to a basis e1;:::; em; *v*1;:::; *v*n of V (see 2.33). Now apply the Gram–Schmidt Procedure (6.31) to e1;:::; em; *v*1;:::; *v*n, producing an orthonormal list

**6.36** e1;:::; em; f1;:::; fnI

here the formula given by the Gram–Schmidt Procedure leaves the ﬁrst m vectors unchanged because they are already orthonormal. The list above is an orthonormal basis of V by 6.28.

Recall that a matrix is called upper triangular if all entries below the diagonal equal 0. In other words, an upper-triangular matrix looks like this:

0B@ \*

: : :

\* 1C ;

0 \*

A

where the 0 in the matrix above indicates that all entries below the diagonal equal 0, and asterisks are used to denote entries on and above the diagonal.

**186** CHAPTER 6 Inner Product Spaces

In the last chapter we showed that if V is a ﬁnite-dimensional complex vector space, then for each operator on V there is a basis with respect to which the matrix of the operator is upper triangular (see 5.27). Now that we

are dealing with inner product spaces, we would like to know whether there exists an *orthonormal* basis with respect to which we have an upper-triangular matrix.

The next result shows that the existence of a basis with respect to which T has an upper-triangular matrix implies the existence of an orthonormal basis with this property. This result is true on both real and complex vector

spaces (although on a real vector space, the hypothesis holds only for some operators).

6.37 Upper-triangular matrix with respect to orthonormal basis

some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Suppose T 2 *L*.V /. If T has an upper-triangular matrix with respect to

Proof Suppose T has an upper-triangular matrix with respect to some basis *v*1;:::; *v*n of V. Thus span.*v*1;:::; *v*j / is invariant under T for each j 1; :::;n (see 5.26).

Apply the Gram–Schmidt Procedure to *v*1;:::; *v*n, producing an orthonor- mal basis e1;:::; en of V. Because

D

span.e1;:::; ej / D span.*v*1;:::; *v*j /

for each j (see 6.31), we conclude that span.e1;:::; ej / is invariant under T for each j 1; :::; n. Thus, by 5.26, T has an upper-triangular matrix with respect to the orthonormal basis e1;:::; en.

D

The next result is an important appli- cation of the result above.

*German mathematician Issai Schur* (*1875–1941*) *published the ﬁrst proof of the next result in 1909.*

6.38 Schur’s Theorem

Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Suppose V is a ﬁnite-dimensional complex vector space and T 2 *L*.V /.

Proof Recall that T has an upper-triangular matrix with respect to some basis of V (see 5.27). Now apply 6.37.

SECTION 6.B Orthonormal Bases **187**

### Linear Functionals on Inner Product Spaces

Because linear maps into the scalar ﬁeld **F** play a special role, we deﬁned a special name for them in Section 3.F. That deﬁnition is repeated below in case you skipped Section 3.F.

6.39 **Deﬁnition *linear functional***

A ***linear functional*** on V is a linear map from V to **F**. In other words, a linear functional is an element of *L*.V; **F**/.

* 1. **Example** The function ' W **F**3 ! **F** deﬁned by

'.z1; z2; z3/ D 2z1 — 5z2 C z3

is a linear functional on **F**3. We could write this linear functional in the form

'.z/ D hz; ui

for every z 2 **F**3, where u D .2; —5; 1/.

Z

（

）

* 1. **Example** The function ' W *P*2.**R**/ ! **R** deﬁned by

'.p/ D

1

p.t/ cos.兀t/ dt

1

is a linear functional on *P*2.**R**/ (here the inner product on *P*2.**R**/ is multi- plication followed by integration on Œ—1; 1]; see 6.33). It is not obvious that there exists u 2 *P*2.**R**/ such that

'.p/ D hp; ui

for every p 2 *P*2.**R**/ [we cannot take u.t/ D cos.兀t/ because that function is not an element of *P*2.**R**/].

*The next result is named in honor of Hungarian mathematician Frigyes Riesz* (*1880–1956*)*, who proved several results early in the twen- tieth century that look very much like the result below.*

2

If u V, then the map that sends *v* to *v*; u is a linear functional on V. The next result shows that every linear

functional on V is of this form. Ex-

h i

ample 6.41 above illustrates the power

of the next result because for the linear functional in that example, there is no

obvious candidate for u.

**188** CHAPTER 6 Inner Product Spaces

6.42 Riesz Representation Theorem

Suppose V is ﬁnite-dimensional and ' is a linear functional on V. Then there is a unique vector u 2 V such that

'.*v*/ D h*v*; ui

for every *v* 2 V.

Proof First we show there exists a vector u 2 V such that '.*v*/ D h*v*; ui for every *v* 2 V. Let e1;:::; en be an orthonormal basis of V. Then

'.*v*/ D '.h*v*; e1ie1 C. . .C h*v*; enien/

D h*v*; e1i'.e1/ C. . .C h*v*; eni'.en/

D h*v*; '.e1/e1 C ... C '.en/eni

for every *v* 2 V, where the ﬁrst equality comes from 6.30. Thus setting

**6.43** u D '.e1/e1 C ... C '.en/en;

we have '.*v*/ D h*v*; ui for every *v* 2 V, as desired.

Now we prove that only one vector u 2 V has the desired behavior.

Suppose u1; u2 2 V are such that

'.*v*/ D h*v*; u1iD h*v*; u2i

for every *v* 2 V. Then

0 D h*v*; u1i— h*v*; u2iD h*v*; u1 — u2i

for every *v* 2 V. Taking *v* D u1 — u2 shows that u1 — u2 D 0. In other words,

u1 D u2, completing the proof of the uniqueness part of the result.

Z

Z

6.44 **Example** Find u 2 *P*2.**R**/ such that

1

1

for every p 2 *P*2.**R**/.

p.t/（cos.兀t/）

dt D

1

p.t/u.t/ dt

1

SECTION 6.B Orthonormal Bases **189**

Solution Let '.p/ D R 1 p.t/（cos.兀t/） dt. Applying formula 6.43 from

1

the proof above, and using the orthonormal basis from Example 6.33, we have

Z 1 q 1 （

u.x/ D

cos.兀t/

1

2

） q 1

Z 1 q 3 （

1

） q 3

Z 1 q 45 （ 2 1 ）（ ） q 45 （ 2 1 ）

dt

2 C

2 t

cos.兀t/

dt

2 x

C

1

8

t

— 3

cos.兀t/

dt

8

x

— 3

:

A bit of calculus shows that

u.x/ D— 45 （x2 — 1 ）:

D h i 2

2冗*2*

3

Suppose V is ﬁnite-dimensional and ' a linear functional on V. Then 6.43 gives a formula for the vector u that satisﬁes '.*v*/ *v*; u for all *v* V. Speciﬁcally, we have

u D '.e1/e1 C ... C '.en/en:

The right side of the equation above seems to depend on the orthonormal basis e1;:::; en as well as on '. However, 6.42 tells us that u is uniquely determined by '. Thus the right side of the equation above is the same regardless of which orthonormal basis e1;:::; en of V is chosen.

EXERCISES 6.B

1. (a) Suppose 0 2 **R**. Show that .cos 0; sin 0 /; .— sin 0; cos 0/ and

.cos 0; sin 0 /; .sin 0; — cos 0/ are orthonormal bases of **R**2.

(b) Show that each orthonormal basis of **R**2 is of the form given by one of the two possibilities of part (a).

1. Suppose e1;:::; em is an orthonormal list of vectors in V. Let *v* V. Prove that

2

k*v*k

2

2

D jh*v*; e1ij

C. . .C jh*v*; emij2

if and only if *v* 2 span.e1;:::; em/.

1. Suppose T .**R**3/ has an upper-triangular matrix with respect to the basis .1; 0; 0/, (1, 1, 1), .1; 1; 2/. Find an orthonormal basis of **R**3 (use the usual inner product on **R**3) with respect to which T has an

2 *L*

upper-triangular matrix.

**190** CHAPTER 6 Inner Product Spaces

1. Suppose n is a positive integer. Prove that

1 cos x

cos 2x

cos nx

sin x

sin 2x

sin nx

p2兀;

p兀 ;

p兀 ;:::;

p兀 ; p兀 ;

p兀 ;:::; p兀

is an orthonormal list of vectors in C Œ—兀; 兀], the vector space of contin- uous real-valued functions on Œ—兀; 兀] with inner product

Z 冗

冗

hf; giD

f .x/g.x/ dx:

[*The orthonormal list above is often used for modeling periodic phenom- ena such as tides.*]

1. On *P*2.**R**/, consider the inner product given by

Z 1

0

hp; qiD

p.x/q.x/ dx:

Apply the Gram–Schmidt Procedure to the basis 1; x; x2 to produce an orthonormal basis of *P*2.**R**/.

1. Find an orthonormal basis of 2.**R**/ (with inner product as in Exercise 5) such that the differentiation operator (the operator that takes p to p0) on *P*2.**R**/ has an upper-triangular matrix with respect to this basis.

*P*

1. Find a polynomial q 2 *P*2.**R**/ such that

Z

for every p 2 *P*2.**R**/.

2

p（ 1 ） D

1

p.x/q.x/ dx

0

1. Find a polynomial q 2 *P*2.**R**/ such that

Z

Z

1

p.x/.cos 兀x/ dx

D

0

1

p.x/q.x/ dx

0

for every p 2 *P*2.**R**/.

1. What happens if the Gram–Schmidt Procedure is applied to a list of vectors that is not linearly independent?

SECTION 6.B Orthonormal Bases **191**

1. Suppose V is a real inner product space and *v*1;:::; *v*m is a linearly inde- pendent list of vectors in V. Prove that there exist exactly 2m orthonormal lists e1;:::; em of vectors in V such that

span.*v*1;:::; *v*j / D span.e1;:::; ej /

for all j 2 f1; :::; mg.

1. Suppose h.; .i1 and h.; .i2 are inner products on V such that h*v*; *w*i1 D 0 if and only if h*v*; *w*i2 D 0. Prove that there is a positive number c such that h*v*; *w*i1 D ch*v*; *w*i2 for every *v*; *w* 2 V.
2. Suppose V is ﬁnite-dimensional and ; 1, ; 2 are inner products on V with corresponding norms 1 and 2. Prove that there exists a positive number c such that

k. k k. k

h. .i h. .i

k*v*k1 三 ck*v*k2

for every *v* 2 V.

1. Suppose *v*1;:::; *v*m is a linearly independent list in V. Show that there exists *w* 2 V such that h*w*; *v*j i >0 for all j 2 f1; :::; mg.
2. Suppose e1;:::; en is an orthonormal basis of V and *v*1;:::; *v*n are vectors in V such that

1

kej — *v*j k < pn

for each j . Prove that *v*1;:::; *v*n is a basis of V.

1. Suppose C**R**.Œ—1; 1]/ is the vector space of continuous real-valued func- tions on the interval Œ—1; 1] with inner product given by

Z 1

hf; giD

1

f .x/g.x/ dx

for f; g 2 C**R**.Œ—1; 1]/. Let ' be the linear functional on C**R**.Œ—1; 1]/

deﬁned by '.f / f .0/. Show that there does not exist g C**R**.Œ 1; 1]/

D 2 —

such that

'.f / D hf; gi

for every f 2 C**R**.Œ—1; 1]/.

[*The exercise above shows that the Riesz Representation Theorem* (*6.42*)

*does not hold on inﬁnite-dimensional vector spaces without additional hypotheses on* V *and* '*.*]

**192** CHAPTER 6 Inner Product Spaces

1. Suppose **F C**, V is ﬁnite-dimensional, T .V /, all the eigenvalues of T have absolute value less than 1, and E> 0. Prove that there exists a positive integer m such that kT m*v*k三 Ek*v*k for every *v* 2 V.

D 2 *L*

1. For u 2 V, let ˆu denote the linear functional on V deﬁned by

.ˆu/.*v*/ D h*v*; ui

for *v* 2 V.

* 1. Show that if **F R**, then ˆ is a linear map from V to V 0. (Recall from Section 3.F that V 0 .V; **F**/ and that V 0 is called the dual space of V.)

D *L*

D

* 1. Show that if **F** D **C** and V ¤ f0g, then ˆ is not a linear map.
  2. Show that ˆ is injective.
  3. Suppose **F R** and V is ﬁnite-dimensional. Use parts (a) and (c) and a dimension-counting argument (but without using 6.42) to show that ˆ is an isomorphism from V onto V 0.

D

[*Part* (*d*) *gives an alternative proof of the Riesz Representation Theorem* (*6.42*) *when* **F R***. Part* (*d*) *also gives a natural isomorphism* (*meaning that it does not depend on a choice of basis*) *from a ﬁnite-dimensional real inner product space onto its dual space.*]

D

SECTION 6.C Orthogonal Complements and Minimization Problems **193**

## *Orthogonal Complements and*Minimization Problems

6.C

### Orthogonal Complements

6.45 **Deﬁnition *orthogonal complement,*** U ?

If U is a subset of V, then the ***orthogonal complement*** of U, denoted U ?, is the set of all vectors in V that are orthogonal to every vector in U :

U ? D f*v* 2 V W h*v*; uiD 0 for every u 2 U g:

For example, if U is a line in **R**3, then U ? is the plane containing the origin that is perpendicular to U. If U is a plane in **R**3, then U ? is the line containing the origin that is perpendicular to U.

6.46 Basic properties of orthogonal complement

1. If U is a subset of V, then U ? is a subspace of V.
2. f0g? D V.
3. V ? D f0g.
4. If U is a subset of V, then U \ U ? c f0g.
5. If U and W are subsets of V and U c W, then W ? c U ?.

Proof

1. Suppose U is a subset of V. Then h0; uiD 0 for every u 2 U ; thus

0 2 U ?.

Suppose *v*; *w* 2 U ?. If u 2 U, then

h*v* C *w*; uiD h*v*; uiC h*w*; uiD 0 C 0 D 0:

Thus *v* C *w* 2 U ?. In other words, U ? is closed under addition.

Similarly, suppose 入 2 **F** and *v* 2 U ?. If u 2 U, then

h入*v*; uiD 入h*v*; uiD 入. 0 D 0:

Thus 入*v* U ?. In other words, U ? is closed under scalar multiplica- tion. Thus U ? is a subspace of V.

2

**194** CHAPTER 6 Inner Product Spaces

1. Suppose *v* 2 V. Then h*v*; 0iD 0, which implies that *v* 2 f0g?. Thus

f0g? D V.

1. Suppose *v* 2 V ?. Then h*v*; *v*iD 0, which implies that *v* D 0. Thus

V ? D f0g.

1. Suppose U is a subset of V and *v* 2 U \ U ?. Then h*v*; *v*iD 0, which implies that *v* D 0. Thus U \ U ? c f0g.
2. Suppose U and W are subsets of V and U c W. Suppose *v* 2 W ?. Then h*v*; uiD 0 for every u 2 W, which implies that h*v*; uiD 0 for every u 2 U. Hence *v* 2 U ?. Thus W ? c U ?.

Recall that if U; W are subspaces of V, then V is the direct sum of U and W (written V U W ) if each element of V can be written in exactly one way as a vector in U plus a vector in W (see 1.40).

D ˚

The next result shows that every ﬁnite-dimensional subspace of V leads to a natural direct sum decomposition of V.

6.47 Direct sum of a subspace and its orthogonal complement

Suppose U is a ﬁnite-dimensional subspace of V. Then

V D U ˚ U ?:

Proof First we will show that

* 1. V D U C U ?:

To do this, suppose *v* V. Let e1;:::; em be an orthonormal basis of U. Obviously

2

* 1. *v* D h„*v*; e1ie1 C .ƒ. ‚. C h*v*; emiem… C „*v* — h*v*; e1ie1 —ƒ‚. . .— h*v*; emiem… :

u

*w*

Let u and *w* be deﬁned as in the equation above. Clearly u 2 U. Because

e1;:::; em is an orthonormal list, for each j D 1; :::;m we have

h*w*; ej iD h*v*; ej i— h*v*; ej i D 0:

Thus *w* is orthogonal to every vector in span.e1;:::; em/. In other words, *w* U ?. Thus we have written *v* u *w*, where u U and *w* U ?, completing the proof of 6.48.

2 D C 2 2

From 6.46(d), we know that U \ U ? D f0g. Along with 6.48, this implies that V D U ˚ U ? (see 1.45).

SECTION 6.C Orthogonal Complements and Minimization Problems **195**

Now we can see how to compute dim U ? from dim U.

6.50 Dimension of the orthogonal complement

Suppose V is ﬁnite-dimensional and U is a subspace of V. Then

dim U ? D dim V — dim U:

Proof The formula for dim U ? follows immediately from 6.47 and 3.78.

The next result is an important consequence of 6.47.

6.51 The orthogonal complement of the orthogonal complement

Suppose U is a ﬁnite-dimensional subspace of V. Then

U D .U ?/?:

Proof First we will show that

**6.52** U c .U ?/?:

To do this, suppose u U. Then u; *v* 0 for every *v* U ? (by the deﬁnition of U ?). Because u is orthogonal to every vector in U ?, we have

2 h i D 2

u 2 .U ?/?, completing the proof of 6.52.

To prove the inclusion in the other direction, suppose *v* 2 .U

?/?

. By

6.47, we can write *v* u *w*, where u U and *w* U ?. We have *v* u *w* U ?. Because *v* .U ?/? and u .U ?/? (from 6.52), we have *v* u .U ?/?. Thus *v* u U ? .U ?/?, which implies that *v* u is orthogonal to itself, which implies that *v* u 0, which implies that *v* u, which implies that *v* U. Thus .U ?/? U, which along with 6.52

D 2 c

— D

— 2 — 2 \ —

— D 2 2 2

D C 2 2

completes the proof.

We now deﬁne an operator *P*U for each ﬁnite-dimensional subspace of V.

6.53 **Deﬁnition *orthogonal projection,*** PU

Suppose U is a ﬁnite-dimensional subspace of V. The ***orthogonal***

For *v* 2 V, write *v* D u C *w*, where u 2 U and *w* 2 U ?. Then PU *v* D u.

***projection*** of V onto U is the operator PU 2 *L*.V / deﬁned as follows:

**196** CHAPTER 6 Inner Product Spaces

The direct sum decomposition V D U ˚ U ? given by 6.47 shows that

each *v* ?2 V can be uniquely written in the form *v* D u C *w* with u 2 U and

*w* 2 U . Thus PU *v* is well deﬁned.

6.54 **Example** Suppose x 2 V with x ¤ 0 and U D span.x/. Show that

for every *v* 2 V.

D

U

P *v* h*v*; xix kxk2

Solution Suppose *v* 2 V. Then

*v* D h*v*; xix C *v* — h*v*; xix ;

kxk2

kxk2

where the ﬁrst term on the right is in span.x/ (and thus in U ) and the second term on the right is orthogonal to x (and thus is in U ?/. Thus PU *v* equals the ﬁrst term on the right, as desired.

6.55 Properties of the orthogonal projection PU

Suppose U is a ﬁnite-dimensional subspace of V and *v* 2 V. Then

1. PU 2 *L*.V / ;
2. PU u D u for every u 2 U ;
3. PU *w* D 0 for every *w* 2 U ?;
4. range PU D U ;
5. null PU D U ?;
6. *v* — PU *v* 2 U ?;
7. PU 2 D PU ;
8. kPU *v*k三 k*v*k;
9. for every orthonormal basis e1;:::; em of U,

PU *v* D h*v*; e1ie1 C. . .C h*v*; emiem:

SECTION 6.C Orthogonal Complements and Minimization Problems **197**

Proof

1. To show that PU is a linear map on V, suppose *v*1; *v*2 2 V. Write

*v*1 D u1 C *w*1 and *v*2 D u2 C *w*2

with u1; u2 2 U and *w*1; *w*2 2 U ?. Thus PU *v*1 D u1 and PU *v*2 D u2.

Now

*v*1 C *v*2 D .u1 C u2/ C .*w*1 C *w*2/;

where u1 C u2 2 U and *w*1 C *w*2 2 U ?. Thus

PU .*v*1 C *v*2/ D u1 C u2 D PU *v*1 C PU *v*2:

Similarly, suppose 入 2 **F**. The equation *v* D u C *w* with u 2 U and *w* 2 U ? implies that 入*v* D 入u C 入*w* with 入u 2 U and 入*w* 2 U ?. Thus PU .入*v*/ D 入u D 入PU *v*.

Hence PU is a linear map from V to V.

1. Suppose u 2 U. We can write u D u C 0, where u 2 U and 0 2 U ?. Thus PU u D u.
2. Suppose *w* 2 U ?. We can write *w* D 0 C*w*, where 0 2 U and *w* 2 U ?. Thus PU *w* D 0.
3. The deﬁnition of PU implies that range PU c U. Part (b) implies that

U c range PU. Thus range PU D U.

1. Part (c) implies that U ? c null PU. To prove the inclusion in the other direction, note that if *v* 2 null PU then the decomposition given by 6.47 must be *v* D 0 C *v*, where 0 2 U and *v* 2 U ?. Thus null PU c U ?.
2. If *v* D u C *w* with u 2 U and *w* 2 U ?, then

*v* — PU *v* D *v* — u D *w* 2 U ?:

1. If *v* D u C *w* with u 2 U and *w* 2 U ?, then

.PU 2/*v* D PU .PU *v*/ D PU u D u D PU *v*:

1. If *v* D u C *w* with u 2 U and *w* 2 U ?, then

kPU *v*k2 D kuk 三 kuk C k*w*k2 D k*v*k ;

2

2

2

where the last equality comes from the Pythagorean Theorem.

1. The formula for PU *v* follows from equation 6.49 in the proof of 6.47.

**198** CHAPTER 6 Inner Product Spaces

### Minimization Problems

The following problem often arises: given a subspace U of V and a point *v* V, ﬁnd a point u U such that *v* u is as small as possible. The

*The remarkable simplicity of the so- lution to this minimization problem has led to many important applica- tions of inner product spaces out- side of pure mathematics.*

k — k

2 2

next proposition shows that this mini- mization problem is solved by taking

u D PU *v*.

6.56 Minimizing the distance to a subspace

Suppose U is a ﬁnite-dimensional subspace of V, *v* 2 V, and u 2 U. Then

k*v* — PU *v*k三 k*v* — uk:

Furthermore, the inequality above is an equality if and only if u D PU *v*.

Proof We have

#### 6.57

k*v* — PU *v*k 三 k*v* — PU *v*k C kPU *v* — uk

2

2 2 2

D k.*v* — PU *v*/ C .PU *v* — u/k

2

D k*v* — uk ;

where the ﬁrst line above holds because 0 PU *v* u 2, the second line above comes from the Pythagorean Theorem [which applies because *v* PU *v* U ? by 6.55(f), and PU *v* u U ], and the third line above holds by simple algebra. Taking square roots gives the desired inequality.

三 k — k

— 2 — 2

Our inequality above is an equality if and only if 6.57 is an equality, which happens if and only if kPU *v* — ukD 0, which happens if and only if u D PU *v*.

*v*

*U*

*PUv*

0

PU *v is the closest point in* U *to v.*

SECTION 6.C Orthogonal Complements and Minimization Problems **199**

The last result is often combined with the formula 6.55(i) to compute explicit solutions to minimization problems.

* 1. **Example** Find a polynomial u with real coefﬁcients and degree at most 5 that approximates sin x as well as possible on the interval Œ 兀; 兀], in

the sense that

—

Z

冗

— j sin x — u.x/j dx

2

冗

is as small as possible. Compare this result to the Taylor series approximation.

Solution Let C**R**Œ—兀; 兀] denote the real inner product space of continuous real-valued functions on Œ—兀; 兀] with inner product

* 1. Z 冗

冗

hf; giD

f .x/g.x/ dx:

Let *v* 2 C**R**Œ—兀; 兀] be the function deﬁned by *v*.x/ D sin x. Let U denote the subspace of C**R**Œ 兀; 兀] consisting of the polynomials with real coefﬁcients and degree at most 5. Our problem can now be reformulated as follows:

Find u 2 U such that k*v* — uk is as small as possible.

—

To compute the solution to our ap- proximation problem, ﬁrst apply the Gram–Schmidt Procedure (using the in-

*A computer that can perform inte- grations is useful here.*

ner product given by 6.59) to the basis 1; x; x2; x3; x4; x5 of U, producing an orthonormal basis e1; e2; e3; e4; e5; e6 of U. Then, again using the inner product given by 6.59, compute PU *v* using 6.55(i) (with m 6). Doing this computation shows that PU *v* is the function u deﬁned by

D

**6.60** u.x/ D 0:987862x — 0:155271x3 C 0:00564312x5;

where the 兀’s that appear in the exact answer have been replaced with a good decimal approximation.

By 6.56, the polynomial u above is the best approximation to sin x on

Œ—兀; 兀] using polynomials of degRree at most 5 (here “best approximation”

means in the sense of minimizing

冗

冗

j sin x — u.x/j2 dx). To see how good

this approximation is, the next ﬁgure shows the graphs of both sin x and our approximation u.x/ given by 6.60 over the interval Œ—兀; 兀].

**200** CHAPTER 6 Inner Product Spaces

1

一3

3

一1

*Graphs on* Œ 兀 ; 兀 ] *of* sin x (*blue*) *and its approximation* u.x/ (*red*) *given by 6.60.*

Our approximation 6.60 is so accurate that the two graphs are almost identical—our eyes may see only one graph! Here the blue graph is placed

—

almost exactly over the red graph. If you are viewing this on an electronic device, try enlarging the picture above, especially near 3 or 3, to see a small gap between the two graphs.

—

Another well-known approximation to sin x by a polynomial of degree 5

is given by the Taylor polynomial

#### 6.61

x3 x5 x — 3Š C 5Š :

To see how good this approximation is, the next picture shows the graphs of both sin x and the Taylor polynomial 6.61 over the interval Œ—兀; 兀].

1

一3

3

一1

*Graphs on* Œ—兀; 兀] *of* sin x (*blue*) *and the Taylor polynomial 6.61* (*red*)*.*

The Taylor polynomial is an excellent approximation to sin x for x near 0. But the picture above shows that for x > 2, the Taylor polynomial is not so accurate, especially compared to 6.60. For example, taking x 3, our approximation 6.60 estimates sin 3 with an error of about 0:001, but the Taylor series 6.61 estimates sin 3 with an error of about 0:4. Thus at x 3, the error

D

D

j j

in the Taylor series is hundreds of times larger than the error given by 6.60. Linear algebra has helped us discover an approximation to sin x that improves upon what we learned in calculus!

SECTION 6.C Orthogonal Complements and Minimization Problems **201**

EXERCISES 6.C

1. Suppose *v*1;:::; *v*m 2 V. Prove that

f*v*1;:::; *v*mg? D span.*v*1;:::; *v*m/ ?:

（ ）

1. Suppose U is a ﬁnite-dimensional subspace of V. Prove that U ? 0

D f g

if and only if U V.

D

[*Exercise 14*(*a*) *shows that the result above is not true without the hy- pothesis that* U *is ﬁnite-dimensional.*]

1. Suppose U is a subspace of V with basis u1;:::; um and

u1;:::; um; *w*1;:::; *w*n

is a basis of V. Prove that if the Gram–Schmidt Procedure is applied to the basis of V above, producing a list e1;:::; em; f1;:::; fn, then e1;:::; em is an orthonormal basis of U and f1;:::; fn is an orthonor- mal basis of U ?.

1. Suppose U is the subspace of **R**4 deﬁned by

（ ）

U D span .1; 2; 3; —4/; .—5; 4; 3; 2/ :

Find an orthonormal basis of U and an orthonormal basis of U ?.

1. Suppose V is ﬁnite-dimensional and U is a subspace of V. Show that

PU ? D I — PU, where I is the identity operator on V.

1. Suppose U and W are ﬁnite-dimensional subspaces of V. Prove that

PU PW D 0 if and only if hu; *w*iD 0 for all u 2 U and all *w* 2 W.

1. Suppose V is ﬁnite-dimensional and P .V / is such that P 2 P and every vector in null P is orthogonal to every vector in range P . Prove that there exists a subspace U of V such that P D PU.

2 *L* D

1. Suppose V is ﬁnite-dimensional and P .V / is such that P 2 P

2 *L* D

and

kP *v*k三 k*v*k

for every *v* 2 V. Prove that there exists a subspace U of V such that

P D PU.

1. Suppose T 2 *L*.V / and U is a ﬁnite-dimensional subspace of V. Prove that U is invariant under T if and only if PU TPU D TPU.

**202** CHAPTER 6 Inner Product Spaces

1. Suppose V is ﬁnite-dimensional, T .V /, and U is a subspace of V. Prove that U and U ? are both invariant under T if and only if PU T D TPU.

2 *L*

1. In **R**4, let

（ ）

U D span .1; 1; 0; 0/; .1; 1; 1; 2/ :

Find u 2 U such that ku — .1; 2; 3; 4/k is as small as possible.

1. Find p 2 *P*3.**R**/ such that p.0/ D 0, p0.0/ D 0, and

Z

1

j2 C 3x — p.x/j dx

0

2

is as small as possible.

1. Find p 2 *P*5.**R**/ that makes

Z

冗

冗

j sin x — p.x/j2 dx

as small as possible.

[*The polynomial 6.60 is an excellent approximation to the answer to this*

*exercise, but here you are asked to ﬁnd the exact solution, which involves powers of* 兀*. A computer that can perform symbolic integration will be useful.*]

1. Suppose C**R**.Œ—1; 1]/ is the vector space of continuous real-valued func- tions on the interval Œ—1; 1] with inner product given by

Z 1

hf; giD

1

f .x/g.x/ dx

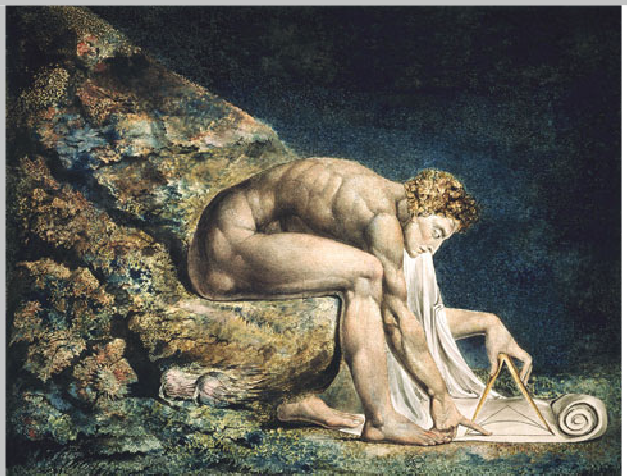
for f; g 2 C**R**.Œ—1; 1]/. Let U be the subspace of C**R**.Œ—1; 1]/ deﬁned

by

U D ff 2 C**R**.Œ—1; 1]/ W f .0/ D 0g:

* 1. Show that U ? D f0g.
  2. Show that 6.47 and 6.51 do not hold without the ﬁnite-dimensional hypothesis.

# *Operators on Inner Product* Spaces



*Isaac Newton*

(*1642–1727*)*, as*

*envisioned by British poet and artist William Blake in this 1795 painting.*

7

CHAPTER

The deepest results related to inner product spaces deal with the subject to which we now turn—operators on inner product spaces. By exploiting properties of the adjoint, we will develop a detailed description of several important classes of operators on inner product spaces.

A new assumption for this chapter is listed in the second bullet point below:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V and W denote ﬁnite-dimensional inner product spaces over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

adjoint

Spectral Theorem positive operators isometries

Polar Decomposition

Singular Value Decomposition

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

203

DOI 10.1007/978-3-319-11080-6 7

**204** CHAPTER 7 Operators on Inner Product Spaces

## *Self-Adjoint and Normal Operators*

7.A

### Adjoints

7.2 **Deﬁnition *adjoint,*** T \*

such that

Suppose T 2 *L*.V; W /. The ***adjoint*** of T is the function T \* W W ! V

hT *v*; *w*iD h*v*;T \**w*i

for every *v* 2 V and every *w* 2 W.

To see why the deﬁnition above makes sense, suppose T .V; W /. Fix *w* W. Consider the linear func- tional on V that maps *v* V to T *v*; *w* ; this linear functional depends on T and

*The word* ***adjoint*** *has another meaning in linear algebra. We do not need the second meaning in this book. In case you encounter the second meaning for adjoint elsewhere, be warned that the two meanings for adjoint are unrelated to each other.*

2 h i

2

2 *L*

*w*. By the Riesz Representation Theo-

rem (6.42), there exists a unique vector in V such that this linear functional is

given by taking the inner product with it. We call this unique vector T \**w*. In

other words, T \**w* is the unique vector in V such that hT *v*; *w*iD h*v*;T \**w*i for every *v* 2 V.

* 1. **Example** Deﬁne T W **R**3 ! **R**2 by

T .x1; x2; x3/ D .x2 C 3x3; 2x1/:

Find a formula for T \*.

Solution Here T \* will be a function from **R**2 to **R**3. To compute T \*, ﬁx a point .y1; y2/ 2 **R**2. Then for every .x1; x2; x3/ 2 **R**3 we have

h.x1; x2; x3/; T \*.y1; y2/iD hT .x1; x2; x3/; .y1; y2/i

D h.x2 C 3x3; 2x1/; .y1; y2/i D x2y1 C 3x3y1 C 2x1y2

D h.x1; x2; x3/; .2y2; y1; 3y1/i:

Thus

T \*.y1; y2/ D .2y2; y1; 3y1/:

SECTION 7.A Self-Adjoint and Normal Operators **205**

* 1. **Example** Fix u 2 V and x 2 W. Deﬁne T 2 *L*.V; W / by

T *v* D h*v*; uix

for every *v* 2 V. Find a formula for T \*.

Solution Fix *w* 2 W. Then for every *v* 2 V we have

h*v*;T \**w*iD hT *v*; *w*i

˝ ˛

D h*v*; uix; *w* D h*v*; uihx; *w*i D *v*; h*w*; xiu :

˝ ˛

Thus

T \**w*D h*w*; xiu:

In the two examples above, T \* turned out to be not just a function but a linear map. This is true in general, as shown by the next result.

The proofs of the next two results use a common technique: ﬂip T \* from one side of an inner product to become T on the other side.

7.5 The adjoint is a linear map

If T 2 *L*.V; W /, then T \* 2 *L*.W; V /.

Proof Suppose T 2 *L*.V; W /. Fix *w*1; *w*2 2 W. If *v* 2 V, then

h*v*;T \*.*w*1 C *w*2/iD hT *v*; *w*1 C *w*2i

D hT *v*; *w*1iC hT *v*; *w*2i D h*v*;T \**w*1iC h*v*;T \**w*2i D h*v*;T \**w*1 C T \**w*2i;

which shows that T \*.*w*1 C *w*2/ D T \**w*1 C T \**w*2.

Fix *w* 2 W and 入 2 **F**. If *v* 2 V, then

h*v*;T \*.入*w*/iD hT *v*; 入*w*i

D 入N hT *v*; *w*i D 入N h*v*;T \**w*i D h*v*; 入T \**w*i;

which shows that T \*.入*w*/ 入T \**w*.

D

Thus T \* is a linear map, as desired.

**206** CHAPTER 7 Operators on Inner Product Spaces

7.6 Properties of the adjoint

1. .S C T /\* D S\* C T \* for all S; T 2 *L*.V; W /;
2. .入T /\* D 入N T \* for all 入 2 **F** and T 2 *L*.V; W /;
3. .T \*/\* D T for all T 2 *L*.V; W /;
4. I\* D I, where I is the identity operator on V ; (e)

is an inner product space over **F**).

.ST /\* D T \*S\* for all T 2 *L*.V; W / and S 2 *L*.W; U / (here U

Proof

1. Suppose S; T 2 *L*.V; W /. If *v* 2 V and *w* 2 W, then

h*v*; .S C T /\**w*iD h.S C T /*v*; *w*i

D hS *v*; *w*iC hT *v*; *w*i D h*v*;S\**w*iC h*v*;T \**w*i D h*v*;S\**w* C T \**w*i:

Thus .S C T /\**w* D S\**w* C T \**w*, as desired.

1. Suppose 入 2 **F** and T 2 *L*.V; W /. If *v* 2 V and *w* 2 W, then

h*v*; .入T /\**w*iD h入T *v*; *w*iD 入hT *v*; *w*iD 入h*v*;T \**w*iD h*v*; 入N T \**w*i:

Thus .入T /\**w* D 入N T \**w*, as desired.

1. Suppose T 2 *L*.V; W /. If *v* 2 V and *w* 2 W, then

h*w*; .T \*/\**v*iD hT \**w*; *v*iD h*v*;T \**w*iD hT *v*; *w*iD h*w*;T *v*i:

Thus .T \*/\**v* D T *v*, as desired.

1. If *v*;u 2 V, then

h*v*;I\*uiD hI *v*; uiD h*v*; ui:

Thus I\*u D u, as desired.

1. Suppose T 2 *L*.V; W / and S 2 *L*.W; U /. If *v* 2 V and u 2 U, then

h*v*; .ST /\*uiD hST *v*; uiD hT *v*;S\*uiD h*v*;T \*.S\*u/i:

Thus .ST /\*u D T \*.S\*u/, as desired.

SECTION 7.A Self-Adjoint and Normal Operators **207**

The next result shows the relationship between the null space and the range of a linear map and its adjoint. The symbol used in the proof means “if and only if”; this symbol could also be read to mean “is equivalent to”.

()

7.7 Null space and range of T \*

Suppose T 2 *L*.V; W /. Then

1. null T \* D .range T /?;
2. range T \* D .null T /?;
3. null T D .range T \*/?;
4. range T D .null T \*/?.

Proof We begin by proving (a). Let *w* 2 W. Then

*w* 2 null T \* () T \**w* D 0

() h*v*;T \**w*iD 0 for all *v* 2 V () hT *v*; *w*iD 0 for all *v* 2 V () *w* 2 .range T /?:

Thus null T \* .range T /?, proving (a).

D

If we take the orthogonal complement of both sides of (a), we get (d), where we have used 6.51. Replacing T with T \* in (a) gives (c), where we have used 7.6(c). Finally, replacing T with T \* in (d) gives (b).

7.8 **Deﬁnition *conjugate transpose***

The ***conjugate transpose*** of an m-by-n matrix is the n-by-m matrix ob- tained by interchanging the rows and columns and then taking the complex conjugate of each entry.

7.9 **Example**

The conjugate transpose of the matrix

*If* **F** D **R***, then the conjugate trans-*

*pose of a matrix is the same as its* ***transpose****, which is the matrix ob- tained by interchanging the rows and columns.*

is the matrix

0 2 6 1

3 — 4i 5 :

@

2 3 C 4i 7

6 5 8i

A

7 —8i

**208** CHAPTER 7 Operators on Inner Product Spaces

The next result shows how to com- pute the matrix of T \* from the matrix of T.

*The adjoint of a linear map does not depend on a choice of basis. This explains why this book em- phasizes adjoints of linear maps instead of conjugate transposes of matrices.*

Caution: Remember that the result below applies only when we are dealing

with orthonormal bases. With respect to nonorthonormal bases, the matrix of T \* does not necessarily equal the conjugate transpose of the matrix of T.

7.10 The matrix of T \*

Let T 2 *L*.V; W /. Suppose e1;:::; en is an orthonormal basis of V and

f1;:::; fm is an orthonormal basis of W. Then

is the conjugate transpose of

*M*（T \*; .f1;:::; fm/; .e1;:::; en/）

*M*（T; .e1;:::; en/; .f1;:::; fm/）:

\*

Proof I（n this proof, we will write ）*M*.T / instead of the longer expres-

sion *M*（ T; .e1;:::; en/; .f1;:::; fm/） ; we will also write *M*.T / instead

*M*

*M*

of T \*; .f1;:::; fm/; .e1;:::; en/ .

Recall that we obtain the kth column of .T / by writing T ek as a linear

combination of the fj ’s; the scalars used in this linear combination then

become the kth column of .T /. Because f1;:::; fm is an orthonormal basis of W, we know how to write Tek as a linear combination of the fj ’s (see 6.30):

*M*

T ek D hT ek; f1if1 C. . .C hT ek; fmifm:

*M* h i

Thus the entry in row j , column k, of .T / is Tek; fj .

Replacing T with T \* and interchanging the roles played by the e’s and f ’s, we see that the entry in row j , column k, of *M*.T \*/ is hT \*fk; ej i, which equals hfk;Tej i, which equals hTej ; fki, which equals the complex conjugate of the entry in row k, column j, of *M*.T /. In other words, *M*.T \*/ is the conjugate transpose of *M*.T /.

SECTION 7.A Self-Adjoint and Normal Operators **209**

### Self-Adjoint Operators

Now we switch our attention to operators on inner product spaces. Thus instead of considering linear maps from V to W, we will be focusing on linear maps from V to V ; recall that such linear maps are called operators.

7.11 **Deﬁnition *self-adjoint***

An operator T 2 *L*.V / is called ***self-adjoint*** if T D T \*. In other words,

T 2 *L*.V / is self-adjoint if and only if

hT *v*; *w*iD h*v*;T *w*i

for all *v*; *w*2 V.

7.12 **Example** Suppose T is the operator on **F**2 whose matrix (with re- spect to the standard basis) is

2 b :

3 7

Find all numbers b such that T is self-adjoint.

Solution The operator T is self-adjoint if and only if b D 3 (because

*M*.T / D *M*.T \*/ if and only if b D 3; recall that *M*.T \*/ is the conjugate transpose of *M*.T /—see 7.10).

You should verify that the sum of two self-adjoint operators is self-adjoint and that the product of a real scalar and a self-adjoint operator is self-adjoint.

A good analogy to keep in mind (es-

*Some mathematicians use the term* ***Hermitian*** *instead of self-adjoint, honoring French mathematician Charles Hermite, who in 1873 pub-*

*lished the ﬁrst proof that* e *is not a*

*zero of any polynomial with integer*

*coefﬁcients.*

pecially when **F C**) is that the adjoint on .V / plays a role similar to complex conjugation on **C**. A complex number z is real if and only if z z; thus a self-

*L*

D

D

D N

adjoint operator (T T \*) is analogous

to a real number.

We will see that the analogy discussed above is reﬂected in some important properties of self-adjoint operators, beginning with eigenvalues in the next result.

If **F** D **R**, then by deﬁnition every eigenvalue is real, so the next result is interesting only when **F** D **C**.

**210** CHAPTER 7 Operators on Inner Product Spaces

7.13 Eigenvalues of self-adjoint operators are real

Every eigenvalue of a self-adjoint operator is real.

Proof Suppose T is a self-adjoint operator on V. Let 入be an eigenvalue of

T, and let *v* be a nonzero vector in V such that T *v* D 入*v*. Then

入k*v*k2 D h入*v*; *v*iD hT *v*; *v*iD h*v*;T *v*iD h*v*; 入*v*iD 入N k*v*k2:

Thus 入 D 入N , which means that 入is real, as desired.

The next result is false for real inner product spaces. As an example, consider the operator T 2 *L*.**R**2/ that is a counterclockwise rotation of 90ı around the origin; thus T .x; y/ D .—y; x/. Obviously T *v* is orthogonal to *v* for every *v* 2 **R**2, even though T ¤ 0.

7.14 Over **C**, T *v* is orthogonal to *v* for all *v* only for the 0 operator

Suppose V is a complex inner product space and T 2 *L*.V /. Suppose

hT *v*; *v*iD 0

for all *v* 2 V. Then T D 0.

Proof We have

hT u; *w*iD hT .u C *w*/; u C *w*i— hT .u — *w*/; u — *w*i

4

hT .u C i*w*/; u C i*w*i— hT .u — i*w*/; u — i*w*i i 4

C

for all u; *w* 2 V, as can be veriﬁed by computing the right side. Note that each term on the right side is of the form hT *v*; *v*i for appropriate *v* 2 V. Thus our hypothesis implies that hT u; *w*iD 0 for all u; *w* 2 V. This implies that T D 0 (take *w* D T u).

The next result is false for real inner product spaces, as shown by consider- ing any operator on a real inner product space that is not self-adjoint.

*The next result provides another ex- ample of how self-adjoint opera- tors behave like real numbers.*

SECTION 7.A Self-Adjoint and Normal Operators **211**

7.15 Over **C**, hT *v*; *v*i is real for all *v* only for self-adjoint operators

Suppose V is a complex inner product space and T 2 *L*.V /. Then T is

self-adjoint if and only if

hT *v*; *v*i2 **R**

for every *v* 2 V.

Proof Let *v* 2 V. Then

hT *v*; *v*i—hT *v*; *v*iD hT *v*; *v*i—h*v*;T *v*iD hT *v*; *v*i—hT \**v*; *v*iD h.T —T \*/*v*; *v*i:

If T *v*; *v* **R** for every *v* V, then the left side of the equation above equals 0, so .T T \*/*v*; *v* 0 for every *v* V. This implies that T T \* 0 (by 7.14). Hence T is self-adjoint.

h — iD 2 — D

h i2 2

Conversely, if T is self-adjoint, then the right side of the equation above equals 0, so hT *v*; *v*iD hT *v*; *v*i for every *v* 2 V. This implies that hT *v*; *v*i2 **R** for every *v* 2 V, as desired.

On a real inner product space V, a nonzero operator T might satisfy T *v*; *v* 0 for all *v* V. However, the next result shows that this cannot happen for a self-adjoint operator.

h iD 2

7.16 If T D T \* and hT *v*; *v*iD 0 for all *v*, then T D 0

Suppose T is a self-adjoint operator on V such that

hT *v*; *v*iD 0

for all *v* 2 V. Then T D 0.

Proof We have already proved this (without the hypothesis that T is self- adjoint) when V is a complex inner product space (see 7.14). Thus we can assume that V is a real inner product space. If u; *w* 2 V, then

**7.17** hT u; *w*iD hT .u C *w*/; u C *w*i— hT .u — *w*/; u — *w*i I

4

this is proved by computing the right side using the equation

hT *w*; uiD h*w*;T uiD hT u; *w*i;

where the ﬁrst equality holds because T is self-adjoint and the second equality holds because we are working in a real inner product space.

Each term on the right side of 7.17 is of the form hT *v*; *v*i for appropriate *v*.

Thus hT u; *w*iD 0 for all u; *w* 2 V. This implies that T D 0 (take *w* D T u).

**212** CHAPTER 7 Operators on Inner Product Spaces

### Normal Operators

7.18 **Deﬁnition *normal***

mutes with its adjoint.

* In other words, T 2 *L*.V / is normal if

TT \* D T \*T:

* An operator on an inner product space is called ***normal*** if it com-

Obviously every self-adjoint operator is normal, because if T is self-adjoint then T \* D T.

7.19 **Example** Let T be the operator on **F**2 whose matrix (with respect to the standard basis) is

2 —3 :

3 2

Show that T is not self-adjoint and that T is normal.

Solution This operator is not self-adjoint because the entry in row 2, column 1 (which equals 3) does not equal the complex conjugate of the entry in row 1, column 2 (which equals 3).

—

The matrix of TT \* equals

2 —3 2 3 ; which equals 13 0 :

3 2 —3 2

0 13

Similarly, the matrix of T \*T equals

2 3 2 —3 ; which equals 13 0 :

—3 2 3 2

0 13

Because TT \* and T \*T have the same matrix, we see that TT \* T \*T. Thus T is normal.

D

In the next section we will see why normal operators are worthy of special

attention.

The next result provides a simple characterization of normal operators.

*The next result implies that*

null T D null T \* *for every normal*

*operator* T*.*

SECTION 7.A Self-Adjoint and Normal Operators **213**

7.20 T is normal if and only if kT *v*kD kT \**v*k for all *v*

An operator T 2 *L*.V / is normal if and only if

kT *v*kD kT \**v*k

for all *v* 2 V.

Proof Let T .V /. We will prove both directions of this result at the same time. Note that

T is normal () T \*T — TT \* D 0

2 *L*

() h.T \*T — TT \*/*v*; *v*iD 0 for all *v* 2 V () hT \*T *v*; *v*iD hTT \**v*; *v*i for all *v* 2 V () kT *v*k2 D kT \**v*k2 for all *v* 2 V;

where we used 7.16 to establish the second equivalence (note that the operator T \*T TT \* is self-adjoint). The equivalence of the ﬁrst and last conditions above gives the desired result.

—

Compare the next corollary to Exercise 2. That exercise states that the eigenvalues of the adjoint of each operator are equal (as a set) to the complex conjugates of the eigenvalues of the operator. The exercise says nothing about eigenvectors, because an operator and its adjoint may have different eigenvectors. However, the next corollary implies that a normal operator and its adjoint have the same eigenvectors.

7.21 For T normal, T and T \* have the same eigenvectors

Suppose T 2 *L*.V / is normal and *v* 2 V is an eigenvector of T with

eigenvalue 入. Then *v* is also an eigenvector of T \* with eigenvalue 入N .

Proof Because T is normal, so is T 入I, as you should verify. Using 7.20, we have

—

0 D k.T — 入I /*v*kD k.T — 入I /\**v*kD k.T \* — 入N I/*v*k:

Hence *v* is an eigenvector of T \* with eigenvalue 入N , as desired.

Because every self-adjoint operator is normal, the next result applies in particular to self-adjoint operators.

**214** CHAPTER 7 Operators on Inner Product Spaces

7.22 Orthogonal eigenvectors for normal operators

Suppose T 2 *L*.V / is normal. Then eigenvectors of T corresponding to

distinct eigenvalues are orthogonal.

Proof Suppose ˛; ˇ are distinct eigenvalues of T, with corresponding eigen- vectors u; *v*. Thus Tu ˛u and T *v* ˇ*v*. From 7.21 we have T \**v* ˇN*v*.

Thus

D D D

.˛ — ˇ/hu; *v*iD h˛u; *v*i— hu; ˇN*v*i

D hT u; *v*i— hu; T \**v*i D 0:

Because ˛ ˇ, the equation above implies that u; *v* 0. Thus u and *v* are orthogonal, as desired.

¤ h iD

EXERCISES 7.A

1. Suppose n is a positive integer. Deﬁne T 2 *L*.**F**n/ by

T .z1;:::; zn/ D .0; z1;:::; zn 1/:

Find a formula for T \*.z1;:::; zn/.

1. Suppose T 2 *L*.V / and 入 2 **F**. Prove that 入is an eigenvalue of T if and only if 入N is an eigenvalue of T \*.
2. Suppose T .V / and U is a subspace of V. Prove that U is invariant under T if and only if U ? is invariant under T \*.

2 *L*

1. Suppose T 2 *L*.V; W /. Prove that
   1. T is injective if and only if T \* is surjective;
   2. T is surjective if and only if T \* is injective.
2. Prove that

and

dim null T \* D dim null T C dim W — dim V

dim range T \* D dim range T

for every T 2 *L*.V; W /.

SECTION 7.A Self-Adjoint and Normal Operators **215**

1. Make *P*2.**R**/ into an inner product space by deﬁning

Z 1

hp; qiD

p.x/q.x/ dx:

Deﬁne T 2 *L P*2.**R**/ by T .a0 C a1x C a2x2/ D a1x.

（ ）

0

* 1. Show that T is not self-adjoint.
  2. The matrix of T with respect to the basis .1; x; x2/ is

0@ 0 0 0 1A

0 1 0

0 0 0

:

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

1. Suppose S; T 2 *L*.V / are self-adjoint. Prove that ST is self-adjoint if and only if ST D TS.
2. Suppose V is a real inner product space. Show that the set of self-adjoint operators on V is a subspace of *L*.V /.
3. Suppose V is a complex inner product space with V ¤ f0g. Show that the set of self-adjoint operators on V is not a subspace of *L*.V /.
4. Suppose dim V 乏 2. Show that the set of normal operators on V is not a subspace of *L*.V /.
5. Suppose P 2 *L*.V / is such that P 2 D P . Prove that there is a subspace

U of V such that P D PU if and only if P is self-adjoint.

1. Suppose that T is a normal operator on V and that 3 and 4 are eigepnvalues

of T. Prove that there exists a vector *v* 2 V such that k*v*k D

kT *v*kD 5.

2 and

1. Give an example of an operator T .**C**4/ such that T is normal but not self-adjoint.

2 *L*

1. Suppose T is a normal operator on V. Suppose also that *v*; *w* V satisfy the equations

2

k*v*kD k*w*kD 2; T *v* D 3*v*; T *w* D 4*w*:

Show that kT .*v* C *w*/kD 10.

**216** CHAPTER 7 Operators on Inner Product Spaces

1. Fix u; x 2 V. Deﬁne T 2 *L*.V / by

T *v* D h*v*; uix

for every *v* 2 V.

* 1. Suppose **F R**. Prove that T is self-adjoint if and only if u; x is linearly dependent.

D

* 1. Prove that T is normal if and only if u; x is linearly dependent.

1. Suppose T 2 *L*.V / is normal. Prove that

range T D range T \*:

1. Suppose T 2 *L*.V / is normal. Prove that

null T k D null T and range T k D range T for every positive integer k.

1. Prove or give a counterexample: If T .V / and there exists an ortho- normal basis e1;:::; en of V such that Tej T \*ej for each j , then T is normal.

k k D k k

2 *L*

1. Suppose T 2 *L*.**C**3/ is normal and T .1; 1; 1/ D .2; 2; 2/. Suppose

.z1; z2; z3/ 2 null T. Prove that z1 C z2 C z3 D 0.

1. Suppose T .V; W / and **F R**. Let ˆV be the isomorphism from V onto the dual space V 0 given by Exercise 17 in Section 6.B, and let ˆW be the corresponding isomorphism from W onto W 0. Show that if ˆV and

2 *L* D

ˆW are used to identify V and W with V 0 and W 0, then T \* is identiﬁed with the dual map T 0. More precisely, show that ˆV ı T \* D T 0 ı ˆW.

1. Fix a positive integer n. In the inner product space of continuous real- valued functions on Œ 兀; 兀] with inner product

—

Z 冗

let

hf; giD

f .x/g.x/ dx;

冗

V D span.1; cos x; cos 2x;:::; cos nx; sin x; sin 2x;:::; sin nx/:

* 1. Deﬁne D .V / by Df f 0. Show that D\* D. Conclude that D is normal but not self-adjoint.

2 *L* D D —

* 1. Deﬁne T 2 *L*.V / by Tf D f 00. Show that T is self-adjoint.

SECTION 7.B The Spectral Theorem **217**

## *The Spectral Theorem*

7.B

Recall that a diagonal matrix is a square matrix that is 0 everywhere except possibly along the diagonal. Recall also that an operator on V has a diagonal matrix with respect to a basis if and only if the basis consists of eigenvectors

of the operator (see 5.41).

The nicest operators on V are those for which there is an *orthonormal* basis of V with respect to which the operator has a diagonal matrix. These are precisely the operators T .V / such that there is an orthonormal basis of V consisting of eigenvectors of T. Our goal in this section is to prove the

2 *L*

Spectral Theorem, which characterizes these operators as the normal operators when **F C** and as the self-adjoint operators when **F R**. The Spectral Theorem is probably the most useful tool in the study of operators on inner product spaces.

D D

Because the conclusion of the Spectral Theorem depends on **F**, we will break the Spectral Theorem into two pieces, called the Complex Spectral Theorem and the Real Spectral Theorem. As is often the case in linear algebra, complex vector spaces are easier to deal with than real vector spaces. Thus we present the Complex Spectral Theorem ﬁrst.

### The Complex Spectral Theorem

The key part of the Complex Spectral Theorem (7.24) states that if **F C** and T .V / is normal, then T has a diagonal matrix with respect to some orthonormal basis of V. The next example illustrates this conclusion.

2 *L*

D

2 *L*

7.23 **Example** Consider the normal operator T .**C**2/ from Example 7.19, whose matrix (with respect to the standard basis) is

2 —3 :

3 2

As you can verify, .pi;1/ ; . pi;1/ is an orthonormal basis of **C**2 consisting of

2 2

eigenvectors of T, and with respect to this basis the matrix of T is the diagonal

:

matrix

2 C 3i 0

0 2 — 3i

In the next result, the equivalence of (b) and (c) is easy (see 5.41). Thus we prove only that (c) implies (a) and that (a) implies (c).

**218** CHAPTER 7 Operators on Inner Product Spaces

7.24 Complex Spectral Theorem

Suppose **F** D **C** and T 2 *L*.V /. Then the following are equivalent:

1. T is normal.
2. V has an orthonormal basis consisting of eigenvectors of T.
3. T has a diagonal matrix with respect to some orthonormal basis of V.

Proof First suppose (c) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. The matrix of T \* (with respect to the same basis) is obtained by taking the conjugate transpose of the matrix of T ; hence T \* also has a diagonal matrix. Any two diagonal matrices commute; thus T commutes with T \*, which means that T is normal. In other words, (a) holds. Now suppose (a) holds, so T is normal. By Schur’s Theorem (6.38), there is an orthonormal basis e1;:::; en of V with respect to which T has an

upper-triangular matrix. Thus we can write

（ ） 0B a1;1 ::: a1;n 1C

**7.25** *M* T; .e1;:::; en/

D @

: : : :

0 an;n

A :

We will show that this matrix is actually a diagonal matrix.

We see from the matrix above that

2

2

and

kTe1k

D ja1;1j

kT \*e1k2 D ja1;1j2 C ja1;2j2 C. . .C ja1;nj2:

Because T is normal, Te1 T \*e1 (see 7.20). Thus the two equations above imply that all entries in the ﬁrst row of the matrix in 7.25, except possibly the ﬁrst entry a1;1, equal 0.

k kD k k

Now from 7.25 we see that

kTe2k D ja2;2j

2

2

(because a1;2 D 0, as we showed in the paragraph above) and

kT \*e2k2 D ja2;2j2 C ja2;3j2 C. . .C ja2;nj2:

Because T is normal, T e2 T \*e2 . Thus the two equations above imply that all entries in the second row of the matrix in 7.25, except possibly the diagonal entry a2;2, equal 0.

k kD k k

Continuing in this fashion, we see that all the nondiagonal entries in the matrix 7.25 equal 0. Thus (c) holds.

### The Real Spectral Theorem

SECTION 7.B The Spectral Theorem **219**

We will need a few preliminary results, which apply to both real and complex inner product spaces, for our proof of the Real Spectral Theorem.

You could guess that the next result is true and even discover its proof by thinking about quadratic polynomials with real coefﬁcients. Speciﬁcally, sup-

*This technique of completing the square can be used to derive the quadratic formula.*

pose b; c **R** and b2 < 4c. Let x be a

2

real number. Then

2 b 2 b2

x

C bx C c D

x C 2

C

c —

> 0:

4

In particular, x2 bx c is an invertible real number (a convoluted way of saying that it is not 0). Replacing the real number x with a self-adjoint operator (recall the analogy between real numbers and self-adjoint operators),

C C

we are led to the result below.

7.26 Invertible quadratic expressions

Then

Suppose T 2 *L*.V / is self-adjoint and b; c 2 **R** are such that b2 < 4c.

T 2 C bT C cI

is invertible.

Proof Let *v* be a nonzero vector in V. Then

h.T

2

C bT C cI/*v*; *v*iD hT

2*v*; *v*iC bhT *v*; *v*iC ch*v*; *v*i

2

D hT *v*;T *v*iC bhT *v*; *v*iC ck*v*k

2 2

乏 kT *v*k — jbjkT *v*kk*v*kC ck*v*k

C

c —

k*v*k

jbjk*v*k 2

2

D

kT *v*k—

b2 2

4

> 0;

where the third line above holds by the Cauchy–Schwarz Inequality (6.15). The last inequality implies that .T 2 bT cI/*v* 0. Thus T 2 bT cI is injective, which implies that it is invertible (see 3.69).

C C ¤ C C

We know that every operator, self-adjoint or not, on a ﬁnite-dimensional nonzero complex vector space has an eigenvalue (see 5.21). Thus the next result tells us something new only for real inner product spaces.

**220** CHAPTER 7 Operators on Inner Product Spaces

7.27 Self-adjoint operators have eigenvalues

Suppose V ¤ f0g and T 2 *L*.V / is a self-adjoint operator. Then T has

an eigenvalue.

Proof We can assume that V is a real inner product space, as we have already noted. Let n D dim V and choose *v* 2 V with *v* ¤ 0. Then

*v*;T *v*;T 2*v*;:::;T n*v*

cannot be linearly independent, because V has dimension n and we have n 1

C

vectors. Thus there exist real numbers a0;:::; an, not all 0, such that

0 D a0*v* C a1T *v* C ... C anT n*v*:

Make the a’s the coefﬁcients of a polynomial, which can be written in factored form (see 4.17) as

a0 C a1x C ... C anxn

D c.x C b1x C c1/ ... .x C bM x C cM /.x — 入1/ x — 入m/;

2

2

where c is a nonzero real number, each bj , cj , and 入j is real, each bj 2 is less than 4cj , m C M 乏 1, and the equation holds for all real x. We then have

0 D a0*v* C a1T *v* C ... C anT n*v*

D .a0I C a1T C ... C anT /*v*

n

D c.T 2 C b1T C c1I/ ... .T 2 C bM T C cM I/.T — 入1I/ ... .T — 入mI /*v*:

By 7.26, each T 2 bj T cj I is invertible. Recall also that c 0. Thus the equation above implies that m>0 and

C C ¤

0 D .T — 入1I/ ... .T — 入mI /*v*:

Hence T 入j I is not injective for at least one j . In other words, T has an eigenvalue.

—

The next result shows that if U is a subspace of V that is invariant under a self-adjoint operator T, then U ? is also invariant under T. Later we will show that the hypothesis that T is self-adjoint can be replaced with the weaker hypothesis that T is normal (see 9.30).

SECTION 7.B The Spectral Theorem **221**

7.28 Self-adjoint operators and invariant subspaces

Suppose T 2 *L*.V / is self-adjoint and U is a subspace of V that is

invariant under T. Then

1. U ? is invariant under T ;
2. T jU 2 *L*.U / is self-adjoint;
3. T jU ? 2 *L*.U ?/ is self-adjoint.

Proof To prove (a), suppose *v* 2 U ?. Let u 2 U. Then

hT *v*; uiD h*v*;T uiD 0;

where the ﬁrst equality above holds because T is self-adjoint and the second equality above holds because U is invariant under T (and hence Tu U ) and because *v* U ?. Because the equation above holds for each u U, we conclude that T *v* U ?. Thus U ? is invariant under T, completing the proof

2

2 2

2

of (a).

To prove (b), note that if u; *v* 2 U, then

h.T jU /u; *v*iD hT u; *v*iD hu; T *v*iD hu; .T jU /*v*i:

Thus T U is self-adjoint.

j

Now (c) follows from replacing U with U ?

by (a).

in (b), which makes sense

We can now prove the next result, which is one of the major theorems in linear algebra.

7.29 Real Spectral Theorem

Suppose **F** D **R** and T 2 *L*.V /. Then the following are equivalent:

1. T is self-adjoint.
2. V has an orthonormal basis consisting of eigenvectors of T.
3. T has a diagonal matrix with respect to some orthonormal basis of V.

**222** CHAPTER 7 Operators on Inner Product Spaces

Proof First suppose (c) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. A diagonal matrix equals its transpose. Hence T T \*, and thus T is self-adjoint. In other words, (a) holds.

D

We will prove that (a) implies (b) by induction on dim V. To get started, note that if dim V 1, then (a) implies (b). Now assume that dim V > 1 and

D

that (a) implies (b) for all real inner product spaces of smaller dimension.

Suppose (a) holds, so T .V / is self-adjoint. Let u be an eigenvector of T with u 1 (7.27 guarantees that T has an eigenvector, which can then be divided by its norm to produce an eigenvector with norm 1). Let U D span.u/. Then U is a 1-dimensional subspace of V that is invariant

k k D

2 *L*

under T. By 7.28(c), the operator T U ? .U ?/ is self-adjoint.

j 2 *L*

By our induction hypothesis, there is an orthonormal basis of U ?

consist-

ing of eigenvectors of T U ? . Adjoining u to this orthonormal basis of U ? gives an orthonormal basis of V consisting of eigenvectors of T, completing the proof that (a) implies (b).

j

We have proved that (c) implies (a) and that (a) implies (b). Clearly (b) implies (c), completing the proof.

7.30 **Example** Consider the self-adjoint operator T on **R**3 whose matrix

(with respect to the standard basis) is

0@

As you can verify,

14 —13 8

—13 14 8

8 8 —7

A1 :

.1; —1; 0/; .1; 1; 1/; .1; 1; —2/

p2 p3 p6

is an orthonormal basis of **R**3 consisting of eigenvectors of T, and with respect

to this basis, the matrix of T is the diagonal matrix

0@

27 0 0

0 9 0

0 0 —15

A1 :

If **F C**, then the Complex Spectral Theorem gives a complete descrip- tion of the normal operators on V. A complete description of the self-adjoint operators on V then easily follows (they are the normal operators on V whose

D

eigenvalues all are real; see Exercise 6).

D

If **F R**, then the Real Spectral Theorem gives a complete description of the self-adjoint operators on V. In Chapter 9, we will give a complete description of the normal operators on V (see 9.34).

EXERCISES 7.B

SECTION 7.B The Spectral Theorem **223**

1. True or false (and give a proof of your answer): There exists T .**R**3/ such that T is not self-adjoint (with respect to the usual inner product) and such that there is a basis of **R**3 consisting of eigenvectors of T.

2 *L*

1. Suppose that T is a self-adjoint operator on a ﬁnite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that T 2 — 5T C 6I D 0.
2. Give an example of an operator T 2 *L*.**C**3/ such that 2 and 3 are the only eigenvalues of T and T 2 — 5T C 6I ¤ 0.
3. Suppose **F C** and T .V /. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

D 2 *L*

V D E.入1;T / ˚ ... ˚ E.入m;T /;

where 入1;:::; 入m denote the distinct eigenvalues of T.

1. Suppose **F R** and T .V /. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

D 2 *L*

V D E.入1;T / ˚ ... ˚ E.入m;T /;

where 入1;:::; 入m denote the distinct eigenvalues of T.

1. Prove that a normal operator on a complex inner product space is self- adjoint if and only if all its eigenvalues are real.

[*The exercise above strengthens the analogy* (*for normal operators*) *between self-adjoint operators and real numbers.*]

1. Suppose V is a complex inner product space and T 2 *L*.V / is a normal operator such that T 9 D T 8. Prove that T is self-adjoint and T 2 D T.
2. Give an example of an operator T on a complex vector space such that

T 9 D T 8 but T 2 ¤ T.

1. Suppose V is a complex inner product space. Prove that every normal operator on V has a square root. (An operator S 2 *L*.V / is called a ***square root*** of T 2 *L*.V / if S 2 D T.)

**224** CHAPTER 7 Operators on Inner Product Spaces

1. Give an example of a real inner product space V and T .V / and real numbers b; c with b2 < 4c such that T 2 bT cI is not invertible. [*The exercise above shows that the hypothesis that* T *is self-adjoint is*

C C

2 *L*

*needed in 7.26, even for real vector spaces.*]

1. Prove or give a counterexample: every self-adjoint operator on V has a cube root. (An operator S 2 *L*.V / is called a ***cube root*** of T 2 *L*.V / if S 3 D T.)
2. Suppose T 2 *L*.V / is self-adjoint, 入 2 **F**, and E > 0. Suppose there exists *v* 2 V such that k*v*kD 1 and

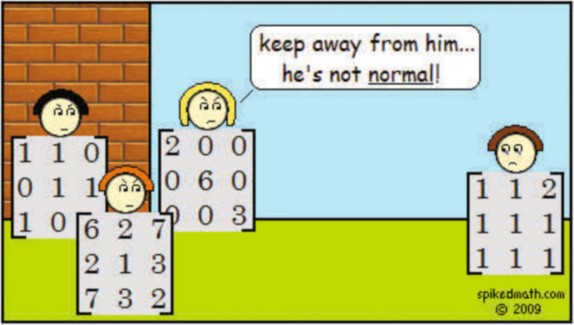
kT *v* — 入*v*k < E:

Prove that T has an eigenvalue 入0 such that j入— 入0j < E.

1. Give an alternative proof of the Complex Spectral Theorem that avoids Schur’s Theorem and instead follows the pattern of the proof of the Real Spectral Theorem.
2. Suppose U is a ﬁnite-dimensional real vector space and T .U /. Prove that U has a basis consisting of eigenvectors of T if and only if there is an inner product on U that makes T into a self-adjoint operator.

2 *L*

1. Find the matrix entry below that is covered up.



SECTION 7.C Positive Operators and Isometries **225**

## *Positive Operators and Isometries*

7.C

### Positive Operators

7.31 **Deﬁnition *positive operator***

An operator T 2 *L*.V / is called ***positive*** if T is self-adjoint and

hT *v*; *v*i乏 0

for all *v* 2 V.

If V is a complex vector space, then the requirement that T is self-adjoint can be dropped from the deﬁnition above (by 7.15).

7.32 **Example *positive operators***

1. If U is a subspace of V, then the orthogonal projection PU is a positive operator, as you should verify.
2. If T 2 *L*.V / is self-adjoint and b; c 2 **R** are such that b2 < 4c, then

T 2 C bT C cI is a positive operator, as shown by the proof of 7.26.

7.33 **Deﬁnition *square root***

An operator R is called a ***square root*** of an operator T if R2 D T.

7.34 **Example** If T .**F**3/ is deﬁned by T .z1; z2; z3/ .z3; 0; 0/, then the operator R .**F**3/ deﬁned by R.z1; z2; z3/ .z2; z3; 0/ is a square root of T.

The characterizations of the positive operators in the next result correspond to characterizations of the nonnegative

*The positive operators correspond*

*terminology would use the term nonnegative instead of positive. However, operator theorists consis- tently call these the positive opera- tors, so we will follow that custom.*

*to the numbers* Œ0; 1/*, so better*

2 *L* D

2 *L* D

numbers among **C**. Speciﬁcally, a com- plex number z is nonnegative if and only if it has a nonnegative square root,

corresponding to condition (c). Also,

z is nonnegative if and only if it has a real square root, corresponding to condition (d). Finally, z is nonnegative if and only if there exists a complex number *w* such that z D *w*N *w*, corresponding to condition (e).

**226** CHAPTER 7 Operators on Inner Product Spaces

7.35 Characterization of positive operators

Let T 2 *L*.V /. Then the following are equivalent:

1. T is positive;
2. T is self-adjoint and all the eigenvalues of T are nonnegative;
3. T has a positive square root;
4. T has a self-adjoint square root;
5. there exists an operator R 2 *L*.V / such that T D R\*R.

Proof We will prove that (a) (b) (c) (d) (e) (a).

First suppose (a) holds, so that T is positive. Obviously T is self-adjoint (by the deﬁnition of a positive operator). To prove the other condition in (b),

) ) ) ) )

suppose 入is an eigenvalue of T. Let *v* be an eigenvector of T corresponding to 入. Then

0 三 hT *v*; *v*iD h入*v*; *v*iD 入h*v*; *v*i:

Thus 入 is a nonnegative number. Hence (b) holds.

Now suppose (b) holds, so that T is self-adjoint and all the eigenvalues of T are nonnegative. By the Spectral Theorem (7.24 and 7.29), there is an orthonormal basis e1;:::; en of V consisting of eigenvectors of T. Let入1;:::; 入n be the eigenvalues of T corresponding to e1;:::; en,; thus each入j is a nonnegative number. Let R be the linear map from V to V such that

q

Rej D 入j ej

for j 1; :::;n (see 3.5). Then R is a positive operator, as you should verify. Furthermore, R2ej 入j ej T ej for each j , which implies that R2 T. Thus R is a positive square root of T. Hence (c) holds.

D D D

D

Clearly (c) implies (d) (because, by deﬁnition, every positive operator is self-adjoint).

Now suppose (d) holds, meaning that there exists a self-adjoint operator R on V such that T R2. Then T R\*R (because R\* R). Hence (e) holds.

D D D

\* Finally, suppose (e) holds. Let R 2 *L*.V / be such that T D R\*R. Then

T D .R\*R/\* D R\*.R\*/\* D R\*R D T. Hence T is self-adjoint. To

complete the proof that (a) holds, note that

hT *v*; *v*iD hR\*R*v*; *v*iD hR*v*; R*v*i乏 0

for every *v* 2 V. Thus T is positive.

SECTION 7.C Positive Operators and Isometries **227**

Each nonnegative number has a unique nonnegative square root. The next result shows that positive operators enjoy a similar property.

*Some mathematicians also use the term* ***positive semideﬁnite opera- tor****, which means the same as posi- tive operator.*

7.36 Each positive operator has only one positive square root

Every positive operator on V has a unique positive square root.

Proof Suppose T .V / is positive. Suppose *v* V is an eigenvector of T.

Thus there exists 入 0 such that T *v*

*A positive operator can have in- ﬁnitely many square roots (al- though only one of them can be positive). For example, the identity*

*operator on* V *has inﬁnitely many square roots if* dim V > 1*.*

乏 D

2

2 *L*

入*v*.

Let R be a positive squarpe root of T.

We will prove that R*v* D 入*v*. This

will imply that the behavior of R on the eigenvectors of T is uniquely deter- mined. Because there is a basis of V consisting of eigenvectors of T (by the

Spectral Theorem), this wpill imply that R is uniquely determined.

To prove that R*v* D 入 *v*, note that the Spectral Theorem asserts that

there is an orthonormal basis e1;:::; en of V consisting of eigenvectors of R. Because R is a positive operator, all its eigenvalues are nonnegative. Thus there exist nonnegative numbers 入1;:::; 入n such that Rej 入j ej for j 1; :::; n.

p

D

D

Because e1;:::; en is a basis of V, we can write

*v* D a1e1 C ... C anen

for some numbers a1;:::; an p2 **F**. Thus p

R*v* D a1

入1e1 C ... C an

入nen

and hence

R2*v* D a1入1e1 C ... C an入nen:

But R2 D T, and T *v* D 入*v*. Thus the equation above implies

a1入e1 C ... C an入en D a1入1e1 C ... C an入nen:

The equation above implies that aj .入— 入j / D 0 for j D 1; :::; n. Hence

X

and thus

*v* D

fj W 入*j* D入g

X

aj ej ;

as desired.

R*v* D

fj W 入*j* D入g

aj p入ej D p入*v*;

**228** CHAPTER 7 Operators on Inner Product Spaces

### Isometries

Operators that preserve norms are sufﬁciently important to deserve a name:

7.37 **Deﬁnition *isometry***

* An operator S 2 *L*.V / is called an ***isometry*** if

kS *v*kD k*v*k

for all *v* 2 V.

* In other words, an operator is an isometry if it preserves norms.

For example, 入 I is an isometry whenever 入 **F** satisﬁes 入 1. We

*The Greek word* ***isos*** *means equal; the Greek word* ***metron*** *means measure. Thus* ***isometry*** *literally means equal measure.*

2 j jD

will see soon that if **F C**, then the next example includes all isometries.

D

2 *L* D

* 1. **Example** Suppose 入1;:::; 入n are scalars with absolute value 1 and S .V / satisﬁes Sej 入j ej for some orthonormal basis e1;:::; en of V. Show that S is an isometry.

Solution Suppose *v* 2 V. Then

* 1. *v* D h*v*; e1ie1 C. . .C h*v*; enien

and

* 1. k*v*k2 D jh*v*; e1ij2 C. . .C jh*v*; enij2;

where we have used 6.30. Applying S to both sides of 7.39 gives

S *v* D h*v*; e1iSe1 C. . .C h*v*; eniSen D 入1h*v*; e1ie1 C ... C 入nh*v*; enien:

The last equation, along with the equation j入j jD 1, shows that

* 1. kS *v*k2 D jh*v*; e1ij2 C. . .C jh*v*; enij2:

Comparing 7.40 and 7.41 shows that *v* S *v* . In other words, S is an isometry.

k kD k k

SECTION 7.C Positive Operators and Isometries **229**

The next result provides several con- ditions that are equivalent to being an isometry. The equivalence of (a) and (b) shows that an operator is an isometry if and only if it preserves inner products. The equivalence of (a) and (c) [or (d)] shows that an operator is an isometry if and only if the list of columns of its

*An isometry on a real inner product space is often called an* ***orthogonal*** *operator. An isometry on a com- plex inner product space is often called a* ***unitary*** *operator. We use the term isometry so that our re- sults can apply to both real and complex inner product spaces.*

matrix with respect to every [or some] basis is orthonormal. Exercise 10 implies that in the previous sentence we can replace “columns” with “rows”.

7.42 Characterization of isometries

Suppose S 2 *L*.V /. Then the following are equivalent:

1. S is an isometry;
2. hSu; S *v*iD hu; *v*i for all u; *v* 2 V ;
3. Se1;:::;Sen is orthonormal for every orthonormal list of vectors

e1;:::; en in V ;

1. there exists an orthonormal basis e1;:::; en of V such that

Se1;:::;Sen is orthonormal;

1. S\*S D I;
2. SS\* D I;
3. S\* is an isometry;
4. S is invertible and S 1 D S \*.

Proof First suppose (a) holds, so S is an isometry. Exercises 19 and 20 in Section 6.A show that inner products can be computed from norms. Because S preserves norms, this implies that S preserves inner products, and hence

(b) holds. More precisely, if V is a real inner product space, then for every

u; *v* 2 V we have

hSu; S *v*iD .kSu C S *v*k

2

2

D .kS.u C *v*/k

2

2

* kSu — S *v*k
* kS.u — *v*/k

2

2

/=4

/=4

D .ku C *v*k

D hu; *v*i;

* ku — *v*k

/=4

**230** CHAPTER 7 Operators on Inner Product Spaces

where the ﬁrst equality comes from Exercise 19 in Section 6.A, the second equality comes from the linearity of S, the third equality holds because S is an isometry, and the last equality again comes from Exercise 19 in Section 6.A. If V is a complex inner product space, then use Exercise 20 in Section 6.A instead of Exercise 19 to obtain the same conclusion. In either case, we see

that (b) holds.

Now suppose (b) holds, so S preserves inner products. Suppose that e1;:::; en is an orthonormal list of vectors in V. Then we see that the list Se1;:::;Sen is orthonormal because Sej ;Sek ej ; ek . Thus (c) holds.

h iD h i

Clearly (c) implies (d).

Now suppose (d) holds. Let e1;:::; en be an orthonormal basis of V such that Se1;:::;Sen is orthonormal. Thus

hS\*Sej ; ekiD hej ; eki

for j; k 1; :::;n [because the term on the left equals Sej ;Sek and

D h i

.Se1;:::;Sen/ is orthonormal]. All vectors u; *v* V can be written as linear combinations of e1;:::; en, and thus the equation above implies that S\*Su; *v* u; *v* . Hence S\*S I; in other words, (e) holds.

2

h iD h i D

Now suppose (e) holds, so that S\*S I. In general, an operator S need not commute with S \*. However, S\*S I if and only if SS\* I; this is a special case of Exercise 10 in Section 3.D. Thus SS\* I, showing that (f)

D

D D

D

holds.

Now suppose (f) holds, so SS\* D I. If *v* 2 V, then

kS\**v*k2 D hS \**v*;S\**v*iD hSS \**v*; *v*iD h*v*; *v*iD k*v*k2:

Thus S\* is an isometry, showing that (g) holds.

Now suppose (g) holds, so S\* is an isometry. We know that (a) ) (e) and

(a) ) (f) because we have shown (a) ) (b) ) (c) ) (d) ) (e) ) (f). Using the implications (a) ) (e) and (a) ) (f) but with S replaced with S\* [and using the equation .S\*/\* D S], we conclude that SS\* D I and S\*S D I.

Thus S is invertible and S 1 D S\*; in other words, (h) holds.

Now suppose (h) holds, so S is invertible and S 1 D S \*. Thus S\*S D I.

If *v* 2 V, then

kS *v*k2 D hS *v*;S*v*iD hS\*S *v*; *v*iD h*v*; *v*iD k*v*k2:

Thus S is an isometry, showing that (a) holds.

) ) ) ) ) ) ) )

We have shown (a) (b) (c) (d) (e) (f) (g) (h) (a),

completing the proof.

SECTION 7.C Positive Operators and Isometries **231**

The previous result shows that every isometry is normal [see (a), (e), and

1. of 7.42]. Thus characterizations of normal operators can be used to give descriptions of isometries. We do this in the next result in the complex case and in Chapter 9 in the real case (see 9.36).

7.43 Description of isometries when **F** D **C**

Suppose V is a complex inner product space and S 2 *L*.V /. Then the

following are equivalent:

1. S is an isometry.
2. There is an orthonormal basis of V consisting of eigenvectors of S

whose corresponding eigenvalues all have absolute value 1.

Proof We have already shown (see Example 7.38) that (b) implies (a).

To prove the other direction, suppose (a) holds, so S is an isometry. By the Complex Spectral Theorem (7.24), there is an orthonormal basis e1;:::; en of V consisting of eigenvectors of S. For j 1; :::; n , let 入 j be the eigenvalue corresponding to ej . Then

2 f g

j入j jD k入j ej kD kSej kD kej kD 1:

Thus each eigenvalue of S has absolute value 1, completing the proof.

EXERCISES 7.C

2 *L*

* 1. Prove or give a counterexample: If T .V / is self-adjoint and there exists an orthonormal basis e1;:::; en of V such that Tej ; ej 0 for each j , then T is a positive operator.
  2. Suppose T is a positive operator on V. Suppose *v*; *w* 2 V are such that

h i乏

T *v* D *w* and T *w* D *v*:

Prove that *v* D *w*.

* 1. Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that T jU 2 *L*.U / is a positive operator on U.
  2. Suppose T .V; W /. Prove that T \*T is a positive operator on V and

2 *L*

TT \* is a positive operator on W.

**232** CHAPTER 7 Operators on Inner Product Spaces

1. Prove that the sum of two positive operators on V is positive.
2. Suppose T .V / is positive. Prove that T k is positive for every positive integer k.

2 *L*

1. Suppose T is a positive operator on V. Prove that T is invertible if and only if

for every *v* 2 V with *v* ¤ 0.

hT *v*; *v*i >0

1. Suppose T 2 *L*.V /. For u; *v* 2 V, deﬁne hu; *v*iT by

hu; *v*iT D hT u; *v*i:

Prove that h.; .iT is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product h.; .i).

1. Prove or disprove: the identity operator on **F**2 has inﬁnitely many self- adjoint square roots.
2. Suppose S 2 *L*.V /. Prove that the following are equivalent:
   1. S is an isometry;
   2. hS\*u; S\**v*iD hu; *v*i for all u; *v* 2 V ;
   3. S\*e1;:::;S\*em is an orthonormal list for every orthonormal list of vectors e1;:::; em in V ;
   4. S\*e1;:::;S\*en is an orthonormal basis for some orthonormal basis e1;:::; en of V.
3. Suppose T1; T2 are normal operators on *L*.**F**3/ and both operators have 2; 5; 7 as eigenvalues. Prove that there exists an isometry S 2 *L*.**F**3/ such that T1 D S\*T2S.
4. Give an example of two self-adjoint operators T1; T2 .**F**4/ such that the eigenvalues of both operators are 2; 5; 7 but there does not exist an isometry S .**F**4/ such that T1 S\*T2S. Be sure to explain why

2 *L*

2 *L* D

there is no isometry with the required property.

1. Prove or give a counterexample: if S .V / and there exists an ortho- normal basis e1;:::; en of V such that Sej 1 for each ej , then S is an isometry.

k kD

2 *L*

1. Let T be the second derivative operator in Exercise 21 in Section 7.A. Show that —T is a positive operator.

SECTION 7.D Polar Decomposition and Singular Value Decomposition **233**

## *Polar Decomposition and Singular*Value Decomposition

7.D

### Polar Decomposition

Recall our analogy between **C** and .V /. Under this analogy, a complex number z corresponds to an operator T, and z corresponds to T \*. The real numbers (z z) correspond to the self-adjoint operators (T T \*), and the

D N D

N

*L*

nonnegative numbers correspond to the (badly named) positive operators.

Another distinguished subset of **C** is the unit circle, which consists of the complex numbers z such that z 1. The condition z 1 is equivalent to the condition zz 1. Under our analogy, this would correspond to the condition T \*T I, which is equivalent to T being an isometry (see 7.42).

D

N D

j jD j jD

In other words, the unit circle in **C** corresponds to the isometries.

Continuing with our analogy, note that each complex number z except 0

can be written in the form

z D z jzjD z pzNz;

jzj

jzj

where the ﬁrst factor, namely, z=jzj, is an element of the unit circle. Our

analogy leads ups to guess that each operator T 2 *L*.V / can be written as an

isometry times T \*T . That guess is indeed correct, as we now prove after

deﬁning the obvious notation, which is justiﬁed by 7.36.

7.44 **Notation**

pT

If T is a positive operator, then pT denotes the unique positive square root of T.

Now we can state and prove the Polar Decomposition, which gives a

beautiful description of an arbitrary operator pon V. Note that T \*T is a

positive operator for every T 2 *L*.V /, and thus T \*T is well deﬁned.

7.45 Polar Decomposition

Suppose T 2 *L*.V /. Then there exists an isometry S 2 *L*.V / such that

T D SpT \*T:

**234** CHAPTER 7 Operators on Inner Product Spaces

Proof If *v* 2 V, then

kT *v*k2 D hT *v*;T *v*iD hT \*T *v*; *v*i

D hpT \*T pT \*T *v*; *v*i D hpT \*T *v*; pT \*T *v*i D kpT \*T *v*k2:

Thus

* 1. kT *v*kD kpT \*T *v*k

for all *v* 2 V.

Deﬁne a linear map S1 W range

T \*T ! range T by

p

* 1. S1.pT \*T *v*/ D T *v*:

The idepa of the proof is to extend S1 to an isometry S 2 *L*.V / such that

T D S T \*T . Now for the details.

First we mpust check thapt S1 is well deﬁned. To do this, suppose *v*1; *v*2 2 V

are such that T \*T *v*1 D T \*T *v*2. For the deﬁnition given by 7.47 to make

sense, we must show that T *v*1 D T *v*2. Note that

kT *v*1 — T *v*2kD kT .*v*1 — *v*2/k

p

D k T \*T .*v*1 — *v*2/k

p p

D k T \*T *v*1 — T \*T *v*2k D 0;

where the second equality holds by 7.46. The equation above shows that

T *v*1 D T *v*2, so S1 is indeed well deﬁned. You should verify that S1 is a

linear map.

p

We see from 7.47 that S1 maps range T

and 7.47 imply that

\*T onto range T. Clearly 7.46

p kS1ukD kuk

for all u 2 range T \*T .

*The rest of the proof extends* S1 *to an isometry* S *on all of* V*.*

In particular, S1 is injective. Thus from the Fundamental Theorem of Lin- ear Maps (3.22), applied to S1, we have

dim range pT \*T D dim range T:

SECTION 7.D Polar Decomposition and Singular Value Decomposition **235**

This implies that dim.range pT \*T /? D dpim .range T /? (see 6.50).

Thus orthonormal bases e1;:::; em of .range T \*T /? and f1;:::; fm

of .range T /? can be chosen; the key point here is that these two ortho-

normal baseps have the same length (denoted m). Now deﬁne a linear map

S2 W .range T \*T /? ! .range T /? by

S2.a1e1 C . .. C amem/ D a1f1 C ... C amfm:

For all *w* 2 .range pT \*T /?, we have kS2*w*kD k*w*k (fromp6.25).

Now let Spbe the operator on V that equals S1 on range T \*T and equals

S2 on .range T \*T /?. More precisely, recall that each *v* 2 V can be written

uniquely in the form

* 1. *v* D u C *w*;

p

where u 2 range T \*T and *w* 2 .range pT \*T /? (see 6.47). For *v* 2 V

with decomposition as above, deﬁne S *v* by

S *v* D S1u C S2*w*:

For each *v* 2 V we have

S.pT \*T *v*/ D S1.pT \*T *v*/ D T *v*;

so T SpT \*T , as desired. All that remains is to show that S is an isometry.

D

However, this follows easily from two uses of the Pythagorean Theorem: if

*v* 2 V has decomposition as in 7.48, then

kS *v*k D kS1u C S2*w*k2 D kS1uk C kS2*w*k2 D kuk C k*w*k2 D k*v*k I

2

2

2

2

the second equality holds because S1u 2 range T and S2*w* 2 .range T /?.

The Polar Decomposition (7.45) states that each operator on V is the product of an isometry and a positive operator. Thus we can write each operator on V as the product of two operators, each of which comes from a class that we can completely describe and that we understand reasonably

well. The isometries are described by 7.43 and 9.36; the positive operators

are described by the Spectral Theorem (7.24 and 7.29).

p

Speciﬁcally, consider the case **F** D **C**, and suppose T D S T \*T is a

Polar Decomposition of an operator T 2 *L*.V /, where S is an isometry. Then

there is an orthonormal basis of V with respect to which S has a dipagonal

matrix, and there is an orthonormal basis of V with respect to which T \*T

has a diagonal matrix. **Warning:** there may not existpan orthonormal basis

that simultaneously puts the matrices of both S and T \*T into these nice

diagonal forms. In other words, S may require one orthonormal basis and

p

T \*T may require a different orthonormal basis.

**236** CHAPTER 7 Operators on Inner Product Spaces

### Singular Value Decomposition

The eigenvalues of an operator tell us something about the behavior of the operator. Another collection of numbers, called the singular values, is also

useful. Recall that eigenspaces and the notation E are deﬁned in 5.36.

7.49 **Deﬁnition *singular values***

Suppose

of T \*T , with each eigenvalue 入repeated dim E.入; T \*T/ times.

p

T

2 *L*.V /. The ***singular values*** of T are the eigenvalues

p

The singular values of T arpe all nonnegative, because they are the eigen-

values of the positive operator T \*T .

7.50 **Example** Deﬁne T 2 *L*.**F**4/ by

T .z1; z2; z3; z4/ D .0; 3z1; 2z2; —3z4/:

Find the singular values of T.

Solution A calculation shows T \*T .z1; z2; z3; z4/ .9z1; 4z2; 0; 9z4/, as you should verify. Thus

D

pT \*T .z1; z2; z3; z4/ D .3z1; 2z2; 0; 3z4/;

and we see that the eigenvalues of pT \*T are 3; 2; 0 and

dim E.3; pT \*T/ D 2; dim E.2; pT \*T/ D 1; dim E.0; pT \*T/ D 1:

Hence the singular values of T are 3; 3; 2; 0.

Note that 3 and 0 are the only eigenvalues of T. Thus in this case, the collection of eigenvalues did not pick up the number 2 that appears in the deﬁnition (and hence the behavior) of T, but the collection of singular values does include 2.

—

Each T 2 *L*.V / has dim V singular values, as can be seen by applying

the Spectral Theorem pand 5.41 [see especially part (e)] to the positive (hence

self-adjoint) operator T \*T . For example, the operator T deﬁned in Exam-

ple 7.50 on the four-dimensional vector space **F**4 has four singular values (they are 3; 3; 2; 0), as we saw above.

The next result shows that every operator on V has a clean description in

terms of its singular values and two orthonormal bases of V.

SECTION 7.D Polar Decomposition and Singular Value Decomposition **237**

Proof By the Spectral Theorempapplied to pT \*T , there is an orthonormal

7.51 Singular Value Decomposition

Suppose T 2 *L*.V / has singular values s1;:::; sn. Then there exist

orthonormal bases e1;:::; en and f1;:::; fn of V such that

T *v* D s1h*v*; e1if1 C ... C snh*v*; enifn

for every *v* 2 V.

basis e1;:::; en of V such that

We have

T \*T ej D sj ej for j D 1; :::; n.

*v* D h*v*; e1ie1 Cp. . .C h*v*; enien

for every *v* 2 V (see 6.30). Apply T \*T to both sides of this equation,

getting

p

T \*T *v* D s1h*v*; e1ie1 C ... C snh*v*; enien

for every *v* 2 V. By the PolarpDecomposition (see 7.45), there is an isometry

1. .V / such that T S T \*T . Apply S to both sides of the equation

2 *L* D

above, getting

T *v* D s1h*v*; e1iSe1 C ... C snh*v*; eniSen

for every *v* V. For each j , let fj Sej . Because S is an isometry, f1;:::; fn is an orthonormal basis of V (see 7.42). The equation above now becomes

2 D

T *v* D s1h*v*; e1if1 C ... C snh*v*; enifn

for every *v* 2 V, completing the proof.

When we worked with linear maps from one vector space to a second vector space, we considered the matrix of a linear map with respect to a basis of the ﬁrst vector space and a basis of the second vector space. When dealing with operators, which are linear maps from a vector space to itself, we almost always use only one basis, making it play both roles.

The Singular Value Decomposition allows us a rare opportunity to make

good use of two different bases for the matrix of an operator. To do this, suppose T .V /. Let s1;:::; sn denote the singular values of T, and let e1;:::; en and f1;:::; fn be orthonormal bases of V such that the Singular

2 *L*

Value Decomposition 7.51 holds. Because T ej D sj fj for each j, we have

（ ） 0B s1 0 1C

*M* T; .e1;:::; en/; .f1;:::; fn/

D @

: : :

0 sn

A :

**238** CHAPTER 7 Operators on Inner Product Spaces

In other words, every operator on V has a diagonal matrix with respect to some orthonormal bases of V, provided that we are permitted to use two different bases rather than a single basis as customary when working with

operators.

Singular values and the Singular Value Decomposition have many applica- tions (some are given in the exercises), including applications in computational

linear algebra. To compute numeric approximations to the singular values of an operator T, ﬁrst compute T \*T and then compute approximations to the eigenvalues of T \*T (good techniques exist for approximating eigenvalues

of positive operators). The nonnegative square roots of these (approximate) eigenvalues of T \*T will be the (approximate) singular values of T. In other words, the singular values of T can be approximated without computing the

squpare root of T \*T. The next result helps justify working with T \*T instead

of T \*T .

7.52 Singular values without taking square root of an operator

Suppose T 2 *L*.V /. Then the singular values of T are the nonnegative

square roots of the eigenvalues of T \*T, with each eigenvalue 入repeated dim E.入; T \*T/ times.

Proof The Spectral Theorem implies that there are an orthonormal basis

e1;:::; en and nonnegative numbers p入1;::: ; 入n supch that T \*T ej D 入j ej

D

for j 1; :::; n. It is easy to see that

which implies the desired result.

T \*T ej D

入j ej for j D 1; :::; n,

EXERCISES 7.D

1. Fix u; x 2 V with u ¤ 0. Deﬁne T 2 *L*.V / by

T *v* D h*v*; uix

for every *v* 2 V. Prove that

pT \*T *v* kxk *v*; u u

D h i

kuk

for every *v* 2 V.

1. Give an example of T .**C**2/ such that 0 is the only eigenvalue of T

2 *L*

and the singular values of T are 5; 0.

SECTION 7.D Polar Decomposition and Singular Value Decomposition **239**

1. Suppose T 2 *L*.V /. Prove that there exists an isometry S 2 *L*.V / such

that p \*

T D TT S:

1. Suppose T 2 *L*.V / and s is a singular value of T. Prove that there exists a vector *v* 2 V such that k*v*kD 1 and kT *v*kD s.
2. Suppose T .**C**2/ is deﬁned by T .x; y/ . 4y; x/. Find the singu- lar values of T.

2 *L* D —

1. Find the singular values of the differentiation operator D .**R**2/ deﬁned by Dp p0, where the inner product on .**R**2/ is as in Example 6.33.

D *P*

2 *P*

1. Deﬁne T 2 *L*.**F**3/ by

T .z1; z2; z3/ D .z3; 2z1; 3z2/:

Find (explicitly) an isometry S 2 *L*.**F**3/ such that T D SpT \*T .

1. Suppose T 2 *L*.V /, S 2 *L*.V / is an isometry, andpR 2 *L*.V / is a

positive operator such that T D SR. Prove that R D T \*T .

[*The exercise above shows that if we write* T *as the product of an isometry*

*and a positive operator* (p*as in the Polar Decomposition 7.45*)*, then the*

*positive operator equals* T \*T *.*]

1. Suppose T 2 *L*.V /. Prove that T is invertible pif and only if there exists

a unique isometry S 2 *L*.V / such that T D S T \*T .

1. Suppose T .V / is self-adjoint. Prove that the singular values of T

2 *L*

equal the absolute values of the eigenvalues of T, repeated appropriately.

1. Suppose T 2 *L*.V /. Prove that T and T \* have the same singular values.
2. Prove or give a counterexample: if T .V /, then the singular values of T 2 equal the squares of the singular values of T.

2 *L*

1. Suppose T .V /. Prove that T is invertible if and only if 0 is not a singular value of T.

2 *L*

1. Suppose T .V /. Prove that dim range T equals the number of nonzero singular values of T.

2 *L*

1. Suppose S .V /. Prove that S is an isometry if and only if all the singular values of S equal 1.

2 *L*

**240** CHAPTER 7 Operators on Inner Product Spaces

1. Suppose T1; T2 2 *L*.V /. Prove that T1 and T2 have the same singular values if and only if there exist isometries S1; S2 2 *L*.V / such that T1 D S1T2S2.
2. Suppose T 2 *L*.V / has singular value decomposition given by

T *v* D s1h*v*; e1if1 C ... C snh*v*; enifn

for every *v* V, where s1;:::; sn are the singular values of T and

2

e1;:::; en and f1;:::; fn are orthonormal bases of V.

* 1. Prove that if *v* 2 V, then

T \**v* D s1h*v*; f1ie1 C ... C snh*v*; fnien:

* 1. Prove that if *v* 2 V, then

T \*T *v* D s12h*v*; e1ie1 C ... C sn2h*v*; enien:

* 1. Prove that if *v* 2 V, then

pT \*T *v* D s1h*v*; e1ie1 C ... C snh*v*; enien:

* 1. Suppose T is invertible. Prove that if *v* 2 V, then

T 1*v* D h*v*; f1ie1 C ... C h*v*; fnien

s1 sn

for every *v* 2 V.

1. Suppose T .V /. Let s denote the smallest singular value of T, and let s denote the largest singular value of T.

2 *L* O

* 1. Prove that sOk*v*k三 kT *v*k三 sk*v*k for every *v* 2 V.
  2. Suppose 入is an eigenvalue of T. Prove that sO 三 j入j三 s.

1. Suppose T 2 *L*.V /. Show that T is uniformly continuous with respect to the metric d on V deﬁned by d.u; *v*/ D ku — *v*k.
2. Suppose S; T .V /. Let s denote the largest singular value of S, let t denote the largest singular value of T, and let r denote the largest singular value of S C T. Prove that r 三 s C t .

2 *L*

# *Operators on Complex* Vector Spaces



*Hypatia, the 5th century Egyptian mathematician and philosopher, as envisioned around 1900 by Alfred Seifert.*

8

CHAPTER

In this chapter we delve deeper into the structure of operators, with most of the attention on complex vector spaces. An inner product does not help with

this material, so we return to the general setting of a ﬁnite-dimensional vector space. To avoid some trivialities, we will assume that V 0 . Thus our assumptions for this chapter are as follows:

¤ f g

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a ﬁnite-dimensional nonzero vector space over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

generalized eigenvectors and generalized eigenspaces characteristic polynomial and the Cayley–Hamilton Theorem decomposition of an operator

minimal polynomial Jordan Form

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

241

DOI 10.1007/978-3-319-11080-6 8

**242** CHAPTER 8 Operators on Complex Vector Spaces

## *Generalized Eigenvectors and Nilpotent*Operators

8.A

### Null Spaces of Powers of an Operator

We begin this chapter with a study of null spaces of powers of an operator.

8.2 Sequence of increasing null spaces

Suppose T 2 *L*.V /. Then

f0gD null T 0 c null T 1 c . . .c null T k c null T kC1 c . . . :

Proof Suppose k is a nonnegative integer and *v*2 null T k. Then T k*v* D 0, and hence T kC1*v* D T .T k*v*/ D T .0/ D 0. Thus *v* 2 null T kC1. Hence null T k c null T kC1, as desired.

The next result says that if two consecutive terms in this sequence of subspaces are equal, then all later terms in the sequence are equal.

8.3 Equality in the sequence of null spaces

Suppose T 2 *L*.V /. Suppose m is a nonnegative integer such that

null T m D null T mC1. Then

null T m D null T mC1 D null T mC2 D null T mC3 D . . . :

Proof Let k be a positive integer. We want to prove that

null T mCk D null T mCkC1:

We already know from 8.2 that null T mCk null T mCkC1.

c

To prove the inclusion in the other direction, suppose *v* null T mCkC1.

2

Then

Hence

T mC1.T k*v*/ D T mCkC1*v* D 0:

T k*v* 2 null T mC1 D null T m:

Thus T mCk*v* D T m.T k*v*/ D 0, which means that *v* 2 null T mCk. This implies that null T mCkC1 c null T mCk, completing the proof.

SECTION 8.A Generalized Eigenvectors and Nilpotent Operators **243**

The proposition above raises the question of whether there exists a non- negative integer m such that null T m null T mC1. The proposition below shows that this equality holds at least when m equals the dimension of the vector space on which T operates.

D

8.4 Null spaces stop growing

Suppose T 2 *L*.V /. Let n D dim V. Then

null T n D null T nC1 D null T nC2 D . . . :

Proof We need only prove that null T n null T nC1 (by 8.3). Suppose this is not true. Then, by 8.2 and 8.3, we have

f0gD null T 0 ¨ null T 1 ¨ ... ¨ null T n ¨ null T nC1;

D

where the symbol ¨ means “contained in but not equal to”. At each of the strict inclusions in the chain above, the dimension increases by at least 1. Thus dim null T nC1 n 1, a contradiction because a subspace of V cannot have a larger dimension than n.

乏 C

Unfortunately, it is not true that V null T range T for each T .V /.

D ˚ 2 *L*

However, the following result is a useful substitute.

8.5 V is the direct sum of null T dim V and range T dim V

Suppose T 2 *L*.V /. Let n D dim V. Then

V D null T n ˚ range T n:

Proof First we show that

* 1. .null T n/ \ .range T n/ D f0g:

Suppose *v* 2 .null T n/ \ .range T n/. Then T n*v* D 0, and there exists u 2 V such that *v* D T nu. Applying T n to both sides of the last equation shows that T n*v* D T 2nu. Hence T 2nu D 0, which implies that T nu D 0 (by 8.4). Thus *v* D T nu D 0, completing the proof of 8.6.

Now 8.6 implies that null T n C range T n is a direct sum (by 1.45). Also,

dim.null T n ˚ range T n/ D dim null T n C dim range T n D dim V;

where the ﬁrst equality above comes from 3.78 and the second equality comes from the Fundamental Theorem of Linear Maps (3.22). The equation above

implies that null T n ˚ range T n D V, as desired.

**244** CHAPTER 8 Operators on Complex Vector Spaces

* 1. **Example** Suppose T 2 *L*.**F**3/ is deﬁned by

T .z1; z2; z3/ D .4z2; 0; 5z3/:

For this operator, null T C range T is not a direct sum of subspaces, because null T D f.z1; 0; 0/ W z1 2 **F**g and range T D f.z1; 0; z3/ W z1; z3 2 **F**g. Thus null T \ range T ¤ f0g and hence null T C range T is not a direct sum. Also note that null T C range T ¤ **F**3.

However, we have T 3.z1; z2; z3/ D .0; 0; 125z3/. Thus we see that null T 3 D f.z1; z2; 0/ W z1; z2 2 **F**g and range T 3 D f.0; 0; z3/ W z3 2 **F**g. Hence **F**3 D null T 3 ˚ range T 3.

### Generalized Eigenvectors

Unfortunately, some operators do not have enough eigenvectors to lead to a good description. Thus in this subsection we introduce the concept of generalized eigenvectors, which will play a major role in our description of the structure of an operator.

To understand why we need more than eigenvectors, let’s examine the question of describing an operator by decomposing its domain into invariant

subspaces. Fix T .V /. We seek to describe T by ﬁnding a “nice” direct

2 *L*

sum decomposition

V D U1 ˚ ... ˚ Um;

where each Uj is a subspace of V invariant under T. The simplest possible nonzero invariant subspaces are 1-dimensional. A decomposition as above where each Uj is a 1-dimensional subspace of V invariant under T is possible if and only if V has a basis consisting of eigenvectors of T (see 5.41). This happens if and only if V has an eigenspace decomposition

**8.8** V D E.入1;T / ˚ ... ˚ E.入m;T /;

where 入1;:::; 入m are the distinct eigenvalues of T (see 5.41).

The Spectral Theorem in the previous chapter shows that if V is an inner

product space, then a decomposition of the form 8.8 holds for every normal

operator if **F C** and for every self-adjoint operator if **F R** because operators of those types have enough eigenvectors to form a basis of V (see 7.24 and 7.29).

D D

SECTION 8.A Generalized Eigenvectors and Nilpotent Operators **245**

Sadly, a decomposition of the form 8.8 may not hold for more general oper- ators, even on a complex vector space. An example was given by the operator in 5.43, which does not have enough eigenvectors for 8.8 to hold. General- ized eigenvectors and generalized eigenspaces, which we now introduce, will remedy this situation.

8.9 **Deﬁnition *generalized eigenvector***

Suppose T 2 *L*.V / and 入 is an eigenvalue of T. A vector *v* 2 V is called

a ***generalized eigenvector*** of T corresponding to 入 if *v* ¤ 0 and

.T — 入I / *v* D 0

for some positive integer j .

j

Although j is allowed to be an arbi- trary integer in the equation

.T — 入I /j *v* D 0

*Note that we do not deﬁne the con- cept of a generalized eigenvalue, because this would not lead to any-*

*thing new. Reason: if* .T — 入I /j *is*

*not injective for some positive inte-*

*ger* j *, then* T — 入I *is not injective,*

*and hence* 入 *is an eigenvalue of* T*.*

in the deﬁnition of a generalized eigen- vector, we will soon prove that every

generalized eigenvector satisﬁes this equation with j D dim V.

8.10 **Deﬁnition *generalized eigenspace,*** G.入; T /

Suppose T 2 *L*.V / and 入 2 **F**. The ***generalized eigenspace*** of T corre-

sponding to 入, denoted G.入; T /, is deﬁned to be the set of all generalized eigenvectors of T corresponding to 入, along with the 0 vector.

Because every eigenvector of T is a generalized eigenvector of T (take j 1 in the deﬁnition of generalized eigenvector), each eigenspace is contained in the corresponding generalized eigenspace. In other words, if

D

T 2 *L*.V / and 入 2 **F**, then

E.入; T / c G.入; T /:

The next result implies that if T .V / and 入 **F**, then G. 入 ; T / is a subspace of V (because the null space of each linear map on V is a subspace of V ).

2 *L* 2

**246** CHAPTER 8 Operators on Complex Vector Spaces

8.11 Description of generalized eigenspaces

Suppose T 2 *L*.V / and 入 2 **F**. Then G.入; T / D null.T — 入I /dim V .

Proof Suppose *v* null.T 入I /dim V . The deﬁnitions imply *v* G.入; T /. Thus G.入; T / null.T 入I /dim V .

� —

2 — 2

Conversely, suppose *v* G.入; T /. Thus there is a positive integer j such

2

that

*v* 2 null.T — 入I /j :

From 8.2 and 8.4 (with T — 入I replacing T ), we get *v* 2 null.T — 入I /dim V . Thus G.入; T / c null.T — 入I /dim V , completing the proof.

8.12 **Example** Deﬁne T 2 *L*.**C**3/ by

T .z1; z2; z3/ D .4z2; 0; 5z3/:

1. Find all eigenvalues of T, the corresponding eigenspaces, and the corresponding generalized eigenspaces.
2. Show that **C**3 is the direct sum of generalized eigenspaces correspond- ing to the distinct eigenvalues of T.

Solution

1. A routine use of the deﬁnition of eigenvalue shows that the eigenvalues of T are 0 and 5. The corresponding eigenspaces are easily seen to be E.0; T / D f.z1; 0; 0/ W z1 2 **C**g and E.5; T / D f.0; 0; z3/ W z3 2 **C**g.

Note that this operator T does not have enough eigenvectors to span its domain **C**3.

We have T 3.z1; z2; z3/ D .0; 0; 125z3/ for all z1; z2; z3 2 **C**. Thus

8.11 implies that G.0; T / D f.z1; z2; 0/ W z1; z2 2 **C**g.

We have .T — 5I/3.z1; z2; z3/ D .—125z1 C 300z2; —125z2; 0/. Thus

8.11 implies that G.5; T / D f.0; 0; z3/ W z3 2 **C**g.

1. The results in part (a) show that **C**3 D G.0; T / ˚ G.5; T /.

SECTION 8.A Generalized Eigenvectors and Nilpotent Operators **247**

One of our major goals in this chapter is to show that the result in part (b) of the example above holds in general for operators on ﬁnite-dimensional complex vector spaces; we will do this in 8.21.

We saw earlier (5.10) that eigenvectors corresponding to distinct eigenval- ues are linearly independent. Now we prove a similar result for generalized eigenvectors.

8.13 Linearly independent generalized eigenvectors

Let T 2 *L*.V /. Suppose 入1;:::; 入m are distinct eigenvalues of T and

*v*1;:::; *v*m are corresponding generalized eigenvectors. Then *v*1;:::; *v*m

is linearly independent.

Proof Suppose a1;:::; am are complex numbers such that

* 1. 0 D a1*v*1 C ... C am*v*m:

Let k be the largest nonnegative integer such that .T — 入1I /k*v*1 ¤ 0. Let

*w* D .T — 入1I /k*v*1:

Thus

.T — 入1I/*w* D .T — 入1I /kC1*w* D 0;

and hence T *w* 入1*w*. Thus .T 入I /*w* .入1 入/*w* for every 入 **F** and hence

D — D — 2

* 1. .T — 入I /n*w* D .入1 — 入/n*w*

for every 入 **F**, where n dim V. Apply the operator

2 D

.T — 入1I /k.T — 入2I /n ... .T — 入mI /n

to both sides of 8.14, getting

0 D a1.T — 入1I/k.T — 入2I /n ... .T — 入mI /n*v*1

n n

D a .T — 入 I/

... .T — 入 I/ *w*

1 2 m

n n

D a1.入1 — 入2/ ... .入1 — 入m/ *w*;

where we have used 8.11 to get the ﬁrst equation above and 8.15 to get the last equation above.

The equation above implies that a1 0. In a similar fashion, aj 0 for each j , which implies that *v*1;:::; *v*m is linearly independent.

D D

**248** CHAPTER 8 Operators on Complex Vector Spaces

### Nilpotent Operators

8.16 **Deﬁnition *nilpotent***

An operator is called ***nilpotent*** if some power of it equals 0.

8.17 **Example *nilpotent operators***

1. The operator N 2 *L*.**F**4/ deﬁned by

N.z1; z2; z3; z4/ D .z3; z4; 0; 0/

is nilpotent because N 2 D 0.

1. The operator of differentiation on *P*m.**R**/ is nilpotent because the

.m C 1/st derivative of every polynomial of degree at most m equals 0. Note that on this space of dimension m C 1, we need to raise the nilpotent operator to the power m C 1 to get the 0 operator.

*The Latin word* ***nil*** *means noth- ing or zero; the Latin word* ***potent*** *means power. Thus* ***nilpotent*** *liter- ally means zero power.*

The next result shows that we never need to use a power higher than the di- mension of the space.

8.18 Nilpotent operator raised to dimension of domain is 0

Suppose N 2 *L*.V / is nilpotent. Then N dim V D 0.

Proof Because N is nilpotent, G.0; N / D V. Thus 8.11 implies that null N dim V D V, as desired.

Given an operator T on V, we want to ﬁnd a basis of V such that the matrix of T with respect to this basis is as simple as possible, meaning that the matrix contains many 0’s.

The next result shows that if N is

*If* V *is a complex vector space, a*

*proof of the next result follows eas-*

*ily from Exercise 7, 5.27, and 5.32. But the proof given here uses sim- pler ideas than needed to prove 5.27, and it works for both real and complex vector spaces.*

nilpotent, then we can choose a basis of V such that the matrix of N with respect to this basis has more than half of its entries equal to 0. Later in this chapter we will do even better.

SECTION 8.A Generalized Eigenvectors and Nilpotent Operators **249**

8.19 Matrix of a nilpotent operator

Suppose N is a nilpotent operator on V. Then there is a basis of V with respect to which the matrix of N has the form

B

0

@

: :

:

0

\*

0

C

A

I

here all entries on and below the diagonal are 0’s.

Proof First choose a basis of null N. Then extend this to a basis of null N 2. Then extend to a basis of null N 3. Continue in this fashion, eventually getting a basis of V (because 8.18 states that null N dim V V ).

Now let’s think about the matrix of N with respect to this basis. The

D

ﬁrst column, and perhaps additional columns at the beginning, consists of all 0’s, because the corresponding basis vectors are in null N. The next set of columns comes from basis vectors in null N 2. Applying N to any such vector, we get a vector in null N ; in other words, we get a vector that is a

linear combination of the previous basis vectors. Thus all nonzero entries in these columns lie above the diagonal. The next set of columns comes from

basis vectors in null N 3. Applying N to any such vector, we get a vector in null N 2; in other words, we get a vector that is a linear combination of the

previous basis vectors. Thus once again, all nonzero entries in these columns lie above the diagonal. Continue in this fashion to complete the proof.

EXERCISES 8.A

1. Deﬁne T 2 *L*.**C**2/ by

T .*w*; z/ D .z; 0/:

Find all generalized eigenvectors of T.

1. Deﬁne T 2 *L*.**C**2/ by

T .*w*; z/ D .—z; *w*/:

Find the generalized eigenspaces corresponding to the distinct eigenval- ues of T.

**250** CHAPTER 8 Operators on Complex Vector Spaces

1. Suppose T 2 *L*.V / is invertible. Prove that G.入; T / D G. 1 ;T 1/ for

入

every 入 2 **F** with 入 ¤ 0.

1. Suppose T 2 *L*.V / and ˛; ˇ 2 **F** with ˛ ¤ ˇ. Prove that

G.˛; T / \ G.ˇ; T / D f0g:

1. Suppose T 2 *L*.V /, m is a positive integer, and *v* 2 V is such that

T m 1*v* ¤ 0 but T m*v* D 0. Prove that

*v*;T *v*;T 2*v*;:::;T m 1*v*

is linearly independent.

1. Suppose T .**C**3/ is deﬁned by T .z1; z2; z3/ .z2; z3; 0/. Prove that T has no square root. More precisely, prove that there does not exist S 2 *L*.**C**3/ such that S 2 D T.

2 *L* D

1. Suppose N .V / is nilpotent. Prove that 0 is the only eigenvalue of N.

2 *L*

1. Prove or give a counterexample: The set of nilpotent operators on V is a subspace of *L*.V /.
2. Suppose S; T 2 *L*.V / and ST is nilpotent. Prove that TS is nilpotent.
3. Suppose that T 2 *L*.V / is not nilpotent. Let n D dim V. Show that

V D null T n 1 ˚ range T n 1.

1. Prove or give a counterexample: If V is a complex vector space and dim V D n and T 2 *L*.V /, then T n is diagonalizable.
2. Suppose N .V / and there exists a basis of V with respect to which N has an upper-triangular matrix with only 0’s on the diagonal. Prove that N is nilpotent.

2 *L*

1. Suppose V is an inner product space and N 2 *L*.V / is normal and nilpotent. Prove that N D 0.
2. Suppose V is an inner product space and N .V / is nilpotent. Prove that there exists an orthonormal basis of V with respect to which N has an upper-triangular matrix.

2 *L*

[*If* F **C***, then the result above follows from Schur’s Theorem* (*6.38*) *without the hypothesis that* N *is nilpotent. Thus the exercise above needs to be proved only when* **F** D **R***.*]

D

SECTION 8.A Generalized Eigenvectors and Nilpotent Operators **251**

1. Suppose N .V / is such that null N dim V 1 null N dim V . Prove that N is nilpotent and that

2 *L* ¤

dim null N j D j

for every integer j with 0 三 j 三 dim V.

1. Suppose T 2 *L*.V /. Show that

V D range T 0 � range T 1 � . . .� range T k � range T kC1 � . . . :

1. Suppose T 2 *L*.V / and m is a nonnegative integer such that

range T m D range T mC1:

Prove that range T k D range T m for all k> m.

1. Suppose T 2 *L*.V /. Let n D dim V. Prove that

range T n D range T nC1 D range T nC2 D . . . :

1. Suppose T 2 *L*.V / and m is a nonnegative integer. Prove that

null T m D null T mC1 if and only if range T m D range T mC1:

1. Suppose T .**C**5/ is such that range T 4 range T 5. Prove that T is nilpotent.

2 *L* ¤

1. Find a vector space W and T .W / such that null T k ¨ null T kC1

2 *L*

and range T k © range T kC1 for every positive integer k.

**252** CHAPTER 8 Operators on Complex Vector Spaces

## *Decomposition of an Operator*

8.B

### Description of Operators on Complex Vector Spaces

We saw earlier that the domain of an operator might not decompose into eigenspaces, even on a ﬁnite-dimensional complex vector space. In this section we will see that every operator on a ﬁnite-dimensional complex vector space has enough generalized eigenvectors to provide a decomposition.

We observed earlier that if T .V /, then null T and range T are invari- ant under T [see 5.3, parts (c) and (d)]. Now we show that the null space and the range of each polynomial of T is also invariant under T.

2 *L*

8.20 The null space and range of p.T / are invariant under T

Suppose T 2 *L*.V / and p 2 *P*.**F**/. Then null p.T / and range p.T / are

invariant under T.

Proof Suppose *v* 2 null p.T /. Then p.T /*v* D 0. Thus

.p.T / .T *v*/ D T p.T /*v* D T .0/ D 0:

（ ） （ ）

Hence T *v* null p.T /. Thus null p.T / is invariant under T, as desired.

2

Suppose *v* range p.T /. Then there exists u V such that *v* p.T /u.

2 2 D

Thus

（ ）

T *v* D T p.T /u D p.T /.T u/:

Hence T *v* 2 range p.T /. Thus range p.T / is invariant under T, as desired.

The following major result shows that every operator on a complex vector space can be thought of as composed of pieces, each of which is a nilpotent operator plus a scalar multiple of the identity. Actually we have already done

the hard work in our discussion of the generalized eigenspaces G.入; T /, so at

this point the proof is easy.

8.21 Description of operators on complex vector spaces

Suppose V is a complex vector space and T 2 *L*.V /. Let 入1;:::; 入m be

the distinct eigenvalues of T. Then

(a) V D G.入1;T / ˚ ... ˚ G.入m;T /;

1. each G.入j ;T / is invariant under T ;
2. each .T — 入j I /jG.入*j* ;T / is nilpotent.

SECTION 8.B Decomposition of an Operator **253**

Proof Let n dim V. Recall that G.入j ;T / null.T 入j I/n for each j (by 8.11). From 8.20 [with p.z/ .z 入 j /n], we get (b). Obviously (c) follows from the deﬁnitions.

D —

D D —

We will prove (a) by induction on n. To get started, note that the desired result holds if n 1. Thus we can assume that n>1 and that the desired result holds on all vector spaces of smaller dimension.

D

Because V is a complex vector space, T has an eigenvalue (see 5.21); thus

m 乏 1. Applying 8.5 to T — 入1I shows that

**8.22** V D G.入1;T / ˚ U;

where U range.T 入 1I/n. Using 8.20 [with p.z/ .z 入 1/n], we see that U is invariant under T. Because G.入1;T / 0 , we have dim U < n. Thus we can apply our induction hypothesis to T U.

j

¤ f g

D — D —

None of the generalized eigenvectors of T U correspond to the eigenvalue 入1, because all generalized eigenvectors of T corresponding to 入1 are in G.入1;T /. Thus each eigenvalue of T U is in 入2;:::; 入m .

j

j f g

By our induction hypothesis, U G.入2;T U / G.入m;T U /.

D j ˚ . . .˚ j

Combining this information with 8.22 will complete the proof if we can show that G.入k;T U / G.入k;T / for k 2; :::; m.

2 f g j c

j D D

Thus ﬁx k 2; :::; m . The inclusion G.入k;T U / G.入k;T / is clear.

To prove the inclusion in the other direction, suppose *v* G.入k;T /. By 8.22, we can write *v v*1 u, where *v*1 G. 入 1;T / and u U. Our induction hypothesis implies that

D C 2 2

2

u D *v*2 C C *v*m;

where each *v*j is in G.入j ;T jU /, which is a subset of G.入j ;T /. Thus

*v* D *v*1 C *v*2 C C *v*m;

Because generalized eigenvectors corresponding to distinct eigenvalues are linearly independent (see 8.13), the equation above implies that each *v*j equals

0 except possibly when j k. In particular, *v*1 0 and thus *v* u U. Because *v* U, we can conclude that *v* G.入k;T U /, completing the

2 2 j

D D D 2

proof.

As we know, an operator on a complex vector space may not have enough eigenvectors to form a basis of the domain. The next result shows that on a complex vector space there are enough generalized eigenvectors to do this.

**254** CHAPTER 8 Operators on Complex Vector Spaces

8.23 A basis of generalized eigenvectors

Suppose V is a complex vector space and T 2 *L*.V /. Then there is a basis

of V consisting of generalized eigenvectors of T.

Proof Choose a basis of each G.入j ;T / in 8.21. Put all these bases together to form a basis of V consisting of generalized eigenvectors of T.

### Multiplicity of an Eigenvalue

If V is a complex vector space and T .V /, then the decomposition of V provided by 8.21 can be a powerful tool. The dimensions of the subspaces involved in this decomposition are sufﬁciently important to get a name.

2 *L*

8.24 **Deﬁnition *multiplicity***

* Suppose T 2 *L*.V /. The ***multiplicity*** of an eigenvalue 入 of T

is deﬁned to be the dimension of the corresponding generalized eigenspace G.入; T /.

* In other words, the multiplicity of an eigenvalue 入 of T equals

dim null.T — 入I /dim V .

The second bullet point above is justiﬁed by 8.11.

8.25 **Example** Suppose T 2 *L*.**C**3/ is deﬁned by

T .z1; z2; z3/ D .6z1 C 3z2 C 4z3; 6z2 C 2z3; 7z3/:

The matrix of T (with respect to the standard basis) is

0 1

@ 6 3 4 A

0 6 2

0 0 7

:

The eigenvalues of T are 6 and 7, as follows from 5.32. You can verify that the generalized eigenspaces of T are as follows:

（ ） （ ）

G.6; T / D span .1; 0; 0/; .0; 1; 0/ and G.7; T / D span .10; 2; 1/ :

Thus the eigenvalue 6 has multiplicity 2 and the eigenvalue 7 has multiplicity 1. The direct sum **C**3 G.6; T / G.7; T / is the decomposition promised by

D ˚

8.21. A basis of **C**3 consisting of generalized eigenvectors of T , as promised

by 8.23, is

.1; 0; 0/; .0; 1; 0/; .10; 2; 1/:

SECTION 8.B Decomposition of an Operator **255**

In Example 8.25, the sum of the multiplicities of the eigenvalues of T equals 3, which is the dimension of the domain of T. The next result shows that this always happens on a complex vector space.

8.26 Sum of the multiplicities equals dim V

Suppose V is a complex vector space and T 2 *L*.V /. Then the sum of the

multiplicities of all the eigenvalues of T equals dim V.

Proof The desired result follows from 8.21 and the obvious formula for the dimension of a direct sum (see 3.78 or Exercise 16 in Section 2.C).

The terms ***algebraic multiplicity*** and ***geometric multiplicity*** are used in some books. In case you encounter this terminology, be aware that the algebraic multiplicity is the same as the multiplicity deﬁned here and the geometric multiplicity is the dimension of the corresponding eigenspace. In

other words, if T 2 *L*.V / and 入is an eigenvalue of T, then

algebraic multiplicity of 入 D dim null.T — 入I /dim V D dim G.入; T /;

geometric multiplicity of 入 D dim null.T — 入I / D dim E.入; T /:

Note that as deﬁned above, the algebraic multiplicity also has a geometric meaning as the dimension of a certain null space. The deﬁnition of multiplicity given here is cleaner than the traditional deﬁnition that involves determinants;

10.25 implies that these deﬁnitions are equivalent.

### Block Diagonal Matrices

To interpret our results in matrix form, we make the following deﬁnition, gener- alizing the notion of a diagonal matrix.

*Often we can understand a matrix better by thinking of it as composed of smaller matrices.*

If each matrix Aj in the deﬁnition

below is a 1-by-1 matrix, then we actually have a diagonal matrix.

8.27 **Deﬁnition *block diagonal matrix***

A ***block diagonal matrix*** is a square matrix of the form

B

A

1

0

@

: :

:

0

Am

1C

where A1;:::; Am are square matrices lying along the diagonal and all the other entries of the matrix equal 0.

A

;

**256** CHAPTER 8 Operators on Complex Vector Spaces

8.28 **Example** The 5-by-5 matrix

0 （ 4 ） 0 0 0 0 1

B 0 2 —3 ! 0 0 C

0 2

0 0

@

0

0

0 0

0 0

1 7

0 1

A

A D 0

B

! C

is a block diagonal matrix with

0B A1 0 1C

A

where

A D @

A2 ;

0 A3

A1 D （ 4 ） ; A2

2 3 D 1 7 ! :

Here the inner matrices in the 5-by-5 matrix above are blocked off to show how we can think of it as a block diagonal matrix.

!—D ; A

0 1

3

0 2

Note that in the next result we get many more zeros in the matrix of T

than are needed to make it upper triangular.

8.29 Block diagonal matrix with upper-triangular blocks

Suppose V is a complex vector space and T 2 *L*.V /. Let 入1;:::; 入m be

the distinct eigenvalues of T, with multiplicities d1;:::; dm. Then there is a basis of V with respect to which T has a block diagonal matrix of the

form

0 A

B@

1

0

: : :

1

0

Am

where each Aj is a dj -by-dj upper-triangular matrix of the form

CA ;

A D

j

B

入

j

@

: : :

\*

入j

C

A

:

0

SECTION 8.B Decomposition of an Operator **257**

Proof Each .T 入j I / G.入*j* ;T / is nilpotent [see 8.21(c)]. For each j , choose a basis of G.入j ;T /, which is a vector space with dimension dj , such that the matrix of .T — 入j I /jG.入*j* ;T / with respect to this basis is as in 8.19. Thus the matrix of T G.入*j* ;T /, which equals .T 入j I/ G.入*j* ;T / 入j I G.入*j* ;T /, with respect to this basis will look like the desired form shown above for Aj .

— j

j — j C j

Putting the bases of the G.入j ;T /’s together gives a basis of V [by 8.21(a)].

The matrix of T with respect to this basis has the desired form.

The 5-by-5 matrix in 8.28 is of the form promised by 8.29, with each of the blocks itself an upper-triangular matrix that is constant along the diagonal

of the block. If T is an operator on a 5-dimensional vector space whose matrix is as in 8.28, then the eigenvalues of T are 4; 2; 1 (as follows from 5.32), with multiplicities 1, 2, 2.

8.30 **Example** Suppose T 2 *L*.**C**3/ is deﬁned by

T .z1; z2; z3/ D .6z1 C 3z2 C 4z3; 6z2 C 2z3; 7z3/:

The matrix of T (with respect to the standard basis) is

@0 1A ;

|  |  |  |
| --- | --- | --- |
| 6 | 3 | 4 |
| 0 | 6 | 2 |
| 0 | 0 | 7 |

which is an upper-triangular matrix but is not of the form promised by 8.29.

As we saw in Example 8.25, the eigenvalues of T are 6 and 7 and the corresponding generalized eigenspaces are

（ ） （ ）

G.6; T / D span .1; 0; 0/; .0; 1; 0/ and G.7; T / D span .10; 2; 1/ :

We also saw that a basis of **C**3 consisting of generalized eigenvectors of T is

.1; 0; 0/; .0; 1; 0/; .10; 2; 1/:

The matrix of T with respect to this basis is

0@ 6 3 0 1A

（ ）

0 6

0

;

0 0 7

which is a matrix of the block diagonal form promised by 8.29.

When we discuss the Jordan Form in Section 8.D, we will see that we can ﬁnd a basis with respect to which an operator T has a matrix with even more 0’s than promised by 8.29. However, 8.29 and its equivalent companion 8.21

are already quite powerful. For example, in the next subsection we will use

8.21 to show that every invertible operator on a complex vector space has a square root.

**258** CHAPTER 8 Operators on Complex Vector Spaces

### Square Roots

Recall that a square root of an operator T .V / is an operator R .V /

2 *L* 2 *L*

such that R2 T (see 7.33). Every complex number has a square root, but

D

not every operator on a complex vector space has a square root. For example,

the operator on **C**3 in Exercise 6 in Section 8.A has no square root. The noninvertibility of that operator is no accident, as we will soon see. We begin by showing that the identity plus any nilpotent operator has a square root.

8.31 Identity plus nilpotent has a square root

Suppose N 2 *L*.V / is nilpotent. Then I C N has a square root.

Proof Consider the Taylor series for the function p1 C x:

**8.32** p1 C x D 1 C a1x C a2x2 C ... :

We will not ﬁnd an explicit formula for the coefﬁcients or worry about whether the inﬁnite sum converges be- cause we will use this equation only as motivation.

*Because* a1 D 1=2*, the formula*

*above shows that* 1 C x=2 *is a*

p

*is small.*

*good estimate for* 1 C x *when* x

Because N is nilpotent, N m 0 for some positive integer m. In 8.32, suppose we replace x with N and 1 with I. Then the inﬁnite sum on the right side becomes a ﬁnite sum (because N j D 0 for all j 乏 m). In other words, we guess that there is a square root of I C N of the form

D

I C a1N C a2N 2 C ... C am 1N m 1:

Having made this guess, we can try to choose a1; a2;:::; am 1 such that the operator above has its square equal to I C N. Now

.ICa1N C a2N 2 C a3N 3 C ... C am 1N m 1/2

D I C 2a1N C .2a2 C a1 /N C .2a3 C 2a1a2/N C ...

2

2

3

C .2am 1 C terms involving a1;:::; am 2/N m 1:

We want the right side of the equation above to equal I N. Hence choose a1 such that 2a1 1 (thus a1 1=2). Next, choose a2 such that 2a2 a12 0 (thus a2 1=8). Then choose a3 such that the coefﬁcient of N 3 on the right side of the equation above equals 0 (thus a3 1=16). Continue in this fashion for j 4; :::;m 1, at each step solving for aj so that the coefﬁcient of N j on the right side of the equation above equals 0. Actually we do not care about the explicit formula for the aj ’s. We need only know that some choice of the aj ’s gives a square root of I C N.

D —

D

D —

D D C D

C

SECTION 8.B Decomposition of an Operator **259**

The previous lemma is valid on real and complex vector spaces. However, the next result holds only on complex vector spaces. For example, the operator

of multiplication by 1 on the 1-dimensional real vector space **R** has no square

—

root.

8.33 Over **C**, invertible operators have square roots

Suppose V is a complex vector space and T 2 *L*.V / is invertible. Then

T has a square root.

Proof Let 入1;:::; 入m be the distinct eigenvalues of T. For each j , there ex- ists a nilpotent operator Nj G.入j ;T / such that T G.入*j* ;T / 入j I Nj [see 8.21(c)]. Because T is invertible, none of the 入j ’s equals 0, so we can

write

（ ）2 *L* j D C

T jG.入*j* ;T /

D 入j

I C Nj

for each j . Clearly Nj =入j is nilpotent, and so I Nj =入j has a square root (by 8.31). Multiplying a square root of the complex number 入j by a square root of I C Nj =入j , we obtain a square root Rj of T jG.入*j* ;T /.

入j

C

A typical vector *v* 2 V can be written uniquely in the form

*v* D u1 C ... C um;

where each uj is in G.入j ;T / (see 8.21). Using this decomposition, deﬁne an operator R 2 *L*.V / by

R*v* D R1u1 C ... C Rmum:

You should verify that this operator R is a square root of T, completing the proof.

By imitating the techniques in this section, you should be able to prove that if V is a complex vector space and T .V / is invertible, then T has a kth root for every positive integer k.

2 *L*

EXERCISES 8.B

2 *L*

1. Suppose V is a complex vector space, N .V /, and 0 is the only eigenvalue of N. Prove that N is nilpotent.
2. Give an example of an operator T on a ﬁnite-dimensional real vector space such that 0 is the only eigenvalue of T but T is not nilpotent.

**260** CHAPTER 8 Operators on Complex Vector Spaces

1. Suppose T .V /. Suppose S .V / is invertible. Prove that T and

2 *L* 2 *L*

S 1TS have the same eigenvalues with the same multiplicities.

1. Suppose V is an n-dimensional complex vector space and T is an oper- ator on V such that null T n 2 null T n 1. Prove that T has at most two distinct eigenvalues.

¤

1. Suppose V is a complex vector space and T .V /. Prove that V has a basis consisting of eigenvectors of T if and only if every generalized eigenvector of T is an eigenvector of T.

2 *L*

[*For* **F** D **C***, the exercise above adds an equivalence to the list in 5.41.*]

1. Deﬁne N 2 *L*.**F**5/ by

N.x1; x2; x3; x4; x5/ D .2x2; 3x3; —x4; 4x5; 0/:

Find a square root of I C N.

1. Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.
2. Suppose T 2 *L*.V / and 3 and 8 are eigenvalues of T. Let n D dim V. Prove that V D .null T n 2/ ˚ .range T n 2/.
3. Suppose A and B are block diagonal matrices of the form

0B A1 0 1C 0B B1 0 1C

A D @

: : :

0 Am

A ; B D @

: : :

0 Bm

A ;

where Aj has the same size as Bj for j 1; :::; m. Show that AB is a

D

block diagonal matrix of the form

0B A1B1 0 1C

AB D @

: : :

0 AmBm

A :

1. Suppose **F** D **C** and T 2 *L*.V /. Prove that there exist D; N 2 *L*.V / such that T D D C N, the operator D is diagonalizable, N is nilpotent, and DN D ND.
2. Suppose T .V / and 入 **F**. Prove that for every basis of V with respect to which T has an upper-triangular matrix, the number of times that 入appears on the diagonal of the matrix of T equals the multiplicity of 入as an eigenvalue of T.

2 *L* 2

SECTION 8.C Characteristic and Minimal Polynomials **261**

2 *L*

## *Characteristic and Minimal Polynomials*

8.C

### The Cayley–Hamilton Theorem

The next deﬁnition associates a polynomial with each operator on V if **F**D **C**. For **F** D **R**, the corresponding deﬁnition will be given in the next chapter.

8.34 **Deﬁnition *characteristic polynomial***

Suppose V is a complex vector space and T 2 *L*.V /. Let 入1;:::; 入m

denote the distinct eigenvalues of T, with multiplicities d1;:::; dm. The

polynomial

.z — 入1/ z — 入m/

is called the ***characteristic polynomial*** of T.

d*1*

d*m*

8.35 **Example** Suppose T .**C**3/ is deﬁned as in Example 8.25. Be- cause the eigenvalues of T are 6, with multiplicity 2, and 7, with multiplicity 1, we see that the characteristic polynomial of T is .z — 6/2.z — 7/.

8.36 Degree and zeros of characteristic polynomial

Suppose V is a complex vector space and T 2 *L*.V /. Then

1. the characteristic polynomial of T has degree dim V ;
2. the zeros of the characteristic polynomial of T are the eigenvalues of T.

Proof Clearly part (a) follows from 8.26 and part (b) follows from the deﬁni- tion of the characteristic polynomial.

Most texts deﬁne the characteristic polynomial using determinants (the two deﬁnitions are equivalent by 10.25). The approach taken here, which is considerably simpler, leads to the following easy proof of the Cayley– Hamilton Theorem. In the next chapter, we will see that this result also holds on real vector spaces (see 9.24).

8.37 Cayley–Hamilton Theorem

Suppose V is a complex vector space and T 2 *L*.V /. Let q denote the

characteristic polynomial of T. Then q.T / D 0.

**262** CHAPTER 8 Operators on Complex Vector Spaces

Proof Let 入1;:::; 入m be the distinct eigenvalues of the operator T, and let d1;:::; dm be the dimensions of the

*English mathematician Arthur Cayley* (*1821–1895*) *published three math papers before complet- ing his undergraduate degree in 1842. Irish mathematician William Rowan Hamilton* (*1805–1865*) *was made a professor in 1827 when he was 22 years old and still an undergraduate!*

corresponding generalized eigenspaces G.入1;T /; :::; G.入m;T /. For each j 2 f1; :::; mg, we know that

.T — 入j I/jG.入*j* ;T / is nilpotent. Thus

we have .T — 入j I /d*j* jG.入 ;T / D 0 (by

*j*

8.18).

Every vector in V is a sum of vectors in G.入1;T /; :::; G.入m;T / (by 8.21). Thus to prove that q.T / 0, we need only show that q.T / G.入*j* ;T / 0 for each j .

D j D

Thus ﬁx j 2 f1; :::; mg. We have

q.T / D .T — 入1I /d*1* ... .T — 入mI /d*m* :

The operators on the right side of the equation above all commute, so we can move the factor .T — 入j I/d*j* to be the last term in the expression on the right.

Because .T — 入j I /d*j* jG.入 ;T / D 0, we conclude that q.T /jG.入 ;T / D 0, as

*j*

*j*

desired.

### The Minimal Polynomial

In this subsection we introduce another important polynomial associated with each operator. We begin with the following deﬁnition.

C C

8.38 **Deﬁnition *monic polynomial***

A ***monic polynomial*** is a polynomial whose highest-degree coefﬁcient equals 1.

8.39 **Example** The polynomial 2 9z2 z7 is a monic polynomial of degree 7.

8.40 Minimal polynomial

Suppose T 2 *L*.V /. Then there is a unique monic polynomial p of

smallest degree such that p.T / D 0.

SECTION 8.C Characteristic and Minimal Polynomials **263**

Proof Let n D dim V. Then the list

I; T; T 2;:::;T n*2*

is not linearly independent in .V /, because the vector space .V / has dimension n2 (see 3.61) and we have a list of length n2 1. Let m be the smallest positive integer such that the list

C

*L L*

**8.41** I; T; T 2;:::;T m

is linearly dependent. The Linear Dependence Lemma (2.21) implies that one of the operators in the list above is a linear combination of the previous ones.

Because m was chosen to be the smallest positive integer such that the list above is linearly dependent, we conclude that T m is a linear combination of I; T; T 2;:::;T m 1. Thus there exist scalars a0; a1; a2;:::; am 1 **F** such

2

that

**8.42** a0I C a1T C a2T 2 C C am 1T m 1 C T m D 0:

Deﬁne a monic polynomial p 2 *P*.**F**/ by

p.z/ D a0 C a1z C a2z2 C C am 1zm 1 C zm:

Then 8.42 implies that p.T / 0.

D

To prove the uniqueness part of the result, note that the choice of m implies that no monic polynomial q 2 *P*.**F**/ with degree smaller than m can satisfy

q.T / D 0. Suppose q 2 *P*.**F**/ is a monic polynomial with degree m and

q.T / D 0. Then .p — q/.T / D 0 and deg.p — q/< m. The choice of m now implies that q D p, completing the proof.

The last result justiﬁes the following deﬁnition.

8.43 **Deﬁnition *minimal polynomial***

Suppose T 2 *L*.V /. Then the ***minimal polynomial*** of T is the unique

monic polynomial p of smallest degree such that p.T / D 0.

The proof of the last result shows that the degree of the minimal polynomial of each operator on V is at most .dim V /2. The Cayley–Hamilton Theorem (8.37) tells us that if V is a complex vector space, then the minimal polynomial of each operator on V has degree at most dim V. This remarkable improvement

also holds on real vector spaces, as we will see in the next chapter.

**264** CHAPTER 8 Operators on Complex Vector Spaces

Suppose you are given the matrix (with respect to some basis) of an operator T .V /. You could program a computer to ﬁnd the minimal polynomial of T as follows: Consider the system of linear equations

2 *L*

* 1. a0*M*.I/ C a1*M*.T / C ... C am 1*M*.T /m 1 D —*M*.T /m

for successive values of m 1; 2; ::: until this system of equations has a solu- tion a0; a1; a2;:::; am 1. The scalars

*Think of this as a system of*

.dim V /2 *linear equations in* m

*variables* a0; a1;:::; am 1*.*

D

a0; a1; a2;:::; am 1;1 will then be the

coefﬁcients of the minimal polynomial of T. All this can be computed using a

familiar and fast (for a computer) process such as Gaussian elimination.

0B 1C

* 1. **Example** Let T be the operator on **C**5 whose matrix (with respect to the standard basis) is

0 0 0 0 —3

1 0 0 0 6

0 1 0 0 0

B@

:

CA

0 0 1 0 0

0 0 0 1 0

Find the minimal polynomial of T.

Solution Because of the large number of 0’s in this matrix, Gaussian elim- ination is not needed here. Simply compute powers of .T /, and then you will notice that there is clearly no solution to 8.44 until m 5. Do the computations and you will see that the minimal polynomial of T equals z5 — 6z C 3.

D

*M*

The next result completely characterizes the polynomials that when applied to an operator give the 0 operator.

8.46 q.T / D 0 implies q is a multiple of the minimal polynomial

Suppose T 2 *L*.V / and q 2 *P*.**F**/. Then q.T / D 0 if and only if q is a

polynomial multiple of the minimal polynomial of T.

Proof Let p denote the minimal polynomial of T.

First we prove the easy direction. Suppose q is a polynomial multiple of p.

Thus there exists a polynomial s 2 *P*.**F**/ such that q D ps. We have

q.T / D p.T /s.T/ D 0 s.T / D 0;

as desired.

SECTION 8.C Characteristic and Minimal Polynomials **265**

To prove the other direction, now suppose q.T / 0. By the Division Algorithm for Polynomials (4.8), there exist polynomials s; r .**F**/ such that

2 *P*

D

**8.47** q D ps C r

and deg r < deg p. We have

0 D q.T / D p.T /s.T/ C r.T / D r.T /:

The equation above implies that r 0 (otherwise, dividing r by its highest- degree coefﬁcient would produce a monic polynomial that when applied to T gives 0; this polynomial would have a smaller degree than the minimal polynomial, which would be a contradiction). Thus 8.47 becomes the equation q D ps. Hence q is a polynomial multiple of p, as desired.

D

The next result is stated only for complex vector spaces, because we have not yet deﬁned the characteristic polynomial when **F R**. However, the result also holds for real vector spaces, as we will see in the next chapter.

D

8.48 Characteristic polynomial is a multiple of minimal polynomial

is a polynomial multiple of the minimal polynomial of T.

Suppose **F** D **C** and T 2 *L*.V /. Then the characteristic polynomial of T

Proof The desired result follows immediately from the Cayley–Hamilton Theorem (8.37) and 8.46.

We know (at least when **F C**) that the zeros of the characteristic polynomial of T are the eigenvalues of T (see 8.36). Now we show that the minimal polynomial has the same zeros (although the multiplicities of these

D

zeros may differ).

8.49 Eigenvalues are the zeros of the minimal polynomial

Let T 2 *L*.V /. Then the zeros of the minimal polynomial of T are

precisely the eigenvalues of T.

Proof Let

p.z/ D a0 C a1z C a2z2 C ... C am 1zm 1 C zm

be the minimal polynomial of T.

**266** CHAPTER 8 Operators on Complex Vector Spaces

First suppose 入 2 **F** is a zero of p. Then p can be written in the form

p.z/ D .z — 入/q.z/;

where q is a monic polynomial with coefﬁcients in **F** (see 4.11). Because

p.T / D 0, we have

0 D .T — 入I/.q.T /*v*/

for all *v* V. Because the degree of q is less than the degree of the minimal polynomial p, there exists at least one vector *v* V such that q.T /*v* 0. The equation above thus implies that 入is an eigenvalue of T, as desired.

2 ¤

2

To prove the other direction, now suppose 入 **F** is an eigenvalue of T. Thus there exists *v* V with *v* 0 such that T *v* 入 *v*. Repeated applications of T to both sides of this equation show that T j *v* 入j *v* for every nonnegative integer j . Thus

D

2 ¤ D

2

0 D p.T /*v* D .a0I C a1T C a2T 2 C ... C am 1T m 1 C T m/*v*

D .a0 C a1入C a2入2 C ... C am 1入m 1 C 入m/*v*

D p.入/*v*:

Because *v* ¤ 0, the equation above implies that p.入/ D 0, as desired.

The next three examples show how our results can be useful in ﬁnding minimal polynomials and in understanding why eigenvalues of some operators cannot be exactly computed.

* 1. **Example** Find the minimal polynomial of the operator T .**C**3/

in Example 8.30.

2 *L*

Solution In Example 8.30 we noted that the eigenvalues of T are 6 and 7. Thus by 8.49, the minimal polynomial of T is a polynomial multiple of

.z 6/.z 7/.

— —

In Example 8.35, we saw that the characteristic polynomial of T is

.z 6/2.z 7/. Thus by 8.48 and the paragraph above, the minimal polyno- mial of T is either .z 6/.z 7/ or .z 6/2.z 7/. A simple computation shows that

— — — —

— —

.T — 6I/.T — 7I/ ¤ 0:

Thus the minimal polynomial of T is .z — 6/2.z — 7/.

SECTION 8.C Characteristic and Minimal Polynomials **267**

* 1. **Example** Find the minimal polynomial of the operator T 2 *L*.**C**3/

deﬁned by T .z1; z2; z3/ D .6z1; 6z2; 7z3/.

Solution It is easy to see that for this operator T, the eigenvalues of T are 6

and 7, and the characteristic polynomial of T is .z 6/2.z 7/.

— —

Thus as in the previous example, the minimal polynomial of T is ei- ther .z — 6/.z — 7/ or .z — 6/2.z — 7/. A simple computation shows that

.T — 6I/.T — 7I/ D 0. Thus the minimal polynomial of T is .z — 6/.z — 7/.

* 1. **Example** What are the eigenvalues of the operator in Example 8.45?

Solution From 8.49 and the solution to Example 8.45, we see that the eigenvalues of T equal the solutions to the equation

z5 — 6z C 3 D 0:

Unfortunately, no solution to this equation can be computed using rational numbers, roots of rational numbers, and the usual rules of arithmetic (a proof of this would take us considerably beyond linear algebra). Thus we cannot ﬁnd

an exact expression for any eigenvalue of T in any familiar form, although

numeric techniques can give good approximations for the eigenvalues of

T. The numeric techniques, which we will not discuss here, show that the eigenvalues for this particular operator are approximately

—1:67; 0:51; 1:40; —0:12 C 1:59i; —0:12 — 1:59i:

The nonreal eigenvalues occur as a pair, with each the complex conjugate of the other, as expected for a polynomial with real coefﬁcients (see 4.15).

EXERCISES 8.C

1. Suppose T 2 *L*.**C**4/ is such that the eigenvalues of T are 3, 5, 8. Prove that .T — 3I/2.T — 5I/2.T — 8I/2 D 0.
2. Suppose V is a complex vector space. Suppose T .V / is such that 5 and 6 are eigenvalues of T and that T has no other eigenvalues. Prove that .T — 5I/n 1.T — 6I/n 1 D 0, where n D dim V.

2 *L*

1. Give an example of an operator on **C**4 whose characteristic polynomial equals .z — 7/2.z — 8/2.

**268** CHAPTER 8 Operators on Complex Vector Spaces

1. Give an example of an operator on **C**4 whose characteristic polyno- mial equals .z — 1/.z — 5/3 and whose minimal polynomial equals

.z — 1/.z — 5/2.

1. Give an example of an operator on **C**4 whose characteristic and minimal polynomials both equal z.z — 1/2.z — 3/.
2. Give an example of an operator on **C**4 whose characteristic polyno- mial equals z.z — 1/2.z — 3/ and whose minimal polynomial equals z.z — 1/.z — 3/.
3. Suppose V is a complex vector space. Suppose T 2 *L*.V / is such that P 2 D P . Prove that the characteristic polynomial of P is zm.z — 1/n, where m D dim null P and n D dim range P .
4. Suppose T .V /. Prove that T is invertible if and only if the constant term in the minimal polynomial of T is nonzero.

2 *L*

1. Suppose T .V / has minimal polynomial 4 5z 6z2 7z3 2z4 z5. Find the minimal polynomial of T 1.

2 *L* C — — C C

1. Suppose V is a complex vector space and T .V / is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T 1. Prove that

2 *L*

1 1

q.z/ D zdim V p

for all nonzero z 2 **C**.

p.0/ z

1. Suppose T 2 *L*.V / is invertible. Prove that there exists a polynomial

p 2 *P*.**F**/ such that T 1 D p.T /.

1. Suppose V is a complex vector space and T .V /. Prove that V has a basis consisting of eigenvectors of T if and only if the minimal polynomial of T has no repeated zeros.

2 *L*

[*For complex vector spaces, the exercise above adds another equivalence to the list given by 5.41.*]

1. Suppose V is an inner product space and T .V / is normal. Prove that the minimal polynomial of T has no repeated zeros.

2 *L*

1. Suppose V is a complex inner product space and S .V / is an isometry. Prove that the constant term in the characteristic polynomial of S has absolute value 1.

2 *L*

SECTION 8.C Characteristic and Minimal Polynomials **269**

1. Suppose T 2 *L*.V / and *v* 2 V.
   1. Prove that there exists a unique monic polynomial p of smallest degree such that p.T /*v* D 0.
   2. Prove that p divides the minimal polynomial of T.
2. Suppose V is an inner product space and T 2 *L*.V /. Suppose

a0 C a1z C a2z2 C C am 1zm 1 C zm

is the minimal polynomial of T. Prove that

a0 C a1z C a2z2 C C am 1zm 1 C zm

is the minimal polynomial of T \*.

1. Suppose **F C** and T .V /. Suppose the minimal polynomial of T has degree dim V. Prove that the characteristic polynomial of T equals the minimal polynomial of T.

D 2 *L*

1. Suppose a0;:::; an 1 **C**. Find the minimal and characteristic polyno- mials of the operator on **C**n whose matrix (with respect to the standard

2

basis) is

0B

1

0 —a0

C

1 0 a

— 1

B : : : C

:

1

: : :

B

CA

@

—a2

:

0 —an 2

1 —an 1

[*The exercise above shows that every monic polynomial is the character- istic polynomial of some operator.*]

1. Suppose V is a complex vector space and T .V /. Suppose that with respect to some basis of V the matrix of T is upper triangular, with 入1;:::; 入n on the diagonal of this matrix. Prove that the characteristic polynomial of T is .z — 入1/ z — 入n/.

2 *L*

1. Suppose V is a complex vector space and V1;:::; Vm are nonzero sub- spaces of V such that V V1 Vm. Suppose T .V / and each Vj is invariant under T. For each j , let pj denote the characteristic polynomial of T jV*j* . Prove that the characteristic polynomial of T equals p1 pm.

D ˚ ... ˚ 2 *L*

**270** CHAPTER 8 Operators on Complex Vector Spaces

## *Jordan Form*

8.D

We know that if V is a complex vector space, then for every T .V / there is a basis of V with respect to which T has a nice upper-triangular matrix (see 8.29). In this section we will see that we can do even better—there is a basis

2 *L*

of V with respect to which the matrix of T contains 0’s everywhere except

possibly on the diagonal and the line directly above the diagonal.

We begin by looking at two examples of nilpotent operators.

* 1. **Example** Let N 2 *L*.**F**4/ be the nilpotent operator deﬁned by

N.z1; z2; z3; z4/ D .0; z1; z2; z3/:

If *v* .1; 0; 0; 0/, then N 3*v*;N 2*v*;N *v*; *v* is a basis of **F**4. The matrix of N

D

with respect to this basis is

0 0

B

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1 | 0 |  |
| 0 | 0 | 1 | 0 |
|  | 0 | 0 |  |
| 0 | 0 | 0 | 0 |

B@ 0

0 1

1 CA

C

The next example of a nilpotent operator has more complicated behavior than the example above.

* 1. **Example** Let N 2 *L*.**F**6/ be the nilpotent operator deﬁned by

N.z1; z2; z3; z4; z5; z6/ D .0; z1; z2; 0; z4; 0/:

Unlike the nice behavior of the nilpotent operator of the previous exam- ple, for this nilpotent operator there does not exist a vector *v* **F**6 such that N 5*v*;N 4*v*;N 3*v*;N 2*v*;N *v*; *v* is a basis of **F**6. However, if we take

2

D D D

*v*1 .1; 0; 0; 0; 0; 0/, *v*2 .0; 0; 0; 1; 0; 0/, and *v*3 .0; 0; 0; 0; 0; 1/, then

N 2*v*1;N *v*1; *v*1;N *v*2; *v*2; *v*3 is a basis of **F**6. The matrix of N with respect to

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0 | 1 |  | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 |

this basis is

0B 0@

B

0 1

0 0 0

A

0 0

0 1C

C

:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |

@B （ 0 ） CA

Here the inner matrices are blocked off to show that we can think of the 6-by-6 matrix above as a block diagonal matrix consisting of a 3-by-3 block with 1’s on the line above the diagonal and 0’s elsewhere, a 2-by-2 block with 1 above the diagonal and 0’s elsewhere, and a 1-by-1 block containing 0.

SECTION 8.D Jordan Form **271**

Our next result shows that every nilpotent operator N .V / behaves similarly to the previous example. Speciﬁcally, there is a ﬁnite collection of vectors *v*1;:::; *v*n V such that there is a basis of V consisting of the vectors

2

2 *L*

of the form N k*v*j , as j varies from 1 to n and k varies (in reverse order) from

0 to the largest nonnegative integer mj such that N m*j v*j 0. For the matrix

¤

interpretation of the next result, see the ﬁrst part of the proof of 8.60.

8.55 Basis corresponding to a nilpotent operator

and nonnegative integers m1;:::; mn such that

Suppose N 2 *L*.V / is nilpotent. Then there exist vectors *v*1;:::; *v*n 2 V

1. N m*1 v*1;:::;N *v*1; *v*1;:::;N m*nv*n;:::;N *v*n; *v*n is a basis of V ;
2. N m*1*C1*v*1 D . . .D N m*n*C1*v*n D 0.

Proof We will prove this result by induction on dim V. To get started, note that the desired result obviously holds if dim V 1 (in that case, the only nilpotent operator is the 0 operator, so take *v*1 to be any nonzero vector and m1 0). Now assume that dim V > 1 and that the desired result holds on all

vector spaces of smaller dimension.

D

D

Because N is nilpotent, N is not injective. Thus N is not surjective (by 3.69) and hence range N is a subspace of V that has a smaller dimension than V. Thus we can apply our induction hypothesis to the restriction operator N range N .range N /. [We can ignore the trivial case range N 0 , because in that case N is the 0 operator and we can choose *v*1;:::; *v*n to be any basis of V and m1 D . . .D mn D 0 to get the desired result.]

j 2 *L* D f g

By our induction hypothesis applied to N jrange N , there exist vectors

*v*1;:::; *v*n 2 range N and nonnegative integers m1;:::; mn such that

**8.56** N m*1 v*1;:::;N *v*1; *v*1;:::;N m*nv*n;:::;N *v*n; *v*n

is a basis of range N and

N m*1*C1*v*1 D . . .D N m*n*C1*v*n D 0:

Because each *v*j is in range N, for each j there exists uj V such that *v*j N uj . Thus N kC1uj N k*v*j for each j and each nonnegative integer k. We now claim that

D D

2

**8.57** N m*1*C1u1;:::;Nu1; u1;:::;N m*n*C1un;:::;Nun; un

R:: 100e CHAPTER 8 Operators on Complex Vector Spaces

is a linearly independent list of vectors in V. To verify this claim, suppose that some linear combination of 8.57 equals 0. Applying N to that linear combination, we get a linear combination of 8.56 equal to 0. However, the

list 8.56 is linearly independent, and hence all the coefﬁcients in our original linear combination of 8.57 equal 0 except possibly the coefﬁcients of the vectors

N m*1*C1u1;:::;N m*n*C1un;

which equal the vectors

N m*1 v*1;:::;N m*nv*n:

Again using the linear independence of the list 8.56, we conclude that those coefﬁcients also equal 0, completing our proof that the list 8.57 is linearly independent.

Now extend 8.57 to a basis

**8.58** N m*1*C1u1;:::;Nu1; u1;:::;N m*n*C1un;:::;Nun; un; *w*1;:::; *w*p

of V (which is possible by 2.33). Each N *w*j is in range N and hence is in the span of 8.56. Each vector in the list 8.56 equals N applied to some vector in the list 8.57. Thus there exists xj in the span of 8.57 such that N *w*j N xj .

D

Now let

unCj D *w*j — xj :

Then N unCj D 0. Furthermore,

N m*1*C1u1;:::;Nu1; u1;:::;N m*n*C1un;:::;Nun; un; unC1;:::; unCp

spans V because its span contains each xj and each unCj and hence each *w*j

(and because 8.58 spans V ).

Thus the spanning list above is a basis of V because it has the same length

as the basis 8.58 (where we have used 2.42). This basis has the required form,

completing the proof.

In the next deﬁnition, the diagonal of each Aj is ﬁlled with some eigenvalue 入j of T, the line directly above the di- agonal of Aj is ﬁlled with 1’s, and all

*French mathematician Camille Jor- dan* (*1838–1922*) *ﬁrst published a proof of 8.60 in 1870.*

other entries in Aj are 0 (to understand why each 入j is an eigenvalue of T, see 5.32). The 入j ’s need not be distinct. Also, Aj may be a 1-by-1 matrix

.入j / containing just an eigenvalue of T.

SECTION 8.D Jordan Form **273**

8.59 **Deﬁnition *Jordan basis***

Suppose T 2 *L*.V /. A basis of V is called a ***Jordan basis*** for T if with

respect to this basis T has a block diagonal matrix

B

A

1

0

@

: :

:

0

Ap

1C

where each Aj is an upper-triangular matrix of the form

A

;

Aj D B@

0B 入j

1

: : :

0

: : :

: :

:

1C

0

1

入j

:

CA

8.60 Jordan Form

Suppose V is a complex vector space. If T 2 *L*.V /, then there is a basis

of V that is a Jordan basis for T.

Proof First consider a nilpotent operator N .V / and the vectors *v*1;:::; *v*n V given by 8.55. For each j , note that N sends the ﬁrst vector in the list N m*j v*j ;:::;N *v*j ; *v*j to 0 and that N sends each vector in this list

2

2 *L*

other than the ﬁrst vector to the previous vector. In other words, 8.55 gives a basis of V with respect to which N has a block diagonal matrix, where each matrix on the diagonal has the form

0B 0 1 0 1C

:

B

: :

:

: :

: :

:

C

: 1

A

@

0 0

Thus the desired result holds for nilpotent operators.

Now suppose T .V /. Let 入1;:::; 入m be the distinct eigenvalues of T. We have the generalized eigenspace decomposition

2 *L*

V D G.入1;T / ˚ ... ˚ G.入m;T /;

where each .T —入j I /jG.入*j* ;T / is nilpotent (see 8.21). Thus some basis of each G.入j ;T / is a Jordan basis for .T 入j I/ G.入*m*;T / (see previous paragraph). Put these bases together to get a basis of V that is a Jordan basis for T.

— j

**274** CHAPTER 8 Operators on Complex Vector Spaces

EXERCISES 8.D

1. Find the characteristic polynomial and the minimal polynomial of the operator N in Example 8.53.
2. Find the characteristic polynomial and the minimal polynomial of the operator N in Example 8.54.
3. Suppose N .V / is nilpotent. Prove that the minimal polynomial of N is zmC1, where m is the length of the longest consecutive string of 1’s that appears on the line directly above the diagonal in the matrix of N with respect to any Jordan basis for N.

2 *L*

1. Suppose T .V / and *v*1;:::; *v*n is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis *v*n;:::; *v*1 obtained by reversing the order of the *v*’s.

2 *L*

1. Suppose T .V / and *v*1;:::; *v*n is a basis of V that is a Jordan basis for T. Describe the matrix of T 2 with respect to this basis.

2 *L*

1. Suppose N .V / is nilpotent and *v*1;:::; *v*n and m1;:::; mn are as in 8.55. Prove that N m*1 v*1;:::;N m*nv*n is a basis of null N.

2 *L*

[*The exercise above implies that* n*, which equals* dim null N*, depends only on* N *and not on the speciﬁc Jordan basis chosen for* N*.*]

1. Suppose p; q .**C**/ are monic polynomials with the same zeros and q is a polynomial multiple of p. Prove that there exists T .Cdeg q/ such that the characteristic polynomial of T is q and the minimal polynomial of T is p.

2 *L*

2 *P*

1. Suppose V is a complex vector space and T .V /. Prove that there does not exist a direct sum decomposition of V into two proper subspaces invariant under T if and only if the minimal polynomial of T is of the form .z — 入/dim V for some 入 2 **C**.

2 *L*

# *Operators on Real Vector Spaces*



*Euclid explaining geometry (from* The School of Athens*, painted by Raphael around 1510).*

9

CHAPTER

In the last chapter we learned about the structure of an operator on a ﬁnite- dimensional complex vector space. In this chapter, we will use our results about operators on complex vector spaces to learn about operators on real vector spaces.

Our assumptions for this chapter are as follows:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a ﬁnite-dimensional nonzero vector space over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

complexiﬁcation of a real vector space complexiﬁcation of an operator on a real vector space

operators on ﬁnite-dimensional real vector spaces have an eigenvalue or a 2-dimensional invariant subspace

characteristic polynomial and the Cayley–Hamilton Theorem description of normal operators on a real inner product space description of isometries on a real inner product space

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

275

DOI 10.1007/978-3-319-11080-6 9

**276** CHAPTER 9 Operators on Real Vector Spaces

## *Complexiﬁcation*

9.A

### Complexiﬁcation of a Vector Space

As we will soon see, a real vector space V can be embedded, in a natural way, in a complex vector space called the complexiﬁcation of V. Each operator on V can be extended to an operator on the complexiﬁcation of V. Our

results about operators on complex vector spaces can then be translated to information about operators on real vector spaces.

We begin by deﬁning the complexiﬁcation of a real vector space.

* 1. **Deﬁnition *complexiﬁcation of*** V***,*** V**C**

Suppose V is a real vector space.

* + - The ***complexiﬁcation*** of V, denoted V**C**, equals V x V. An element of V**C** is an ordered pair .u; *v*/, where u; *v* 2 V, but we will write this as u C i*v*.
    - Addition on V**C** is deﬁned by

.u1 C i*v*1/ C .u2 C i*v*2/ D .u1 C u2/ C i.*v*1 C *v*2/

for u1; *v*1; u2; *v*2 2 V.

* + - Complex scalar multiplication on V**C** is deﬁned by

.a C bi/.u C i*v*/ D .au — b*v*/ C i.a*v* C bu/

for a; b 2 **R** and u; *v* 2 V.

Motivation for the deﬁnition above of complex scalar multiplication comes from usual algebraic properties and the identity i2 1. If you remember the motivation, then you do not need to memorize the deﬁnition above.

D —

We think of V as a subset of V**C** by identifying u 2 V with u C i0. The construction of V**C** from V can then be thought of as generalizing the construction of **C**n from **R**n.

9.3 V**C** is a complex vector space.

Suppose V is a real vector space. Then with the deﬁnitions of addition and scalar multiplication as above, V**C** is a complex vector space.

The proof of the result above is left as an exercise for the reader. Note that the additive identity of V**C** is 0 C i0, which we write as just 0.

SECTION 9.A Complexiﬁcation **277**

Probably everything that you think should work concerning complexiﬁca- tion does work, usually with a straightforward veriﬁcation, as illustrated by the next result.

9.4 Basis of V is basis of V**C**

Suppose V is a real vector space.

1. If *v*1;:::; *v*n is a basis of V (as a real vector space), then *v*1;:::; *v*n

is a basis of V**C** (as a complex vector space).

1. The dimension of V**C** (as a complex vector space) equals the dimen- sion of V (as a real vector space).

Proof To prove (a), suppose *v*1;:::; *v*n is a basis of the real vector space V. Then span.*v*1;:::; *v*n/ in the complex vector space V**C** contains all the vectors *v*1;:::; *v*n;i*v*1;:::;i*v*n. Thus *v*1;:::; *v*n spans the complex vector space V**C**.

To show that *v*1;:::; *v*n is linearly independent in the complex vector space V**C**, suppose 入1;:::; 入n 2 **C** and

入1*v*1 C ... C 入n*v*n D 0:

Then the equation above and our deﬁnitions imply that

.Re 入1/*v*1 C ... C .Re 入n/*v*n D 0 and .Im 入1/*v*1 C. . . C .Im 入n/*v*n D 0:

Because *v*1;:::; *v*n is linearly independent in V, the equations above imply Re 入1 D ... D Re 入n D 0 and Im 入1 D ... D Imn D 0. Thus we have 入 1 入 n 0. Hence *v*1;:::; *v*n is linearly independent in V**C**,

D . . . D D

completing the proof of (a).

Clearly (b) follows immediately from (a).

### Complexiﬁcation of an Operator

Now we can deﬁne the complexiﬁcation of an operator.

9.5 **Deﬁnition *complexiﬁcation of*** T***,*** T**C**

T, denoted T**C**, is the operator T**C** 2 *L*.V**C**/ deﬁned by

T**C**.u C i*v*/ D Tu C iT *v*

for u; *v* 2 V.

Suppose V is a real vector space and T 2 *L*.V /. The ***complexiﬁcation*** of

**278** CHAPTER 9 Operators on Real Vector Spaces

You should verify that if V is a real vector space and T 2 *L*.V /, then T**C**

is indeed in *L*.V**C**/. The key point here is tha（t our deﬁnit）ion of complex scalar

multiplication can be used to show that T**C**

all u; *v* 2 V and all **complex** numbers 入.

入.u C i*v*/

D 入T**C**.u C i*v*/ for

The next example gives a good way to think about the complexiﬁcation of a typical operator.

2 *L* D

9.6 **Example** Suppose A is an n-by-n matrix of real numbers. Deﬁne T .**R**n/ by Tx Ax, where elements of **R**n are thought of as n-by-1 column vectors. Identifying the complexiﬁcation of **R**n with **C**n, we then have T**C**z Az for each z **C**n, where again elements of **C**n are thought of as n-by-1 column vectors.

In other words, if T is the operator of matrix multiplication by A on **R**n,

D 2

then the complexiﬁcation T**C** is also matrix multiplication by A but now acting on the larger domain **C**n.

The next result makes sense because 9.4 tells us that a basis of a real vector space is also a basis of its complexiﬁcation. The proof of the next result follows immediately from the deﬁnitions.

9.7 Matrix of T**C** equals matrix of T

Then *M*.T / D *M*.T**C**/, where both matrices are with respect to the basis

Suppose V is a real vector space with basis *v*1;:::; *v*n and T 2 *L*.V /.

*v*1;:::; *v*n.

The result above and Example 9.6 provide complete insight into complexi- ﬁcation, because once a basis is chosen, every operator essentially looks like Example 9.6. Complexiﬁcation of an operator could have been deﬁned using matrices, but the approach taken here is more natural because it does not depend on the choice of a basis.

We know that every operator on a nonzero ﬁnite-dimensional complex vector space has an eigenvalue (see 5.21) and thus has a 1-dimensional in- variant subspace. We have seen an example [5.8(a)] of an operator on a

nonzero ﬁnite-dimensional real vector space with no eigenvalues and thus no 1-dimensional invariant subspaces. However, we now show that an invariant subspace of dimension 1 or 2 always exists. Notice how complexiﬁcation

leads to a simple proof of this result.

SECTION 9.A Complexiﬁcation **279**

9.8 Every operator has an invariant subspace of dimension 1 or 2

Every operator on a nonzero ﬁnite-dimensional vector space has an invariant subspace of dimension 1 or 2.

Proof Every operator on a nonzero ﬁnite-dimensional complex vector space has an eigenvalue (5.21) and thus has a 1-dimensional invariant subspace.

Hence assume V is a real vector space and T 2 *L*.V /. The complexiﬁca-

tion T**C** has an eigenvalue a C bi (by 5.21), where a; b 2 **R**. Thus there exist u; *v* 2 V, not both 0, such that T**C**.u C i*v*/ D .a C bi/.u C i*v*/. Using the deﬁnition of T**C**, the last equation can be rewritten as

Tu C iT *v* D .au — b*v*/ C .a*v* C bu/i:

Thus

Tu D au — b*v* and T *v* D a*v* C bu:

Let U equal the span in V of the list u; *v*. Then U is a subspace of V with dimension 1 or 2. The equations above show that U is invariant under T, completing the proof.

### The Minimal Polynomial of the Complexiﬁcation

Suppose V is a real vector space and T 2 *L*.V /. Repeated application of the deﬁnition of T**C** shows that

**9.9** .T**C**/n.u C i*v*/ D T nu C iT n*v*

for every positive integer n and all u; *v* 2 V.

Notice that the next result implies that the minimal polynomial of T**C** has

real coefﬁcients.

9.10 Minimal polynomial of T**C** equals minimal polynomial of T

Suppose V is a real vector space and T 2 *L*.V /. Then the minimal

polynomial of T**C** equals the minimal polynomial of T.

Proof Let p 2 *P*.**R**/ den（ote th）e minimal polynomial of T. From 9.9 it is

Suppose q 2 *P*.**C**/ is a monic polynomial such that q.T**C**/ D 0. Then

）

easy to see that p.T**C**/ D

p.T / **C** , and thus p.T**C**/ D 0.

（

q.T**C**/ .u/ D 0 for every u 2 V. Letting r denote the polynomial whose j th

coefﬁcient is the real part of the j th coefﬁcient of q, we see that r is a monic polynomial and r.T / 0. Thus deg q deg r deg p.

D D 乏

The conclusions of the two previous paragraphs imply that p is the minimal polynomial of T**C**, as desired.

**280** CHAPTER 9 Operators on Real Vector Spaces

### Eigenvalues of the Complexiﬁcation

Now we turn to questions about the eigenvalues of the complexiﬁcation of an operator. Again, everything that we expect to work indeed works easily.

We begin with a result showing that the real eigenvalues of T**C** are precisely the eigenvalues of T. We give two different proofs of this result. The ﬁrst

proof is more elementary, but the second proof is shorter and gives some useful insight.

9.11 Real eigenvalues of T**C**

Suppose V is a real vector space, T 2 *L*.V /, and 入 2 **R**. Then 入 is an

eigenvalue of T**C** if and only if 入is an eigenvalue of T.

Proof 1 First suppose 入 is an eigenvalue of T. Then there exists *v* 2 V with *v* ¤ 0 such that T *v* D 入*v*. Thus T**C***v* D 入*v*, which shows that 入 is an eigenvalue of T**C**, completing one direction of the proof.

To prove the other direction, suppose now that 入 is an eigenvalue of T**C**. Then there exist u; *v* 2 V with u C i*v* ¤ 0 such that

T**C**.u C i*v*/ D 入.u C i*v*/:

The equation above implies that Tu D 入u and T *v* D 入*v*. Because u ¤ 0 or

*v* ¤ 0, this implies that 入is an eigenvalue of T, completing the proof.

Proof 2 The (real) eigenvalues of T are the (real) zeros of the minimal polynomial of T (by 8.49). The real eigenvalues of T**C** are the real zeros of the minimal polynomial of T**C** (again by 8.49). These two minimal polynomials are the same (by 9.10). Thus the eigenvalues of T are precisely the real eigenvalues of T**C**, as desired.

Our next result shows that T**C** behaves symmetrically with respect to an eigenvalue 入 and its complex conjugate 入N .

9.12 T**C** — 入I and T**C** — 入N I

Suppose V is a real vector space, T 2 *L*.V /, 入 2 **C**, j is a nonnegative

integer, and u; *v* 2 V. Then

.T**C** — 入I / .u C i*v*/ D 0 if and only if .T**C** — 入I / .u — i*v*/ D 0:

j

N

j

SECTION 9.A Complexiﬁcation **281**

Proof We will prove this result by induction on j . To get started, note that if j 0 then (because an operator raised to the power 0 equals the identity operator) the result claims that u i*v* 0 if and only if u i*v* 0, which is

C D — D

D

clearly true.

Thus assume by induction that j 乏 1 and the desired result holds for j — 1.

Suppose .T**C** — 入I /j .u C i*v*/ D 0. Then

（ ）

* 1. .T**C** — 入I /j 1 .T**C** — 入I/.u C i*v*/ D 0:

Writing 入 D a C bi, where a; b 2 **R**, we have

* 1. .T**C** — 入I/.u C i*v*/ D .T u — au C b*v*/ C i.T *v* — a*v* — bu/

and

* 1. .T**C** — 入N I/.u — i*v*/ D .T u — au C b*v*/ — i.T *v* — a*v* — bu/:

Our induction hypothesis, 9.13, and 9.14 imply that

（ ）

.T**C** — 入N I/j 1 .T u — au C b*v*/ — i.T *v* — a*v* — bu/ D 0:

Now the equation above and 9.15 imply that .T**C** 入N I/j .u i*v*/ 0, completing the proof in one direction.

— — D

The other direction is proved by replacing 入with 入N , replacing *v* with *v*,

—

and then using the ﬁrst direction.

An important consequence of the result above is the next result, which states that if a number is an eigenvalue of T**C**, then its complex conjugate is also an eigenvalue of T**C**.

9.16 Nonreal eigenvalues of T**C** come in pairs

Suppose V is a real vector space, T 2 *L*.V /, and 入 2 **C**. Then 入 is an

eigenvalue of T**C** if and only if 入N is an eigenvalue of T**C**.

Proof Take j D 1 in 9.12.

By deﬁnition, the eigenvalues of an operator on a real vector space are real numbers. Thus when mathematicians sometimes informally mention the complex eigenvalues of an operator on a real vector space, what they have in mind is the eigenvalues of the complexiﬁcation of the operator.

Recall that the multiplicity of an eigenvalue is deﬁned to be the dimension of the generalized eigenspace corresponding to that eigenvalue (see 8.24). The next result states that the multiplicity of an eigenvalue of a complexiﬁcation equals the multiplicity of its complex conjugate.

**282** CHAPTER 9 Operators on Real Vector Spaces

9.17 Multiplicity of 入 equals multiplicity of 入N

Suppose V is a real vector space, T 2 *L*.V /, and 入 2 **C** is an eigenvalue

of T**C**. Then the multiplicity of 入 as an eigenvalue of T**C** equals the multiplicity of 入N as an eigenvalue of T**C**.

Proof Suppose u1 C i*v*1;:::; um C i*v*m is a basis of the generalized eigenspace G.入; T**C**/, where u1;:::; um; *v*1;:::; *v*m 2 V. Then using 9.12, routine arguments show that u1 — i*v*1;:::; um — i*v*m is a basis of the gen-

eralized eigenspace G.入N ; T**C**/. Thus both 入 and 入N

eigenvalues of T**C**.

have multiplicity m as

9.18 **Example** Suppose T 2 *L*.**R**3/ is deﬁned by

T .x1; x2; x3/ D .2x1; x2 — x3; x2 C x3/:0

2 0 0 1

The matrix of T with respect to the standard basis of **R**3 is 0 1 —1 :

@ A

0 1 1

As you can verify, 2 is an eigenvalue of T with multiplicity 1 and T has no

other eigenvalues.

If we identify the complexiﬁcation of **R**3 with **C**3, then the matrix of T**C** with respect to the standard basis of **C**3 is the matrix above. As you can verify, the eigenvalues of T**C** are 2, 1 C i, and 1 — i, each with multiplicity

1. Thus the nonreal eigenvalues of T**C** come as a pair, with each the complex

conjugate of the other and with the same multiplicity, as expected by 9.17.

We have seen an example [5.8(a)] of an operator on **R**2 with no eigenvalues.

The next result shows that no such example exists on **R**3.

9.19 Operator on odd-dimensional vector space has eigenvalue

Every operator on an odd-dimensional real vector space has an eigenvalue.

Proof Suppose V is a real vector space with odd dimension and T 2 *L*.V /. Because the nonreal eigenvalues of T**C** come in pairs with equal multiplicity (by 9.17), the sum of the multiplicities of all the nonreal eigenvalues of T**C** is

an even number.

Because the sum of the multiplicities of all the eigenvalues of T**C** equals the (complex) dimension of V**C** (by Theorem 8.26), the conclusion of the paragraph above implies that T**C** has a real eigenvalue. Every real eigenvalue of T**C** is also an eigenvalue of T (by 9.11), giving the desired result.

SECTION 9.A Complexiﬁcation **283**

### Characteristic Polynomial of the Complexiﬁcation

In the previous chapter we deﬁned the characteristic polynomial of an operator on a ﬁnite-dimensional complex vector space (see 8.34). The next result is a key step toward deﬁning the characteristic polynomial for operators on ﬁnite-dimensional real vector spaces.

9.20 Characteristic polynomial of T**C**

Suppose V is a real vector space and T 2 *L*.V /. Then the coefﬁcients of

the characteristic polynomial of T**C** are all real.

Proof Suppose 入is a nonreal eigenvalue of T**C** with multiplicity m. Then 入N is also an eigenvalue of T**C** with multiplicity m (by 9.17). Thus the characteristic polynomial of T**C** includes factors of .z 入/m and .z 入N /m. Multiplying

together these two factors, we have

— —

（ ）

.z — 入/m.z — 入N /m D z2 — 2.Re 入/z C j入j2 m:

The polynomial above on the right has real coefﬁcients.

The characteristic polynomial of T**C** is the product of terms of the form above and terms of the form .z — t /d , where t is a real eigenvalue of T**C** with multiplicity d. Thus the coefﬁcients of the characteristic polynomial of T**C**

are all real.

Now we can deﬁne the characteristic polynomial of an operator on a ﬁnite-dimensional real vector space to be the characteristic polynomial of its complexiﬁcation.

9.21 **Deﬁnition *Characteristic polynomial***

Suppose V is a real vector space and T 2 *L*.V /. Then the ***characteristic***

***polynomial*** of T is deﬁned to be the characteristic polynomial of T**C**.

9.22 **Example** Suppose T 2 *L*.**R**3/ is deﬁned by

T .x1; x2; x3/ D .2x1; x2 — x3; x2 C x3/:

As we noted in 9.18, the eigenvalues of T**C** are 2, 1 C i, and 1 — i, each with multiplicity 1. Thus the characteristic polynomial of the complexiﬁcation T**C** is .z — 2/ z — .1 C i/ z — .1 — i/ , which equals z3 — 4z2 C 6z — 4. Hence the characteristic polynomial of T is also z3 — 4z2 C 6z — 4.

（ ）（ ）

**284** CHAPTER 9 Operators on Real Vector Spaces

In the next result, the eigenvalues of T are all real (because T is an operator on a real vector space).

9.23 Degree and zeros of characteristic polynomial

Suppose V is a real vector space and T 2 *L*.V /. Then

1. the coefﬁcients of the characteristic polynomial of T are all real;
2. the characteristic polynomial of T has degree dim V ;
3. the eigenvalues of T are precisely the real zeros of the characteristic polynomial of T.

Proof Part (a) holds because of 9.20.

Part (b) follows from 8.36(a).

Part (c) holds because the real zeros of the characteristic polynomial of T are the real eigenvalues of T**C** [by 8.36(a)], which are the eigenvalues of T (by 9.11).

In the previous chapter, we proved the Cayley–Hamilton Theorem (8.37) for complex vector spaces. Now we can also prove it for real vector spaces.

9.24 Cayley–Hamilton Theorem

Then q.T / D 0.

Suppose T 2 *L*.V /. Let q denote the characteristic polynomial of T.

Proof We have already proved this result when V is a complex vector space. Thus assume that V is a real vector space.

The complex case of the Cayley–Hamilton Theorem (8.37) implies that

q.T**C**/ D 0. Thus we also have q.T / D 0, as desired.

9.25 **Example** Suppose T 2 *L*.**R**3/ is deﬁned by

T .x1; x2; x3/ D .2x1; x2 — x3; x2 C x3/:

As we saw in 9.22, the characteristic polynomial of T is z3 4z2 6z 4. Thus the Cayley–Hamilton Theorem implies that T 3 4T 2 6T 4I 0, which can also be veriﬁed by direct calculation.

— C — D

— C —

We can now prove another result that we previously knew only in the complex case.

SECTION 9.A Complexiﬁcation **285**

9.26 Characteristic polynomial is a multiple of minimal polynomial

Suppose T 2 *L*.V /. Then

1. the degree of the minimal polynomial of T is at most dim V ;
2. the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

Proof Part (a) follows immediately from the Cayley–Hamilton Theorem.

Part (b) follows from the Cayley–Hamilton Theorem and 8.46.

EXERCISES 9.A

1. Prove 9.3.
2. Verify that if V is a real vector space and T 2 *L*.V /, then T**C** 2 *L*.V**C**/.
3. Suppose V is a real vector space and *v*1;:::; *v*m 2 V. Prove that *v*1;:::; *v*m is linearly independent in V**C** if and only if *v*1;:::; *v*m is linearly independent in V.
4. Suppose V is a real vector space and *v*1;:::; *v*m 2 V. Prove that

*v*1;:::; *v*m spans V**C** if and only if *v*1;:::; *v*m spans V.

1. Suppose that V is a real vector space and S; T 2 *L*.V /. Show that

.S C T /**C** D S**C** C T**C** and that .入T /**C** D 入T**C** for every 入 2 **R**.

1. Suppose V is a real vector space and T .V /. Prove that T**C** is invertible if and only if T is invertible.

2 *L*

1. Suppose V is a real vector space and N .V /. Prove that N**C** is nilpotent if and only if N is nilpotent.

2 *L*

1. Suppose T .**R**3/ and 5; 7 are eigenvalues of T. Prove that T**C** has no nonreal eigenvalues.

2 *L*

1. Prove there does not exist an operator T .**R**7/ such that T 2 T I

2 *L* C C

is nilpotent.

1. Give an example of an operator T .**C**7/ such that T 2 T I

2 *L* C C

is nilpotent.

**286** CHAPTER 9 Operators on Real Vector Spaces

1. Suppose V is a real vector space and T 2 *L*.V /. Suppose there exist b; c 2 **R** such that T 2 C bT C cI D 0. Prove that T has an eigenvalue if and only if b2 乏 4c.
2. Suppose V is a real vector space and T .V /. Suppose there exist b; c **R** such that b2 < 4c and T 2 bT cI is nilpotent. Prove that T has no eigenvalues.

2 C C

2 *L*

1. Suppose V is a real vector space, T .V /, and b; c **R** are such that b2 < 4c. Prove that null.T 2 bT cI/j has even dimension for every positive integer j .

C C

2 *L* 2

1. Suppose V is a real vector space with dim V D 8. Suppose T 2 *L*.V /

is such that T 2 C T C I is nilpotent. Prove that .T 2 C T C I/4 D 0.

1. Suppose V is a real vector space and T .V / has no eigenvalues. Prove that every subspace of V invariant under T has even dimension.

2 *L*

1. Suppose V is a real vector space. Prove that there exists T 2 *L*.V / such that T 2 D —I if and only if V has even dimension.
2. Suppose V is a real vector space and T 2 *L*.V / satisﬁes T 2 D —I. Deﬁne complex scalar multiplication on V as follows: if a; b 2 **R**, then

.a C bi/*v* D a*v* C bT *v*:

* 1. Show that the complex scalar multiplication on V deﬁned above and the addition on V makes V into a complex vector space.
  2. Show that the dimension of V as a complex vector space is half the dimension of V as a real vector space.

1. Suppose V is a real vector space and T .V /. Prove that the following are equivalent:

2 *L*

* 1. All the eigenvalues of T**C** are real.
  2. There exists a basis of V with respect to which T has an upper- triangular matrix.
  3. There exists a basis of V consisting of generalized eigenvectors of T.

1. Suppose V is a real vector space with dim V D n and T 2 *L*.V / is such that null T n 2 ¤ null T n 1. Prove that T has at most two distinct eigenvalues and that T**C** has no nonreal eigenvalues.

SECTION 9.B Operators on Real Inner Product Spaces **287**

## *Operators on Real Inner Product Spaces*

9.B

We now switch our focus to the context of inner product spaces. We will give a complete description of normal operators on real inner product spaces; a key step in the proof of this result (9.34) requires the result from the previous section that an operator on a ﬁnite-dimensional real vector space has an

invariant subspace of dimension 1 or 2 (9.8).

After describing the normal operators on real inner product spaces, we will

use that result to give a complete description of isometries on such spaces.

### Normal Operators on Real Inner Product Spaces

The Complex Spectral Theorem (7.24) gives a complete description of normal operators on complex inner product spaces. In this subsection we will give a complete description of normal operators on real inner product spaces.

We begin with a description of the operators on 2-dimensional real inner

product spaces that are normal but not self-adjoint.

9.27 Normal but not self-adjoint operators

Then the following are equivalent:

Suppose V is a 2-dimensional real inner product space and T 2 *L*.V /.

(a)

(b)

T is normal but not self-adjoint.

The matrix of T with respect to every orthonormal basis of V has the form

a —b

b a

;

with b ¤ 0.

(c)

The matrix of T with respect to some orthonormal basis of V has

the form

a —b

b a

;

with b> 0.

Proof First suppose (a) holds, so that T is normal but not self-adjoint. Let

e1; e2 be an orthonormal basis of V. Suppose

**9.28** *M*（T; .e1; e2/） D a c :

b d

**288** CHAPTER 9 Operators on Real Vector Spaces

Then kTe1k2 D a2 C b2 and kT \*e1k2 D a2 C c2. Because T is normal, kTe1kD kT \*e1k (see 7.20); thus these equations imply that b2 D c2. Thus c D b or c D —b. But c ¤ b, because otherwise T would be self-adjoint, as

can be seen from the matrix in 9.28. Hence c D —b, so

**9.29** *M*（T; .e1; e2/） D a —b :

b d

The matrix of T \* is the transpose of the matrix above. Use matrix multipli- cation to compute the matrices of TT \* and T \*T (do it now). Because T is

normal, these two matrices are equal. Equating the entries in the upper-right corner of the two matrices you computed, you will discover that bd ab. Now b 0, because otherwise T would be self-adjoint, as can be seen from the matrix in 9.29. Thus d a, completing the proof that (a) implies (b).

D

¤

D

Now suppose (b) holds. We want to prove that (c) holds. Choose an orthonormal basis e1; e2 of V. We know that the matrix of T with respect to this basis has the form given by (b), with b 0. If b > 0, then (c) holds and we have proved that (b) implies (c). If b< 0, then, as you should verify,

¤

（ ）

the matrix of T with respect to the orthonormal basis e1; —e2 equals a b ,

where —b> 0; thus in this case we also see that (b) implies (c).

ba

Now suppose (c) holds, so that the matrix of T with respect to some orthonormal basis has the form given in (c) with b > 0. Clearly the matrix of T is not equal to its transpose (because b 0). Hence T is not self-adjoint. Now use matrix multiplication to verify that the matrices of TT \* and T \*T are equal. We conclude that TT \* T \*T. Hence T is normal. Thus (c)

¤

D

implies (a), completing the proof.

The next result tells us that a normal operator restricted to an invariant subspace is normal. This will allow us to use induction on dim V when we prove our description of normal operators (9.34).

9.30 Normal operators and invariant subspaces

subspace of V that is invariant under T. Then

1. U ? is invariant under T ;
2. U is invariant under T \*;
3. .T jU /\* D .T \*/jU ;
4. T jU 2 *L*.U / and T jU ? 2 *L*.U ?/ are normal operators.

Suppose V is an inner product space, T 2 *L*.V / is normal, and U is a

SECTION 9.B Operators on Real Inner Product Spaces **289**

Proof First we will prove (a). Let e1;:::; em be an orthonormal basis of U. Extend to an orthonormal basis e1;:::; em; f1;:::; fn of V (this is possible by 6.35). Because U is invariant under T, each T ej is a linear combination of e1;:::; em. Thus the matrix of T with respect to the basis e1;:::; em; f1;:::; fn is of the form

*M*.T / D

e1 ::: em f1 ::: fn

e1

0

1

: A B

B

em C

I

f1

: 0 C

B@

CA

fn

here A denotes an m-by-m matrix, 0 denotes the n-by-m matrix of all 0’s, B denotes an m-by-n matrix, C denotes an n-by-n matrix, and for convenience the basis has been listed along the top and left sides of the matrix.

For each j 1; :::; m , Tej 2 equals the sum of the squares of the absolute values of the entries in the j th column of A (see 6.25). Hence

2 f g k k

Xm the sum of the squares of the absolute

j D1

values of the entries of

A.

2

**9.31**

kTej k D

For each j 1; :::; m , T \*ej 2 equals the sum of the squares of the absolute values of the entries in the j th rows of A and B. Hence

Xm \*

2 f g k k

**9.32**

kT

ej k2 D

j D1

the sum of the squares of the absolute

values of the entries of

A

and

B

.

Because T is normal, kTej kD kT \*ej k for each j (see 7.20); thus

2

m

X

j D1

kTej k D

m

j D1

X

kT \*ej k2:

This equation, along with 9.31 and 9.32, implies that the sum of the squares of the absolute values of the entries of B equals 0. In other words, B is the matrix of all 0’s. Thus

**290** CHAPTER 9 Operators on Real Vector Spaces

#### 9.33

*M*.T / D

e1 ::: em f1 ::: fn

e1

0

1

: A 0

B

em C

:

f1

: 0 C

B@

CA

fn

This representation shows that Tfk is in the span of f1;:::; fn for each k. Because f1;:::; fn is a basis of U ?, this implies that T *v* U ? whenever *v* U ?. In other words, U ? is invariant under T, completing the proof of (a). To prove (b), note that .T \*/, which is the conjugate transpose of .T /, has a block of 0’s in the lower left corner (because .T /, as given above, has a block of 0’s in the upper right corner). In other words, each T \*ej can be written as a linear combination of e1;:::; em. Thus U is invariant under T \*,

*M*

*M M*

2

2

completing the proof of (b).

To prove (c), let S D T jU 2 *L*.U /. Fix *v* 2 U. Then

hSu; *v*iD hT u; *v*i

D hu; T \**v*i

for all u U. Because T \**v* U [by (b)], the equation above shows that S\**v* T \**v*. In other words, .T U /\* .T \*/ U, completing the proof of (c). To prove (d), note that T commutes with T \* (because T is normal) and that .T U /\* .T \*/ U [by (c)]. Thus T U commutes with its adjoint and hence is normal. Interchanging the roles of U and U ?, which is justiﬁed by

D j D j

2 2

j D j j

(a), shows that T jU ? is also normal, completing the proof of (d).

Our next result shows that normal operators on real inner product spaces come close to having diagonal matrices. Speciﬁcally, we get block diagonal ma- trices, with each block having size at

*Note that if an operator* T *has a*

*block diagonal matrix with respect*

*to some basis, then the entry in each* 1*-by-*1 *block on the diagonal of this matrix is an eigenvalue of* T*.*

most 2-by-2.

We cannot expect to do better than the next result, because on a real inner

product space there exist normal operators that do not have a diagonal matrix with respect to any basis. For example, the operator T .**R**2/ deﬁned by T .x; y/ . y; x/ is normal (as you should verify) but has no eigenvalues; thus this particular T does not have even an upper-triangular matrix with respect to any basis of **R**2.

D —

2 *L*

SECTION 9.B Operators on Real Inner Product Spaces **291**

9.34 Characterization of normal operators when **F** D **R**

Suppose V is a real inner product space and T 2 *L*.V /. Then the follow-

ing are equivalent:

(a)

(b)

T is normal.

There is an orthonormal basis of V with respect to which T has a block diagonal matrix such that each block is a 1-by-1 matrix or a 2-by-2 matrix of the form

a —b

b a

;

with b> 0.

Proof First suppose (b) holds. With respect to the basis given by (b), the matrix of T commutes with the matrix of T \* (which is the transpose of the matrix of T ), as you should verify (use Exercise 9 in Section 8.B for the product of two block diagonal matrices). Thus T commutes with T \*, which means that T is normal, completing the proof that (b) implies (a).

Now suppose (a) holds, so T is normal. We will prove that (b) holds by induction on dim V. To get started, note that our desired result holds if dim V 1 (trivially) or if dim V 2 [if T is self-adjoint, use the Real Spectral Theorem (7.29); if T is not self-adjoint, use 9.27].

D D

Now assume that dim V > 2 and that the desired result holds on vector spaces of smaller dimension. Let U be a subspace of V of dimension 1 that is invariant under T if such a subspace exists (in other words, if T has an eigenvector, let U be the span of this eigenvector). If no such subspace exists, let U be a subspace of V of dimension 2 that is invariant under T (an invariant subspace of dimension 1 or 2 always exists by 9.8).

If dim U 1, choose a vector in U with norm 1; this vector will be an orthonormal basis of U, and of course the matrix of T U .U / is a 1-by-1 matrix. If dim U 2, then T U .U / is normal (by 9.30) but not self-adjoint (otherwise T U, and hence T, would have an eigenvector by 7.27). Thus we can choose an orthonormal basis of U with respect to which the matrix of T jU 2 *L*.U / has the required form (see 9.27).

j

D j 2 *L*

j 2 *L*

D

Now U ? is invariant under T and T U ? is a normal operator on U ?

j

(by 9.30). Thus by our induction hypothesis, there is an orthonormal basis of U ? with respect to which the matrix of T U ? has the desired form. Adjoin- ing this basis to the basis of U gives an orthonormal basis of V with respect to which the matrix of T has the desired form. Thus (b) holds.

j

**292** CHAPTER 9 Operators on Real Vector Spaces

### Isometries on Real Inner Product Spaces

As we will see, the next example is a key building block for isometries on real inner product spaces. Also, note that the next example shows that an isometry on **R**2 may have no eigenvalues.

2

9.35 **Example** Let 0 **R**. Then the operator on **R**2 of counterclockwise rotation (centered at the origin) by an angle of 0 is an isometry, as is geomet-

rically obvious. The matrix of this operator with respect to the standard basis is

—

cos 0 sin 0

sin 0 cos 0 :

If 0 is not an integer multiple of 兀, then no nonzero vector of **R**2 gets mapped

to a scalar multiple of itself, and hence the operator has no eigenvalues.

The next result shows that every isometry on a real inner product space is composed of pieces that are rotations on 2-dimensional subspaces, pieces that equal the identity operator, and pieces that equal multiplication by —1.

9.36 Description of isometries when **F** D **R**

Suppose V is a real inner product space and S 2 *L*.V /. Then the following

are equivalent:

1. S is an isometry.
2. There is an orthonormal basis of V with respect to which S has a block diagonal matrix such that each block on the diagonal is a

1-by-1 matrix containing 1 or —1 or is a 2-by-2 matrix of the form

cos 0 sin 0

— sin 0

cos 0

;

with 0 2 .0; 兀/.

Proof First suppose (a) holds, so S is an isometry. Because S is normal, there is an orthonormal basis of V with respect to which S has a block diagonal matrix such that each block is a 1-by-1 matrix or a 2-by-2 matrix of the form

**9.37** a —b ;

b a

with b> 0 (by 9.34).

SECTION 9.B Operators on Real Inner Product Spaces **293**

If 入 is an entry in a 1-by-1 matrix along the diagonal of the matrix of S (with respect to the basis mentioned above), then there is a basis vector ej such that Sej 入ej . Because S is an isometry, this implies that 入 1.

D j jD

Thus 入 1 or 入 1, because these are the only real numbers with absolute value 1.

D D —

Now consider a 2-by-2 matrix of the form 9.37 along the diagonal of the matrix of S. There are basis vectors ej ; ej C1 such that

Sej D aej C bej C1:

Thus

1 D kej k2 D kSej k2 D a2 C b2:

The equation above, along with the condition b> 0, implies that there exists a number 0 .0; 兀/ such that a cos 0 and b sin 0 . Thus the matrix 9.37 has the required form, completing the proof in this direction.

2 D D

Conversely, now suppose (b) holds, so there is an orthonormal basis of V with respect to which the matrix of S has the form required by the theorem. Thus there is a direct sum decomposition

V D U1 ˚ ... ˚ Um;

where each Uj is a subspace of V of dimension 1 or 2. Furthermore, any two vectors belonging to distinct U ’s are orthogonal, and each SjU*j* is an isometry mapping Uj into Uj . If *v* 2 V, we can write

*v* D u1 C ... C um;

where each uj is in Uj . Applying S to the equation above and then taking norms gives

kS *v*k

2

D kS u1 C ... C Sumk

2 2

2

D kS u1k

2

C. . .C kS umk

2

D ku1k

2

C. . .C kumk

D k*v*k :

Thus S is an isometry, and hence (a) holds.

**294** CHAPTER 9 Operators on Real Vector Spaces

EXERCISES 9.B

1. Suppose S 2 *L*.**R**3/ is an isometry. Prove that there exists a nonzero vector x 2 **R**3 such that S 2x D x.
2. Prove that every isometry on an odd-dimensional real inner product space has 1 or —1 as an eigenvalue.
3. Suppose V is a real inner product space. Show that

hu C i*v*;x C iyiD hu; xiC h*v*; yiC （h*v*; xi— hu; yi）i

for u; *v*; x;y 2 V deﬁnes a complex inner product on V**C**.

1. Suppose V is a real inner product space and T 2 *L*.V / is self-adjoint. Show that T**C** is a self-adjoint operator on the inner product space V**C** deﬁned by the previous exercise.
2. Use the previous exercise to give a proof of the Real Spectral Theorem (7.29) via complexiﬁcation and the Complex Spectral Theorem (7.24).
3. Give an example of an operator T on an inner product space such that T has an invariant subspace whose orthogonal complement is not invariant under T.

[*The exercise above shows that 9.30 can fail without the hypothesis that*

T *is normal.*]

1. Suppose T 2 *L*.V / and T has a block diagonal matrix

0B

1

A1 0

: : C

@

A

:

0 Am

with respect to some basis of V. For j 1; :::; m, let Tj be the operator on V whose matrix with respect to the same basis is a block diagonal matrix with blocks the same size as in the matrix above, with Aj in the j th block, and with all the other blocks on the diagonal equal to identity matrices (of the appropriate size). Prove that T D T1 ... Tm.

D

1. Suppose D is the differentiation operator on the vector space V in Exercise 21 in Section 7.A. Find an orthonormal basis of V such that the matrix of the normal operator D has the form promised by 9.34.

# *Trace and Determinant*



*British mathematician and pioneer computer scientist Ada Lovelace* (*1815–1852*)*, as painted by Alfred Chalon in this 1840 portrait.*

CHAPTER

10

Throughout this book our emphasis has been on linear maps and operators rather than on matrices. In this chapter we pay more attention to matrices as we deﬁne the trace and determinant of an operator and then connect these notions to the corresponding notions for matrices. The book concludes with an explanation of the important role played by determinants in the theory of volume and integration.

Our assumptions for this chapter are as follows:

* 1. **Notation F*,*** V
     + **F** denotes **R** or **C**.
     + V denotes a ﬁnite-dimensional nonzero vector space over **F**.

LEARNING OBJECTIVES FOR THIS CHAPTER

change of basis and its effect upon the matrix of an operator trace of an operator and of a matrix

determinant of an operator and of a matrix determinants and volume

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,

295

DOI 10.1007/978-3-319-11080-6 10

**296** CHAPTER 10 Trace and Determinant

## *Trace*

10.A

For our study of the trace and determinant, we will need to know how the matrix of an operator changes with a change of basis. Thus we begin this chapter by developing the necessary material about change of basis.

### Change of Basis

With respect to every basis of V, the matrix of the identity operator I .V / is the diagonal matrix with 1’s on the diagonal and 0’s elsewhere. We also use the symbol I for the name of this matrix, as shown in the next deﬁnition.

2 *L*

10.2 **Deﬁnition *identity matrix, I***

Suppose n is a positive integer. The n-by-n diagonal matrix

B

1

0

@

: :

:

0

1

C

is called the ***identity matrix*** and is denoted I.

A

Note that we use the symbol I to denote the identity operator (on all vector spaces) and the identity matrix (of all possible sizes). You should always be

able to tell from the context which particular meaning of I is intended. For example, consider the equation .I/ I on the left side I denotes the identity operator, and on the right side I denotes the identity matrix.

*M* D I

If A is a square matrix (with entries in **F**, as usual) with the same size as I, then AI D IA D A, as you should verify.

10.3 **Deﬁnition *invertible, inverse,*** A 1

A square matrix A is called ***invertible*** if there is a square matrix B of

the same size such that AB D BA D I; we call B the ***inverse*** of A and

denote it by A 1.

The same proof as used in 3.54 shows that if A is an invertible square matrix, then there is a unique matrix B such that AB D BA D I (and thus the notation B D A 1 is justiﬁed).

*Some mathematicians use the terms* ***nonsingular****, which means the same as invertible, and* ***singular****, which means the same as noninvertible.*

SECTION 10.A Trace **297**

In Section 3.C we deﬁned the matrix of a linear map from one vector space to another with respect to two bases—one basis of the ﬁrst vector space and another basis of the second vector space. When we study operators, which are linear maps from a vector space to itself, we almost always use the same basis for both vector spaces (after all, the two vector spaces in question are equal). Thus we usually refer to the matrix of an operator with respect to a basis and display at most one basis because we are using one basis in two capacities.

The next result is one of the unusual cases in which we use two different bases even though we have operators from a vector space to itself. It is just a

convenient restatement of 3.43 (with U and W both equal to V ), but now we

are being more careful to include the various bases explicitly in the notation.

The result below holds because we deﬁned matrix multiplication to make it true—see 3.43 and the material preceding it.

The next result deals with the matrix of the identity operator I with respect to two different bases. Note that the kth column of the matrix I; .u1;:::; un/; .*v*1;:::; *v*n/ consists of the scalars needed to write uk

10.4 The matrix of the product of linear maps

Suppose u1;:::; un and *v*1;:::; *v*n and *w*1;:::; *w*n are all bases of V. Suppose S; T 2 *L*.V /. Then

*M*（S; .*v*1;:::; *v*n/; .*w*1;:::; *w*n/）*M*（T; .u1;:::; un/; .*v*1;:::; *v*n/）:

*M* ST; .u1;:::; un/; .*w*1;:::; *w*n/） D

（ ）*M*

as a linear combination of *v*1;:::; *v*n.

10.5 Matrix of the identity with respect to two bases

Suppose u1;:::; un and *v*1;:::; *v*n are bases of V. Then the matrices

*M* I; .u ;:::; u /; .

（

1

n 1

*v* ;:::;

*v* / and *M* I; .

n

）

（

*v* ;:::;

1

*v* /; .u ;:::; u /

n 1

n

are invertible, and each is the inverse of the other.

）

Proof In 10.4, replace *w*j with uj , and replace S and T with I, getting

I D *M* I; .*v*1;:::; *v*n/; .u1;:::; un/ *M* I; .u1;:::; un/; .*v*1;:::; *v*n/ :

（ ） （ ）

Now interchange the roles of the u’s and *v*’s, getting

（ ） （ ）

I D *M* I; .u1;:::; un/; .*v*1;:::; *v*n/ *M* I; .*v*1;:::; *v*n/; .u1;:::; un/ :

These two equations give the desired result.

**298** CHAPTER 10 Trace and Determinant

10.6 **Example** Consider the bases .4; 2/; .5; 3/ and .1; 0/; .0; 1/ of **F**2.

Obviously

*M* I; .4; 2/; .5; 3/ ; .1; 0/; .0; 1/ D 4 5 ;

2 3

（ ） （ ）

because I.4; 2/ 4.1; 0/ 2.0; 1/ and I.5; 3/ 5.1; 0/ 3.0; 1/.

D C D C

The inverse of the matrix above is

2 — 2 ! ;

5

3

—1 2

as you should verify. Thus 10.5 implies that

（ ） （

） 3

— 2 !

*M* I;

.1; 0/; .0; 1/ ;

.4; 2/; .5; 3/ D

:

—1 2

5

2

Now we can see how the matrix of T changes when we change bases. In

Proof In 10.4, replace *w*j with uj and replace S with I, getting

the result below, we have two different bases of V. Recall that the notation

*M* T; .u ;:::; u / is shorthand for *M* T; .u ;:::; u /; .u ;:::; u /

（

1

n

）

（

1

n 1

n

）

10.7 Change of basis formula

Suppose T 2 *L*.V /. Let u1;:::; un and *v*1;:::; *v*n be bases of V. Let

A D *M* I; .u ;:::; u /; .

（

1

n 1

*v* ;:::;

*v* / . Then

n

*M*（T; .u1;:::; un/） D A 1*M*（T; .*v*1;:::; *v*n/）A:

）

（ ） （ ）

**10.8** *M* T; .u1;:::; un/ D A 1*M* T; .u1;:::; un/; .*v*1;:::; *v*n/ ;

where we have used 10.5.

Again use 10.4, this time replacing *w*j with *v*j . Also replace T with I and replace S with T, getting

（ ） （ ）

*M* T; .u1;:::; un/; .*v*1;:::; *v*n/ D *M* T; .*v*1;:::; *v*n/ A:

Substituting the equation above into 10.8 gives the desired result.

SECTION 10.A Trace **299**

### Trace: A Connection Between Operators and Matrices

Suppose T .V / and 入 is an eigenvalue of T. Let n dim V. Re- call that we deﬁned the multiplicity of 入 to be the dimension of the gen- eralized eigenspace G.入; T / (see 8.24) and that this multiplicity equals dim null.T 入 I /n (see 8.11). Recall also that if V is a complex vector space, then the sum of the multiplicities of all the eigenvalues of T equals n

—

2 *L* D

(see 8.26).

In the deﬁnition below, the sum of the eigenvalues “with each eigenvalue repeated according to its multiplicity” means that if 入1;:::; 入m are the distinct eigenvalues of T (or of T**C** if V is a real vector space) with multiplicities d1;:::; dm, then the sum is

d1入1 C C dm入m:

Or if you prefer to list the eigenvalues with each repeated according to its multiplicity, then the eigenvalues could be denoted 入1;:::; 入n (where the index n equals dim V ) and the sum is

入1 C C 入n:

10.9 **Deﬁnition *trace of an operator***

Suppose T 2 *L*.V /.

* If **F** D **C**, then the ***trace*** of T is the sum of the eigenvalues of T,

with each eigenvalue repeated according to its multiplicity.

with each eigenvalue repeated according to its multiplicity.

The trace of T is denoted by trace T.

* If **F** D **R**, then the ***trace*** of T is the sum of the eigenvalues of T**C**,

10.10 **Example** Suppose T .**C**3/ is the operator whose matrix is

0@ 3 —1 —2 1A

2 *L*

3 2 —3

1 2 0

:

Then the eigenvalues of T are 1, 2 3i, and 2 3i, each with multiplicity 1, as you can verify. Computing the sum of the eigenvalues, we ﬁnd that trace T D 1 C .2 C 3i/ C .2 — 3i/; in other words, trace T D 5.

C —

**300** CHAPTER 10 Trace and Determinant

The trace has a close connection with the characteristic polynomial. Sup- pose 入1;:::; 入n are the eigenvalues of T (or of T**C** if V is a real vector space) with each eigenvalue repeated according to its multiplicity. Then by deﬁnition (see 8.34 and 9.21), the characteristic polynomial of T equals

.z — 入1/ z — 入n/:

Expanding the polynomial above, we can write the characteristic polynomial of T in the form

**10.11** zn — .入1 C ... C 入n/zn 1 C ... C .—1/n.入1 入n/:

The expression above immediately leads to the following result.

10.12 Trace and characteristic polynomial

Suppose T 2 *L*.V /. Let n D dim V. Then trace T equals the negative of

the coefﬁcient of zn 1 in the characteristic polynomial of T.

Most of the rest of this section is devoted to discovering how to compute trace T from the matrix of T (with respect to an arbitrary basis).

Let’s start with the easiest situation. Suppose V is a complex vector space,

1. .V /, and we choose a basis of V as in 8.29. With respect to that basis, T has an upper-triangular matrix with the diagonal of the matrix containing precisely the eigenvalues of T, each repeated according to its multiplicity. Thus trace T equals the sum of the diagonal entries of .T / with respect to

*M*

2 *L*

that basis.

The same formula works for the operator T .**C**3/ in Example 10.10 whose trace equals 5. In that example, the matrix is not in upper-triangular form. However, the sum of the diagonal entries of the matrix in that example

2 *L*

equals 5, which is the trace of the operator T.

At this point you should suspect that trace T equals the sum of the diagonal entries of the matrix of T with respect to an arbitrary basis. Remarkably, this suspicion turns out to be true. To prove it, we start by making the following

deﬁnition.

10.13 **Deﬁnition *trace of a matrix***

The ***trace*** of a square matrix A, denoted trace A, is deﬁned to be the sum of the diagonal entries of A.

SECTION 10.A Trace **301**

Now we have deﬁned the trace of an operator and the trace of a square matrix, using the same word “trace” in two different contexts. This would be bad terminology unless the two concepts turn out to be essentially the same.

（ ）D *M*

As we will see, it is indeed true that trace T trace T; .*v*1;:::; *v*n/ , where *v*1;:::; *v*n is an arbitrary basis of V. We will need the following result

for the proof.

10.14 Trace of AB equals trace of BA

If A and B are square matrices of the same size, then

trace.AB/ D trace.BA/:

Proof Suppose

B0 A1;1 ::: A1;n 1C B0 B1;1 ::: B1;n 1C

A D @

:

:

A ; B D @

:

:

A :

An;1 ::: An;n

The j th term on the diagonal of AB equals

Xn

Bn;1 ::: Bn;n

Thus

n

X

trace.AB/ D

kD1

Xn

Aj;k Bk;j :

Aj;k Bk;j

j D1 kD1

X

X

n n

D Bk;j Aj;k

kD1 j D1 n

X

D kth term on the diagonal of BA

kD1

as desired.

D trace.BA/;

Now we can prove that the sum of the diagonal entries of the matrix of an operator is independent of the basis with respect to which the matrix is computed.

**302** CHAPTER 10 Trace and Determinant

10.15 Trace of matrix of operator does not depend on basis

Let T 2 *L*.V /. Suppose u1;:::; un and *v*1;:::; *v*n are bases of V. Then trace *M*（T; .u1;:::; un/） D trace *M*（T; .*v*1;:::; *v*n/）:

Proof Let A D *M*（I; .u1;:::; un/; .*v*1;:::; *v*n/）. Then

（ ） （ （ ） ）*M* D *M*

trace T; .u1;:::; un/ trace A 1 T; .*v*1;:::; *v*n/ A

D trace （*M*（T; .*v*1;:::; *v*n/）A）A 1

（ ）D trace *M* T;.*v* ;:::; *v* / ;1 n

where the ﬁrst equality comes from 10.7 and the second equality follows from 10.14. The third equality completes the proof.

The result below, which is the most important result in this section, states that the trace of an operator equals the sum of the diagonal entries of the matrix of the operator. This theorem does not specify a basis because, by the result above, the sum of the diagonal entries of the matrix of an operator is the same for every choice of basis.

10.16 Trace of an operator equals trace of its matrix

Suppose T 2 *L*.V /. Then trace T D trace *M*.T /.

Proof As noted above, trace .T / is independent of which basis of V we choose (by 10.15). Thus to show that

*M*

trace T D trace *M*.T /

for every basis of V, we need only show that the equation above holds for some basis of V.

As we have already discussed, if V is a complex vector space, then choos- ing the basis as in 8.29 gives the desired result. If V is a real vector space, then applying the complex case to the complexiﬁcation T**C** (which is used to deﬁne trace T ) gives the desired result.

If we know the matrix of an operator on a complex vector space, the result above allows us to ﬁnd the sum of all the eigenvalues without ﬁnding any of the eigenvalues, as shown by the next example.

SECTION 10.A Trace **303**

10.17 **Example** Consider the operator on **C**5 whose matrix is

0 0 0 0 —3

0B 1C

1 0 0 0 6

0 1 0 0 0

B@

:

CA

0 0 1 0 0

0 0 0 1 0

No one can ﬁnd an exact formula for any of the eigenvalues of this operator. However, we do know that the sum of the eigenvalues equals 0, because the sum of the diagonal entries of the matrix above equals 0.

We can use 10.16 to give easy proofs of some useful properties about traces of operators by shifting to the language of traces of matrices, where

certain properties have already been proved or are obvious. The proof of the next result is an example of this technique. The eigenvalues of S T are not, in general, formed from adding together eigenvalues of S and eigenvalues of

C

T. Thus the next result would be difﬁcult to prove without using 10.16.

10.18 Trace is additive

Suppose S; T 2 *L*.V /. Then trace.S C T/ D trace S C trace T.

Proof Choose a basis of V. Then

trace.S C T/ D trace *M*.S C T/

（ ）

D trace *M*.S/ C *M*.T /

D trace *M*.S/ C trace *M*.T /

D trace S C trace T;

where again the ﬁrst and last equalities come from 10.16; the third equality is obvious from the deﬁnition of the trace of a matrix.

The techniques we have developed have the following curious consequence. A generalization of this result to inﬁnite- dimensional vector spaces has impor- tant consequences in modern physics, particularly in quantum theory.

*The statement of the next result does not involve traces, although the short proof uses traces. When- ever something like this happens in mathematics, we can be sure that a good deﬁnition lurks in the back- ground.*

**304** CHAPTER 10 Trace and Determinant

10.19 The identity is not the difference of ST and TS

There do not exist operators S; T 2 *L*.V / such that ST — TS D I.

Proof Suppose S; T 2 *L*.V /. Choose a basis of V . Then trace.ST — TS/ D trace.ST / — trace.TS/

D trace *M*.ST / — trace *M*.TS/

（ ） （ ）

D trace *M*.S/*M*.T / — trace *M*.T /*M*.S/

D 0;

where the ﬁrst equality comes from 10.18, the second equality comes from 10.16, the third equality comes from 3.43, and the fourth equality comes from

10.14. Clearly the trace of I equals dim V, which is not 0. Because ST TS

—

and I have different traces, they cannot be equal.

EXERCISES 10.A

1. Sup（pose T 2 *L*.V /） and *v*1;:::; *v*n is a basis of V. Prove that the matrix
2. Suppose A and B are square matrices of the same size and AB D I. Prove that BA D I.

*M* T; .*v*1;:::; *v*n/

is invertible if and only if T is invertible.

1. Suppose T .V / has the same matrix with respect to every basis of V. Prove that T is a scalar multiple of the identity operator.

2 *L*

1. Suppose u1;:::; un and *v*1;:::; *v*n are bases of V. Let T 2 *L*.V / be the operator such that T *v*k D uk for k D 1; :::; n. Prove that

（ ） （ ）

*M* T; .*v*1;:::; *v*n/ D *M* I; .u1;:::; un/; .*v*1;:::; *v*n/ :

1. Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that A 1BA is an upper-triangular matrix.
2. Give an example of a real vector space V and T .V / such that trace.T 2/< 0.

2 *L*

1. Suppose V is a real vector space, T 2 *L*.V /, and V has a basis consisting of eigenvectors of T. Prove that trace.T 2/ 乏 0.

SECTION 10.A Trace **305**

1. Suppose V is an inner product space and *v*; *w* 2 V. Deﬁne T 2 *L*.V / by

Tu D hu; *v*i*w*. Find a formula for trace T.

1. Suppose P 2 *L*.V / satisﬁes P 2 D P . Prove that

trace P D dim range P:

1. Suppose V is an inner product space and T 2 *L*.V /. Prove that

trace T \* D trace T:

1. Suppose V is an inner product space. Suppose T 2 *L*.V / is a positive operator and trace T D 0. Prove that T D 0.
2. Suppose V is an inner product space and P; Q 2 *L*.V / are orthogonal projections. Prove that trace.PQ/ 乏 0.
3. Suppose T 2 *L*.**C**3/ is the operator whose matrix is

0@ 51 —12 —21 1A

60 —40 —28

57 —68 1

:

Someone tells you (accurately) that 48 and 24 are eigenvalues of T. Without using a computer or writing anything down, ﬁnd the third eigen- value of T.

—

1. Suppose T 2 *L*.V / and c 2 **F**. Prove that trace.cT / D c trace T.
2. Suppose S; T 2 *L*.V /. Prove that trace.ST / D trace.TS/.
3. Prove or give a counterexample: if S; T .V /, then trace.ST /

2 *L* D

.trace S/.trace T /.

1. Suppose T 2 *L*.V / is such that trace.ST / D 0 for all S 2 *L*.V /. Prove that T D 0.
2. Suppose V is an inner product space with orthonormal basis e1;:::; en

and T 2 *L*.V /. Prove that

trace.T \*T/ D kT e1k2 C. . .C kT enk2:

Conclude that the right side of the equation above is independent of which orthonormal basis e1;:::; en is chosen for V.

**306** CHAPTER 10 Trace and Determinant

1. Suppose V is an inner product space. Prove that

hS; T iD trace.ST \*/

deﬁnes an inner product on *L*.V /.

1. Suppose V is a complex inner product space and T .V /. Let

2 *L*

入1;:::; 入n be the eigenvalues of T, repeated according to multiplicity.

Suppose

B0 A1;1 ::: A1;n 1C

@ : : A

An;1 ::: An;n

is the matrix of T with respect to some orthonormal basis of V. Prove that

X

X

n n

j入1j C. . .C j入nj 三 jAj;kj :

2

2

2

kD1 j D1

1. Suppose V is an inner product space. Suppose T 2 *L*.V / and

kT \**v*k三 kT *v*k

for every *v* V. Prove that T is normal.

2

[*The exercise above fails on inﬁnite-dimensional inner product spaces,*

*leading to what are called hyponormal operators, which have a well- developed theory.*]

SECTION 10.B Determinant **307**

## *Determinant*

10.B

### Determinant of an Operator

Now we are ready to deﬁne the determinant of an operator. Notice that the deﬁnition below mimics the approach we took when deﬁning the trace, with the product of the eigenvalues replacing the sum of the eigenvalues.

10.20 **Deﬁnition *determinant of an operator,*** det T

Suppose T 2 *L*.V /.

* If **F** D **C**, then the ***determinant*** of T is the product of the eigenvalues

of T, with each eigenvalue repeated according to its multiplicity.

of T**C**, with each eigenvalue repeated according to its multiplicity.

The determinant of T is denoted by det T.

* If **F** D **R**, then the ***determinant*** of T is the product of the eigenvalues

If 入1;:::; 入m are the distinct eigenvalues of T (or of T**C** if V is a real vector space) with multiplicities d1;:::; dm, then the deﬁnition above implies

det T D 入d*1* ... 入d*m* :

1

m

Or if you prefer to list the eigenvalues with each repeated according to its multiplicity, then the eigenvalues could be denoted 入1;:::; 入n (where the index n equals dim V ) and the deﬁnition above implies

det T D 入1 ... 入n:

10.21 **Example** Suppose T .**C**3/ is the operator whose matrix is

0@ 3 —1 —2 1A

2 *L*

3 2 —3

1 2 0

:

Then the eigenvalues of T are 1, 2 3i, and 2 3i, each with multiplicity 1, as you can verify. Computing the product of the eigenvalues, we ﬁnd that det T D 1 . .2 C 3i/ . .2 — 3i/; in other words, det T D 13.

C —

**308** CHAPTER 10 Trace and Determinant

The determinant has a close connection with the characteristic polynomial. Suppose 入1;:::; 入n are the eigenvalues of T (or of T**C** if V is a real vector space) with each eigenvalue repeated according to its multiplicity. Then the expression for the characteristic polynomial of T given by 10.11 gives the following result.

10.22 Determinant and characteristic polynomial

Suppose T 2 *L*.V /. Let n D dim V. Then det T equals .—1/n times the

constant term of the characteristic polynomial of T.

Combining the result above and 10.12, we have the following result.

10.23 Characteristic polynomial, trace, and determinant

written as

Suppose T 2 *L*.V /. Then the characteristic polynomial of T can be

zn — .trace T /zn 1 C ... C .—1/n.det T /:

We turn now to some simple but important properties of determinants. Later we will discover how to calculate det T from the matrix of T (with respect to an arbitrary basis).

The crucial result below has an easy proof due to our deﬁnition.

10.24 Invertible is equivalent to nonzero determinant

An operator on V is invertible if and only if its determinant is nonzero.

Proof First suppose V is a complex vector space and T .V /. The operator T is invertible if and only if 0 is not an eigenvalue of T. Clearly this happens if and only if the product of the eigenvalues of T is not 0. Thus T is invertible if and only if det T 0, as desired.

2 *L*

¤

Now consider the case where V is a real vector space and T .V /. Again, T is invertible if and only if 0 is not an eigenvalue of T, which happens if and only if 0 is not an eigenvalue of T**C** (because T**C** and T have the same real eigenvalues by 9.11). Thus again we see that T is invertible if and only if det T ¤ 0.

2 *L*

Some textbooks take the result below as the deﬁnition of the characteristic polynomial and then have our deﬁnition of the characteristic polynomial as a consequence.

SECTION 10.B Determinant **309**

10.25 Characteristic polynomial of T equals det.zI — T/

Suppose T 2 *L*.V /. Then the characteristic polynomial of T equals

det.zI — T /.

Proof First suppose V is a complex vector space. If 入 ; z **C**, then 入 is an eigenvalue of T if and only if z 入is an eigenvalue of zI T, as can be seen from the equation

—.T — 入I / D .zI — T/ — .z — 入/I:

— —

2

Raising both sides of this equation to the dim V power and then taking null spaces of both sides shows that the multiplicity of 入 as an eigenvalue of T equals the multiplicity of z 入 as an eigenvalue of zI T.

— —

Let 入1;:::; 入n denote the eigenvalues of T, repeated according to mul- tiplicity. Thus for z 2 **C**, the paragraph above shows that the eigenvalues of zI — T are z — 入1;:::;z — 入n, repeated according to multiplicity. The determinant of zI — T is the product of these eigenvalues. In other words,

det.zI — T/ D .z — 入1/ z — 入n/:

The right side of the equation above is, by deﬁnition, the characteristic poly- nomial of T, completing the proof when V is a complex vector space.

Now suppose V is a real vector space. Applying the complex case to T**C**

gives the desired result.

### Determinant of a Matrix

Our next task is to discover how to compute det T from the matrix of T (with respect to an arbitrary basis). Let’s start with the easiest situation. Suppose V is a complex vector space, T .V /, and we choose a basis of V as in

2 *L*

8.29. With respect to that basis, T has an upper-triangular matrix with the

diagonal of the matrix containing precisely the eigenvalues of T, each repeated according to its multiplicity. Thus det T equals the product of the diagonal entries of .T / with respect to that basis.

*M*

When dealing with the trace in the previous section, we discovered that the formula (trace = sum of diagonal entries) that worked for the upper-triangular matrix given by 8.29 also worked with respect to an arbitrary basis. Could that also work for determinants? In other words, is the determinant of an operator equal to the product of the diagonal entries of the matrix of the operator with respect to an arbitrary basis?

**310** CHAPTER 10 Trace and Determinant

Unfortunately, the determinant is more complicated than the trace. In par- ticular, det T need not equal the product of the diagonal entries of .T / with respect to an arbitrary basis. For example, the operator in Example 10.21 has determinant 13 but the product of the diagonal entries of its matrix equals 0.

*M*

For each square matrix A, we want to deﬁne the determinant of A, denoted

det A, so that det T det .T / regardless of which basis is used to com- pute .T /. We begin our search for the correct deﬁnition of the determinant of a matrix by calculating the determinants of some special operators.

*M*

D *M*

10.26 **Example** Suppose a1;:::; an 2 **F**. Let

A D 0B

B

@

0 an

a1 0

a2 0

1C I

C

: : : : : : A

an 1 0

here all entries of the matrix are 0 except for the upper-right corner and along the line just below the diagonal. Suppose *v*1;:::; *v*n is a basis of V and T 2 *L*.V / is such that *M* T; .*v*1;:::; *v*n/ D A. Find the determinant of T.

（ ）

Solution First assume aj ¤ 0 for each j D 1; :::;n — 1. Note that the list

*v*1;T *v*1;T 2*v*1;:::;T n 1*v*1 equals *v*1; a1*v*2; a1a2*v*3;:::; a1 ... an 1*v*n.

Thus *v*1;T *v*1;:::;T n 1*v*1 is lin- early independent (because the a’s are all nonzero). Hence if p is a monic poly- nomial with degree at most n 1, then

*Computing the minimal polynomial is often an efﬁcient method of ﬁnd- ing the characteristic polynomial, as is done in this example.*

—

p.T /*v*1 0. Thus the minimal poly- nomial of T cannot have degree less than n.

¤

As you should verify, T n*v*j a1 an*v*j for each j . Thus we have T n a1 anI. Hence zn a1 an is the minimal polynomial of T. Be- cause n dim V and the characteristic polynomial is a polynomial multiple of the minimal polynomial (9.26), this implies that zn a1 an is also the characteristic polynomial of T.

— ...

D

D ... — ...

D ...

Thus 10.22 implies that

det T D .—1/n 1a1 ... an:

If some aj equals 0, then T *v*j 0 for some j , which implies that 0 is an eigenvalue of T and hence det T 0. In other words, the formula above also holds if some aj equals 0.

D

D

SECTION 10.B Determinant **311**

Thus in order to have det T det .T /, we will have to make the deter- minant of the matrix in Example 10.26 equal to . 1/n 1a1 an. However, we do not yet have enough evidence to make a reasonable guess about the

— ...

D *M*

proper deﬁnition of the determinant of an arbitrary square matrix.

To compute the determinants of a more complicated class of operators, we introduce the notion of permutation.

10.27 **Deﬁnition *permutation,*** perm n

each of the numbers 1; :::;n exactly once.

* The set of all permutations of .1; : : : ; n/ is denoted perm n.
* A ***permutation*** of .1; : : : ; n/ is a list .m1;:::; mn/ that contains

For example, .2; 3; 4; 5; 1/ perm 5. You should think of an element of perm n as a rearrangement of the ﬁrst n integers.

* 1. **Example** Suppose a1;:::; an **F** and *v*1;:::; *v*n is a basis of V. Consider a permutation .p1;:::; pn/ perm n that can be obtained as fol- lows: break .1; : : : ; n/ into lists of consecutive integers and in each list move the ﬁrst term to the end of that list. For example, taking n D 9, the permutation

.2; 3; 1; 5; 6; 7; 4; 9; 8/

2

2

2

is obtained from .1; 2; 3/; .4; 5; 6; 7/; .8; 9/ by moving the ﬁrst term of each of these lists to the end, producing .2; 3; 1/; .5; 6; 7; 4/; .9; 8/, and then putting these together to form the permutation displayed above.

Let T 2 *L*.V / be the operator such that

T *v*k D ak*v*p*k*

for k D 1; :::; n. Find det T.

Solution This generalizes Example 10.26, because if .p1;:::; pn/ is the permutation .2; 3;:::; n; 1/, then our operator T is the same as the operator T in Example 10.26.

With respect to the basis *v*1;:::; *v*n, the matrix of the operator T is a block

diagonal matrix

A D @

: : :

0 AM

A ;

0B A1 0 1C

where each block is a square matrix of the form of the matrix in 10.26.

**312** CHAPTER 10 Trace and Determinant

Correspondingly, we can write V D V1 ˚ ... ˚ VM , where each Vj is invariant under T and each T jV*j* is of the form of the operator in 10.26. Because det T .det T V*1* / .det T V*M* / (because the dimensions of the generalized eigenspaces in the Vj add up to dim V ), we have

D j ... j

det T D .—1/n*1* 1 ... .—1/n*M*  1a1 ... an;

where Vj has dimension nj (and correspondingly each Aj has size nj -by-nj ) and we have used the result from 10.26.

— ... —

The number . 1/n*1* 1 . 1/n*M*  1 that appears above is called the sign of the corresponding permutation .p1;:::; pn/, denoted sign.p1;:::; pn/ [this is a temporary deﬁnition that we will change to an equivalent deﬁnition

later, when we deﬁne the sign of an arbitrary permutation].

To put this into a form that does not depend on the particular permutation

.p1;:::; pn/, let Aj;k denote the entry in row j , column k, of the matrix A

from Example 10.28. Thus

,D

Aj;k

0 if j ¤ pk;

ak if j D pk.

Example 10.28 shows that we want

* 1. det A D X

（sign.m1;:::; mn/）Am*1*;1 ... Am*n*;nI

.m*1*;:::;m*n*/2perm n

note that each summand is 0 except the one corresponding to the permutation

.p1;:::; pn/ [which is why it does not matter that the sign of the other

permutations is not yet deﬁned].

We can now guess that det A should be deﬁned by 10.29 for an arbitrary square matrix A. This will turn out to be correct. We will now dispense with the motivation and begin the more formal approach. First we will need to

deﬁne the sign of an arbitrary permutation.

10.30 **Deﬁnition *sign of a permutation***

* The ***sign*** of a permutation .m1;:::; mn/ is deﬁned to be 1 if the

number of pairs of integers .j; k/ with 1 三 j < k 三 n such that

j appears after k in the list .m1;:::; mn/ is even and —1 if the

number of such pairs is odd.

* In other words, the sign of a permutation equals 1 if the natural

order has been changed an even number of times and equals —1 if

the natural order has been changed an odd number of times.

SECTION 10.B Determinant **313**

10.31 **Example *sign of permutation***

The only pair of integers .j; k/ with j < k such that j appears after k in the list .2; 1; 3; 4/ is .1; 2/. Thus the permutation .2; 1; 3; 4/ has sign —1.

•

In the permutation .2; 3;:::; n; 1/, the only pairs .j; k/ with j < k that appear with changed order are .1; 2/; .1; 3/; : : : ; .1; n/; because we have n 1 such pairs, the sign of this permutation equals . 1/n 1 (note

•

— —

that the same quantity appeared in Example 10.26).

The next result shows that interchanging two entries of a permutation changes the sign of the permutation.

10.32 Interchanging two entries in a permutation

Interchanging two entries in a permutation multiplies the sign of the permutation by —1.

Proof Suppose we have two permutations, where the second permutation is obtained from the ﬁrst by interchanging two entries. If the two interchanged entries were in their natural order in the ﬁrst permutation, then they no longer

are in the second permutation, and vice versa, for a net change (so far) of 1 or

—1 (both odd numbers) in the number of pairs not in their natural order.

Consider each entry between the two

*Some texts use the term* ***signum****, which means the same as sign.*

interchanged entries. If an intermediate entry was originally in the natural order

with respect to both interchanged entries, then it is now in the natural order with respect to neither interchanged entry. Similarly, if an intermediate entry was originally in the natural order with respect to neither of the interchanged entries, then it is now in the natural order with respect to both interchanged entries. If an intermediate entry was originally in the natural order with respect to exactly one of the interchanged entries, then that is still true. Thus the net

change for each intermediate entry in the number of pairs not in their natural order is 2, 2, or 0 (all even numbers).

—

For all the other entries, there is no change in the number of pairs not in

their natural order. Thus the total net change in the number of pairs not in their natural order is an odd number. Thus the sign of the second permutation

equals —1 times the sign of the ﬁrst permutation.

Our motivation for the next deﬁnition comes from 10.29.

R:: 100JC CHAPTER 10 Trace and Determinant

10.33 **Deﬁnition *determinant of a matrix,*** det A

Suppose A is an n-by-n matrix

A D

B

A

1;1

::: A

1;n

@

:

An;1

:::

:

An;n

C

The ***determinant*** of A, denoted det A, is deﬁned by

A

:

det A D

.m*1*;:::;m*n*/2perm n

X

（sign.m1;:::; mn/）Am*1*;1 ... Am*n*;n:

* 1. **Example *determinants***

If A is the 1-by-1 matrix ŒA1;1], then det A A1;1, because perm 1 has only one element, namely .1/, which has sign 1.

* D
* Clearly perm 2 has only two elements, namely .1; 2/, which has sign 1,

and .2; 1/, whic h has sign —1. Thus

det

A1;1 A1;2

A2;1 A2;2

*The set* perm 3 *contains six ele- ments. In general,* perm n *contains* nŠ *elements. Note that* nŠ *rapidly grows large as* n *increases.*

D A1;1A2;2 — A2;1A1;2:

To make sure you understand this process, you should now ﬁnd the for- mula for the determinant of an arbitrary

3-by-3 matrix using just the deﬁnition

given above.

* 1. **Example** Compute the determinant of an upper-triangular matrix

0B A1;1 \* 1C

A D @

: : :

0 An;n

A :

Solution The permutation .1; 2; :::; n/ has sign 1 and thus contributes a term of A1;1 ... An;n to the sum deﬁning det A in 10.33. Any other permutation

.m1;:::; mn/ 2 perm n contains at least one entry mj with mj > j, which means that Am*j* ;j 0 (because A is upper triangular). Thus all the other

D

terms in the sum in 10.33 make no contribution.

Hence det A A1;1 An;n. In other words, the determinant of an upper- triangular matrix equals the product of the diagonal entries.

D ...

SECTION 10.B Determinant **315**

Suppose V is a complex vector space, T .V /, and we choose a basis of V as in 8.29. With respect to that basis, T has an upper-triangular matrix with the diagonal of the matrix containing precisely the eigenvalues of T,

2 *L*

each repeated according to its multiplicity. Thus Example 10.35 tells us that det T det .T /, where the matrix is with respect to that basis.

D *M*

D *M*

Our goal is to prove that det T det .T / for every basis of V, not just

the basis from 8.29. To do this, we will need to develop some properties of

determinants of matrices. The result below is the ﬁrst of the properties we will need.

10.36 Interchanging two columns in a matrix

Suppose A is a square matrix and B is the matrix obtained from A by interchanging two columns. Then

det A D — det B:

Proof Think of the sum deﬁning det A in 10.33 and the corresponding sum deﬁning det B. The same products of Aj;k’s appear in both sums, although they correspond to different permutations. The permutation corresponding to

a given product of Aj;k’s when computing det B is obtained by interchanging two entries in the corresponding permutation when computing det A, thus multiplying the sign of the permutation by —1 (see 10.32). Hence we see that det A D — det B.

If T .V / and the matrix of T (with respect to some basis) has two equal columns, then T is not injective and hence det T 0. Although this comment makes the next result plausible, it cannot be used in the proof,

D

2 *L*

because we do not yet know that det T D det *M*.T / for every choice of basis.

10.37 Matrices with two equal columns

If A is a square matrix that has two equal columns, then det A D 0.

Proof Suppose A is a square matrix that has two equal columns. Interchang- ing the two equal columns of A gives the original matrix A. Thus from 10.36 (with B D A), we have

det A D — det A;

which implies that det A D 0.

**316** CHAPTER 10 Trace and Determinant

Recall from 3.44 that if A is an n-by-n matrix

B0 A1;1 ::: A1;n 1C

A D @

:

:

A ;

An;1 ::: An;n

then we can think of the kth column of A as an n-by-1 matrix denoted A ;k:

0

BA D

A

;k @

A1;k

:

An;k

1C :

Note that Aj;k, with two subscripts, de- notes an entry of A, whereas A ;k, with a dot as a placeholder and one subscript,

*Some books deﬁne the determinant to be the function deﬁned on the square matrices that is linear as a function of each column sepa- rately and that satisﬁes 10.38 and*

det I D 1*. To prove that such a*

*function exists and that it is unique takes a nontrivial amount of work.*

denotes a column of A. This notation allows us to write A in the form

. A ;1 ::: A ;n /;

which will be useful.

The next result shows that a permutation of the columns of a matrix changes the determinant by a factor of the sign of the permutation.

10.38 Permuting the columns of a matrix

is a permutation. Then

det. A ;m*1* ::: A ;m*n* / D （sign.m1;:::; mn/） det A:

Suppose A D . A ;1 ::: A ;n / is an n-by-n matrix and .m1;:::; mn/

Proof We can transform the matrix . A ;m*1* ::: A ;m*n* / into A through a

series of steps. In each step, we interchange two columns and hence multiply the determinant by 1 (see 10.36). The number of steps needed equals the number of steps needed to transform the permutation .m1;:::; mn/ into the permutation .1; : : : ; n/ by interchanging two entries in each step. The proof is completed by noting that the number of such steps is even if .m1;:::; mn/ has sign 1, odd if .m1;:::; mn/ has sign 1 (this follows from 10.32, along with the observation that the permutation .1; : : : ; n/ has sign 1).

—

—

SECTION 10.B Determinant **317**

The next result about determinants will also be useful.

10.39 Determinant is a linear function of each column

Suppose k; n are positive integers with 1 三 k 三 n. Fix n-by-1 matrices

vector A ;k to

A ;1;:::; A ;n except A ;k. Then the function that takes an n-by-1 column

det. A ;1 :::

A ;k

:::

A ;n /

is a linear map from the vector space of n-by-1 matrices with entries in **F**

to **F**.

Proof The linearity follows easily from 10.33, where each term in the sum contains precisely one entry from the kth column of A.

Now we are ready to prove one of the key properties about determinants of square matrices. This property will enable us to connect the determinant of an operator with the determinant of its

*The result below was ﬁrst proved in 1812 by French mathematicians Jacques Binet and Augustin-Louis Cauchy.*

matrix. Note that this proof is considerably more complicated than the proof of the corresponding result about the trace (see 10.14).

10.40 Determinant is multiplicative

Suppose A and B are square matrices of the same size. Then det.AB/ D det.BA/ D .det A/.det B/:

Proof Write A . A ;1 ::: A ;n /, where each A ;k is an n-by-1 column

D

of A. Also write

B0 B1;1 ::: B1;n 1C

B D @

:

:

A D . B ;1 ::: B ;n /;

Bn;1 ::: Bn;n

where each B ;k is an n-by-1 column of B. Let ek denote the n-by-1 matrix that equals 1 in the kth row and 0Pelsewhere. Note that Aek D A ;k and

Bek D B ;k. Furthermore, B ;k D

n

Bm;kem.

First we will prove det.AB/ D mD1

.det A/.

det

B/. As we observed ear-

lier (see 3.49), the deﬁnition of matrix multiplication easily implies that

AB D . AB ;1 ::: AB ;n /. Thus

**318** CHAPTER 10 Trace and Determinant

det.AB/ D det. AB ;1 ::: AB ;n /

D det. A.Pn Bm ;1em / ::: A.Pn

m*1*D1

*1*

*1*

m*n*D1

*n*

*n*

Bm ;nem / /

D det. Pn

Bm ;1Aem

::: Pn

Bm ;nAem /

m*1*D1 *1 1*

Bm*1*;1 ... Bm*n*;n det. Aem*1* ::: Aem*n* /;

Xn Xn

D

...

m*1*D1

m*n*D1

m*n*D1 *n n*

where the last equality comes from repeated applications of the linearity of det

as a function of one column at a time (10.39). In the last sum above, all terms in which mj mk for some j k can be ignored, because the determinant of a matrix with two equal columns is 0 (by 10.37). Thus instead of summing

D ¤

over all m1;:::; mn with each mj taking on values 1; :::; n, we can sum just over the permutations, where the mj ’s have distinct values. In other words,

det.AB/ D X

.m ;:::;m /2perm n*n1*

D .m*1*;:::;mX*n*/2perm n

Bm*1*;1 ... Bm*n*;n det. Aem*1* ::: Aem*n* /

Bm*1*;1 ... Bm*n*;n（sign.m1;:::; mn/） det A

D .det A/ X （sign.m1;:::; mn/）Bm*1*;1 ... Bm*n*;n

.m*1*;:::;m*n*/2perm n

D .det A/.det B/;

where the second equality comes from 10.38.

In the paragraph above, we proved that det.AB/ .det A/.det B/. In- terchanging the roles of A and B, we have det.BA/ .det B/.det A/. The last equation can be rewritten as det.BA/ .det A/.det B/, completing the

D

D

D

proof.

Now we can prove that the determi- nant of the matrix of an operator is in- dependent of the basis with respect to which the matrix is computed.

*Note the similarity of the proof of the next result to the proof of the analogous result about the trace* (*see 10.15*)*.*

10.41 Determinant of matrix of operator does not depend on basis

Let T 2 *L*.V /. Suppose u1;:::; un and *v*1;:::; *v*n are bases of V. Then

det *M*（T; .u1;:::; un/） D det *M*（T; .*v*1;:::; *v*n/）:

SECTION 10.B Determinant **319**

Proof Let A D *M*（I; .u1;:::; un/; .*v*1;:::; *v*n/）. Then

（ ） （ （ ） ）*M* D *M*

det T; .u1;:::; un/ det A 1 T; .*v*1;:::; *v*n/ A

D det （*M*（T; .*v*1;:::; *v*n/）A）A 1

（ ）D det *M* T;.*v* ;:::; *v* / ;1 n

where the ﬁrst equality follows from 10.7 and the second equality follows from 10.40. The third equality completes the proof.

The result below states that the determinant of an operator equals the determinant of the matrix of the operator. This theorem does not specify a basis because, by the result above, the determinant of the matrix of an operator is the same for every choice of basis.

10.42 Determinant of an operator equals determinant of its matrix

Suppose T 2 *L*.V /. Then det T D det *M*.T /.

Proof As noted above, 10.41 implies that det .T / is independent of which basis of V we choose. Thus to show that det T det .T / for every basis of V, we need only show that the result holds for some basis of V.

As we have already discussed, if V is a complex vector space, then choos- ing a basis of V as in 8.29 gives the desired result. If V is a real vector space, then applying the complex case to the complexiﬁcation T**C** (which is used to deﬁne det T ) gives the desired result.

D *M*

*M*

If we know the matrix of an operator on a complex vector space, the result above allows us to ﬁnd the product of all the eigenvalues without ﬁnding any of the eigenvalues.

10.43 **Example** Suppose T is the operator on **C**5 whose matrix is

0 0 0 0 —3

0B 1C

1 0 0 0 6

0 1 0 0 0

B@

:

CA

0 0 1 0 0

0 0 0 1 0

No one knows an exact formula for any of the eigenvalues of this operator. However, we do know that the product of the eigenvalues equals —3, because the determinant of the matrix above equals —3.

**320** CHAPTER 10 Trace and Determinant

We can use 10.42 to give easy proofs of some useful properties about determinants of operators by shifting to the language of determinants of matrices, where certain properties have already been proved or are obvious. We carry out this procedure in the next result.

10.44 Determinant is multiplicative

Suppose S; T 2 *L*.V /. Then

det.ST / D det.TS/ D .det S/.det T /:

Proof Choose a basis of V. Then

det.ST / D det *M*.ST /

（ ）

D d（ et *M*.S/*M*）（ .T / ）

D

det *M*.S/

det *M*.T /

D .det S/.det T /;

where the ﬁrst and last equalities come from 10.42 and the third equality comes from 10.40.

In the paragraph above, we proved that det.ST / .det S/.det T /. Inter- changing the roles of S and T, we have det.TS/ .det T /.det S /. Because

D

D

multiplication of elements of **F** is commutative, the last equation can be rewritten as det.TS/ D .det S/.det T /, completing the proof.

### The Sign of the Determinant

We proved the basic results of linear algebra before introducing determinants in this ﬁnal chapter. Although determinants have value as a research tool in more advanced subjects, they play little role in basic linear algebra (when the subject is done right).

Determinants do have one important application in undergraduate mathemat- ics, namely, in computing certain vol- umes and integrals. In this subsection we interpret the meaning of the sign of

*Most applied mathematicians agree that determinants should rarely be used in serious numeric calculations.*

the determinant on a real vector space. Then in the ﬁnal subsection we will use the linear algebra we have learned to make clear the connection between determinants and these applications. Thus we will be dealing with a part of analysis that uses linear algebra.

SECTION 10.B Determinant **321**

We will begin with some purely linear algebra results that will also be useful when investigating volumes. Our setting will be inner product spaces. Recall that an isometry on an inner product space is an operator that preserves norms. The next result shows that every isometry has determinant with

absolute value 1.

10.45 Isometries have determinant with absolute value 1

Suppose V is an inner product space and S 2 *L*.V / is an isometry. Then

jdet SjD 1.

Proof First consider the case where V is a complex inner product space. Then all the eigenvalues of S have absolute value 1 (see the proof of 7.43). Thus the product of the eigenvalues of S, counting multiplicity, has absolute value one. In other words, det S 1, as desired.

Now suppose V is a real inner product space. We present two different

j jD

proofs in this case.

Proof 1: With respect to the inner product on the complexiﬁcation V**C** given by Exercise 3 in Section 9.B, it is easy to see that S**C** is an isometry on V**C**. Thus by the complex case that we have already done, we have jdet S**C**jD 1. By deﬁnition of the determinant on real vector spaces, we have det S det S**C** and thus det S 1, completing the proof.

j jD

D

Proof 2: By 9.36, there is an orthonormal basis of V with respect to which

*M*.S/ is a block diagonal matrix, where each block on the diagonal is a

1-by-1 matrix containing 1 or —1 or a 2-by-2 matrix of the form

—

cos 0 sin 0

sin 0 cos 0 ;

with 0 .0; 兀/. Note that the determinant of each 2-by-2 matrix of the form above equals 1 (because cos2 0 sin2 0 1). Thus the determinant of S, which is the product of the determinants of the blocks (see Exercise 6), is the

C D

2

product of 1’s and —1’s. Hence, jdet SjD 1, as desired.

The Real Spectral Theorem 7.29 states that a self-adjoint operator T on a real inner product space has an orthonormal basis consisting of eigenvectors. With respect to such a basis, the number of times each eigenvalue appears on

the diagonal of .T / is its multiplicity. Thus det T equals the product of its

*M*

eigenvalues, counting multiplicity (of course, this holds for every operator,

self-adjoint or not, on a complex vector space).

**322** CHAPTER 10 Trace and Determinant

Recall that if V is an inner product space and T 2 *L*.V /, then T \*pT is a

positive operator and hence haspa unique positive square root, denoted T \*T

(see 7.35 and 7.36). Because T \*T is popsitive, all its eigenvalues are non-

negative (again, see 7.35), and hence det play a role in next example.

T \*T 乏 0. These considerations

10.46 **Example** Suppose V is a real inner product space and T .V / is invertible (and thus det T is either positive or negative). Attach a geometric meaning to the sign of det T.

2 *L*

Solution First we consider an isometry S 2 *L*.V /. By 10.45, the determinant of S equals 1 or —1. Note that

f*v* 2 V W S *v* D —*v*g

is the eigenspace E. 1; S/. Thinking geometrically, we could say that this is the subspace on which S reverses direction. An examination of proof 2 of 10.45 shows that det S 1 if this subspace has even dimension and det S 1 if this subspace has odd dimension.

*We are not formally deﬁning the phrase “reverses direction” be- cause these comments are meant only as an intuitive aid to our un- derstanding.*

—

D

D —

Returning to our arbitrary invertible operator T 2 *L*.V /, by the Polar Decomposition (7.45) there is an isometry S 2 *L*.V / such that

T D SpT \*T:

Now 10.44 tells us that

det T D .det S/.det pT \*T /:

The remarks just before this example pointed out that det pT \*T 0. Thus whether det T is positive or negative depends on whether det S is positive or

乏

negative. As we saw in the paragraph above, this depends on whether the subspace on which S reverses direction has even or odd dimension.

Becauspe T is the product of S and an operator that never reverses direction

(namely, T \*T ), we can reasonably say that whether det T is positive or

negative depends on whether T reverses vectors an even or an odd number of

times.

### Volume

SECTION 10.B Determinant **323**

The next result will be a key tool in our investigation ofpvolume. Recall that

our remarks before Example 10.46 pointed out that det T \*T 乏 0.

10.47 jdet T jD det pT \*T

Suppose V is an inner product space and T 2 *L*.V /. Then

jdet T jD det pT \*T:

Proof

By the Polar Decomposition (7.45), there is an isometry S 2 *L*.V / such

*Another proof of this result is sug- gested in Exercise 8.*

that

Thus

T D SpT

\*T:

jdet T jD jdet Sj det pT \*T

D

D

D det pT \*T;

where the ﬁrst equality follows from 10.44 and the second equality follows from 10.45.

Now we turn to the question of volume in **R**n. Fix a positive integer n for the rest of this subsection. We will consider only the real inner product space **R**n, with its standard inner product.

We would like to assign to each subset Q of **R**n its n-dimensional volume (when n 2, this is usually called area instead of volume). We begin with boxes, where we have a good intuitive notion of volume.

D

10.48 **Deﬁnition *box***

A ***box*** in **R**n is a set of the form

f.y1;:::; yn/ 2 **R** W xj < yj < xj C rj for j D 1; :::; ng;

n

where r1;:::; rn are positive numbers and .x1;:::; xn/ 2 **R**n. The num-

bers r1;:::; rn are called the ***side lengths*** of the box.

You should verify that when n 2, a box is a rectangle with sides parallel to the coordinate axes, and that when n 3, a box is a familiar 3-dimensional box with sides parallel to the coordinate axes.

**324** CHAPTER 10 Trace and Determinant

2 *L* c

The next deﬁnition ﬁts with our intuitive notion of volume, because we deﬁne the volume of a box to be the product of the side lengths of the box.

10.49 **Deﬁnition *volume of a box***

The ***volume*** of a box B in **R**n with side lengths r1;:::; rn is deﬁned to be r1 ... rn and is denoted by volume B.

To deﬁne the volume of an arbitrary set Q **R**n, the idea is to write Q as a subset of a union of many small boxes,

then add up the volumes of these small boxes. As we approximate Q more accurately by unions of small boxes, we get a better estimate of volume Q.

*Readers familiar with outer mea- sure will recognize that concept here.*

c

10.50 **Deﬁnition *volume***

to be the inﬁmum of

volume B1 C volume B2 C ... ;

where the inﬁmum is taken over all sequences B1; B2;::: of boxes in **R**n whose union contains Q.

Suppose Q c **R**n. Then the ***volume*** of Q, denoted volume Q, is deﬁned

We will work only with an intuitive notion of volume. Our purpose in this book is to understand linear algebra, whereas notions of volume belong to analysis (although volume is intimately connected with determinants, as we will soon see). Thus for the rest of this section we will rely on intuitive notions of volume rather than on a rigorous development, although we shall maintain our usual rigor in the linear algebra parts of what follows. Everything said here about volume will be correct if appropriately interpreted—the intuitive approach used here can be converted into appropriate correct deﬁnitions, correct statements, and correct proofs using the machinery of analysis.

10.51 **Notation** T .Q/

For T a function deﬁned on a set Q, deﬁne T .Q/ by

T .Q/ D fTx W x 2 Qg:

For T .**R**n/ and Q **R**n, we seek a formula for volume T .Q/ in terms of T and volume Q. We begin by looking at positive operators.

SECTION 10.B Determinant **325**

10.52 Positive operators change volume by factor of determinant

Suppose T 2 *L*.**R**n/ is a positive operator and Q c **R**n. Then volume T .Q/ D .det T /.volume Q/:

Proof To get a feeling for why this result is true, ﬁrst consider the special case where 入1;:::; 入n are positive numbers and T 2 *L*.**R**n/ is deﬁned by

T .x1;:::; xn/ D .入1x1;:::; 入nxn/:

This operator stretches the j th standard basis vector by a factor of 入j . If B is a box in **R**n with side lengths r1;:::; rn, then T .B/ is a box in **R**n with side lengths 入1r; :::; 入nr. The box T .B/ thus has volume 入1 ... 入nr1 ... rn, whereas the box Q has volume r1 ... rn. Note that det T D 入1 ... 入n. Thus

volume T .B/ D .det T /.volume B/

for every box B in **R**n. Because the volume of Q is approximated by sums of volumes of boxes, this implies that volume T .Q/ .det T /.volume Q/.

D

Now consider an arbitrary positive operator T .**R**n/. By the Real Spectral Theorem (7.29), there exist an orthonormal basis e1;:::; en of **R**n and nonnegative numbers 入1;:::; 入n such that T ej 入j ej for j 1; :::; n. In the special case where e1;:::; en is the standard basis of **R**n, this operator

2 *L*

D D

is the same one as deﬁned in the paragraph above. For an arbitrary orthonor- mal basis e1;:::; en, this operator has the same behavior as the one in the paragraph above—it stretches the j th basis vector in an orthonormal basis by a factor of 入j . Your intuition about volume should convince you that volume

behaves the same with respect to each orthonormal basis. That intuition, and the special case of the paragraph above, should convince you that T multiplies volume by a factor of 入1 ... 入n, which again equals det T.

Our next tool is the following result, which states that isometries do not change volume.

10.53 An isometry does not change volume

Suppose S 2 *L*.**R**n/ is an isometry and Q c **R**n. Then

volume S.Q/ D volume Q:

**326** CHAPTER 10 Trace and Determinant

Proof For x; y 2 **R**n, we have

kSx — SykD kS.x — y/k

D kx — yk:

In other words, S does not change the distance between points. That property alone may be enough to convince you that S does not change volume.

However, if you need stronger persuasion, consider the complete descrip-

tion of isometries on real inner product spaces provided by 9.36. According to 9.36, S can be decomposed into pieces, each of which is the identity on some subspace (which clearly does not change volume) or multiplication by 1 on

—

some subspace (which again clearly does not change volume) or a rotation on a 2-dimensional subspace (which again does not change volume). Or use

9.36 in conjunction with Exercise 7 in Section 9.B to write S as a product of

operators, each of which does not change volume. Either way, you should be convinced that S does not change volume.

Now we can prove that an operator T 2 *L*.**R**n/ changes volume by a factor of jdet T j. Note the huge importance of the Polar Decomposition in the proof.

10.54 T changes volume by factor of jdet T j

Suppose T 2 *L*.**R**n/ and Q c **R**n. Then

volume T .Q/ D jdet T j.volume Q/:

Proof By the Polar Decomposition (7.45), there is an isometry S 2 *L*.V /

such that

T D SpT

\*T:

If Q c **R**n, then T .Q/ D S pT \*T .Q/ . Thus

（ ）

volume T .Q/ volume S pT \*T .Q/

（ ）D

D volume pT \*T .Q/

D .det pT \*T /.volume Q/

D jdet T j.volume Q/;

where the second equality holds because volume is not changed by the isom-

etry S (bpy 10.53), the third equality holds by 10.52 (applied to the positive

operator T \*T ), and the fourth equality holds by 10.47.

SECTION 10.B Determinant **327**

The result that we just proved leads to the appearance of determinants in the formula for change of variables in multivariable integration. To describe this, we will again be vague and intuitive.

Throughout this book, almost all the functions we have encountered have been linear. Thus please be aware that the functions f and (J in the material below are not assumed to be linear.

The next deﬁnition aims at conveying the idea of the integral; it is not intended as a rigorous deﬁnition.

10.55 **Deﬁnition *integral,*** RQ f

If Q c **R** and f

n

is a real-valued function on Q, then the ***integral*** of f

over Q, denoted Q f or Q f .x/ dx, is deﬁned by breaking Q into pieces small enough that f is almost constant on each piece. On each piece,

multiply the (almost constant) value of f by the volume of the piece, then

add up these numbers for all the pieces, getting an approximation to the integral that becomes more accurate as Q is divided into ﬁner pieces.

R

R

Actually, Q in the deﬁnition above needs to be a reasonable set (for example, open or measurable) and f needs to be a reasonable function (for

example, continuous or measurable), but we will not worry about those technicalities. Also, notice that the x in Q f .x/ dx is a dummy variable and could be replaced with any other symbol.

R

Now we deﬁne the notions of differentiable and derivative. Notice that

in this context, the derivative is an operator, not a number as in one-variable calculus. The uniqueness of T in the deﬁnition below is left as Exercise 9.

10.56 **Deﬁnition *differentiable, derivative,*** (J 0.x/

Suppose Q is an open subset of **R**n and (J is a function from Q to **R**n.

For x 2 Q, the function (J is called ***differentiable*** at x if there exists an

operator T 2 *L*.**R**n/ such that

y!0

lim k(J.x C y/ — (J.x/ — Tyk D 0:

If (J is differentiable at x, then the unique operator T 2 *L*.**R**n/ satisfying

kyk

the equation above is called the ***derivative*** of (J at x and is denoted by

(J 0.x/.

**328** CHAPTER 10 Trace and Determinant

The idea of the derivative is that for x ﬁxed and kyk small,

*If* n D 1*, then the derivative in the*

*sense of the deﬁnition above is the operator on* **R** *of multiplication by the derivative in the usual sense of one-variable calculus.*

（ ）

(J.x C y/ � (J.x/ C (J 0.x/ .y/I

because (J 0.x/ .**R**n/, this makes sense.

2 *L*

Suppose Q is an open subset of **R**n and (J is a function from Q to **R**n. We

can write

（ ）

(J.x/ D (J1.x/; : : : ; (Jn.x/ ;

where each (Jj is a function from Q to **R**. The partial derivative of (Jj with respect to the kth coordinate is denoted Dk(Jj . Evaluating this partial derivative at a point x Q gives Dk(Jj .x/. If (J is differentiable at x, then the matrix of (J 0.x/ with respect to the standard basis of **R**n contains Dk(Jj .x/ in row j , column k (this is left as an exercise). In other words,

2

（ ） 0B

**10.57** *M* (J 0.x/

D @

D1(J1.x/ : : : Dn(J1.x/ 1C

D1(Jn.x/ ::: Dn(Jn.x/

:

:

A :

Now we can state the change of variables integration formula. Some additional mild hypotheses are needed for f and (J 0 (such as continuity or measurability), but we will not worry about them because the proof below is

really a pseudoproof that is intended to convey the reason the result is true.

The result below is called a change of variables formula because you can think of y (J.x/ as a change of variables, as illustrated by the two examples that follow the proof.

D

10.58 Change of variables in an integral

Suppose Q is an open subset of **R**n and (J W Q ! **R**n is differentiable at

every point of Q. If f is a real-valued function deﬁned on (J.Q/, then

Z

a.Q/

f .y/ dy D Z f （(J.x/）jdet (J 0.x/j dx:

Q

Proof Let x 2 Q and let r be a small su（bset o）f Q containing x such that f

Adding a ﬁxed vector [such as (J.x/] to each vector in a set produces another set with the same volume. Thus our approximation for (J near x using the derivative shows that

is approximately equal to the constant f

(J.x/

on the set (J.r/.

SECTION 10.B Determinant **329**

[（ ） ]

volume (J.r/ � volume (J 0.x/ .r/ :

Using 10.54 applied to the operator (J 0.x/, this becomes

volume (J.r/ � jdet (J 0.x/j.volume r/:

Let y D (J.x/. M（ ultip）ly the left side of the equation above by f .y/ and the

（ ）

right side by f

getting

(J.x/

[because y D (J.x/, these two quantities are equal],

f .y/ volume (J.r/ � f (J.x/ jdet (J 0.x/j.volume r/:

Now break Q into many small pieces and add the corresponding versions of the equation above, getting the desired result.

The key point when making a change of variables is that the factor of det (J 0.x/ must be included when making a substitution y f .x/, as in the right side of 10.58. We ﬁnish up by illustrating this point with two important

j j D

examples.

* 1. **Example *polar coordinates***

Deﬁne (J W **R**2 ! **R**2 by

(J.r; 0/ D .r cos 0; r sin 0 /;

where we have used r; 0 as the coordinates instead of x1; x2 for reasons that will be obvious to everyone familiar with polar coordinates (and will be a mystery to everyone else). For this choice of (J , the matrix of partial derivatives corresponding to 10.57 is

—

cos 0 r sin 0 sin 0 r cos 0 ;

as you should verify. The determinant of the matrix above equals r, thus explaining why a factor of r is needed when computing an integral in polar coordinates.

For example, note the extra factor of r in the following familiar formula involving integrating a function f over a disk in **R**2:

Z 1 Z p1 x*2* Z 2冗 Z 1

0

1 p

1 x

*2* f .x; y/ dy dx D

f .r cos 0; r sin 0 /r dr d0:

0

**330** CHAPTER 10 Trace and Determinant

* 1. **Example *spherical coordinates***

Deﬁne (J W **R**3 ! **R**3 by

(J.p; '; 0/ D .p sin ' cos 0; p sin ' sin 0; p cos '/;

where we have used p; 0; ' as the coordinates instead of x1; x2; x3 for reasons that will be obvious to everyone familiar with spherical coordinates (and will be a mystery to everyone else). For this choice of (J , the matrix of partial

derivatives corresponding to 10.57 is

0@ —

sin ' cos 0 p cos ' cos 0 p sin ' sin 0 sin ' sin 0 p cos ' sin 0 p sin ' cos 0

cos ' —p sin ' 0

A1 ;

as you should verify. The determinant of the matrix above equals p2 sin ', thus explaining why a factor of p2 sin ' is needed when computing an integral in spherical coordinates.

For example, note the extra factor of p2 sin ' in the following familiar formula involving integrating a function f over a ball in **R**3:

Z 1 Z p1 x*2* Z p1 x*2* y*2*

p *2* p

f .x; y; z/ dz dy dx

*2 2*

1 1 x 1 x y

Z 2冗 Z 冗 Z 1

D

f .p sin ' cos 0; p sin ' sin 0; p cos '/p2

sin ' dp d' d0:

0 0 0

EXERCISES 10.B

2 *L*

1. Suppose V is a real vector space. Suppose T .V / has no eigenvalues. Prove that det T > 0.
2. Suppose V is a real vector space with even dimension and T .V /. Suppose det T < 0. Prove that T has at least two distinct eigenvalues.

2 *L*

1. Suppose T 2 *L*.V / and n D dim V > 2. Let 入1;:::; 入n denote the eigenvalues of T (or of T**C** if V is a real vector space), repeated according to multiplicity.
   1. Find a formula for the coefﬁcient of zn 2 in the characteristic polynomial of T in terms of 入1;:::; 入n.
   2. Find a formula for the coefﬁcient of z in the characteristic polyno- mial of T in terms of 入1;:::; 入n.

SECTION 10.B Determinant **331**

1. Suppose T 2 *L*.V / and c 2 **F**. Prove that det.cT / D cdim V det T.
2. Prove or give a counterexample: if S; T 2 *L*.V /, then det.S C T/ D

det S C det T.

1. Suppose A is a block upper-triangular matrix

0B A1 \* 1C

A D @

: : :

0 Am

A ;

where each Aj along the diagonal is a square matrix. Prove that

det A D .det A1/ det Am/:

1. Suppose A is an n-by-n matrix with real entries. Let S 2 *L*.**C**n/ denote the operator on **C**n whose matrix equals A, and let T 2 *L*.**R**n/ denote the operator on **R**n whose matrix equals A. Prove that trace S D trace T and det S D det T.
2. Suppose V is an inner product space and T 2 *L*.V /. Prove that

det T \* D det T:

Use this to prove that det T det pT \*T , giving a different proof than

j jD

was given in 10.47.

1. Suppose Q is an open subset of **R**n and (J is a function from Q to **R**n. Suppose x Q and (J is differentiable at x. Prove that the operator T .**R**n/ satisfying the equation in 10.56 is unique.

2 *L*

2

[*This exercise shows that the notation* (J 0.x/ *is justiﬁed.*]

1. Suppose T 2 *L*.**R**n/ and x 2 **R**n. Prove that T is differentiable at x and

T 0.x/ D T.

1. Find a suitable hypothesis on (J and then prove 10.57.
2. Let a; b; c be positive numbers. Find the volume of the ellipsoid

n 3 x2 y2 z2 o

.x; y; z/ 2 **R**

W a2 C b2 C c2 <1

by ﬁnding a set Q c **R**3 whose volume you know and an operator

T 2 *L*.**R**3/ such that T .Q/ equals the ellipsoid above.

page 1: Pierre Louis Dumesnil; 1884 copy by Nils Forsberg/Public domain image from *Wikimedia*.

*Photo Credits*

•

page 27: George M. Bergman/Archives of the Mathematisches Forschungsinstitut Oberwolfach.

•

page 51: Gottlieb Biermann; photo by A. Wittmann/Public domain image from *Wikimedia*.

•

* page 117: Mostafa Azizi/Public domain image from *Wikimedia*.
* page 131: Hans-Peter Postel/Public domain image from *Wikimedia*.
* page 163: Public domain image from *Wikimedia*.

page 203: Public domain image from *Wikimedia*. Original painting is in Tate Britain.

•

* page 224: Spiked Math.
* page 241: Public domain image from *Wikimedia*.

page 275: Public domain image from *Wikimedia*. Original fresco is in the Vatican.

•

* page 295: Public domain image from *Wikimedia*.

© Springer International Publishing 2015

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, DOI 10.1007/978-3-319-11080-6

333

A 1, 296

*Symbol Index*

Aj; , 76

Aj;k, 70

A ;k, 76

At, 109

**C**,2

deg, 31

�, 179

det, 307, 314

dim, 44

, 21

˚

Dk, 328

E.入; T /, 155

**F**,4

**F**1, 13

**F**m;n, 73

**F**n,6

**F**S , 14

G.入; T /, 245

I , 52, 296

()

, 207

Im, 118

—1, 31

R

Q f , 327

*L*.V /, 86

*L*.V; W /, 52

*M*.T /, 70, 146

*M*.*v*/, 84

perm, 311

© Springer International Publishing 2015

.**F**/, 30

兀, 97

*P*

m.**F**/, 31

*P*

p.T /, 143

PU , 195

**R**,2

Re, 118

(J 0, 327

¨, 243

TQ , 97

p

T , 233

T 0, 103

T \*, 204

T 1, 80

T .Q/, 324

T**C**, 277

T m, 143

T U , 132, 137

j

T=U , 137

U ?, 193

U 0, 104

hu; *v*i, 166

V , 16

jj jj

*v* , 168

V 0, 101

V=U , 95

—*v*, 15

V**C**, 276

*v* C U , 94

zN, 118

jzj, 118

335

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, DOI 10.1007/978-3-319-11080-6

absolute value, 118 addition

*Index*

in quotient space, 96 of complex numbers, 2 of functions, 14

of linear maps, 55 of matrices, 72

of subspaces, 20

of vectors, 12

of vectors in **F**n,7 additive inverse

in **C**, 3,4

in **F**n,9

in vector space, 12, 15

additivity, 52

adjoint of a linear map, 204 afﬁne subset, 94

algebraic multiplicity, 255 annihilator of a subspace, 104 Apollonius’s Identity, 179

associativity, 3, 12, 56

backward shift, 53, 59, 81, 86, 140

basis, 39

of eigenvectors, 157, 218,

221, 224, 268

of generalized eigenvectors, 254

Binet, Jacques, 317

Blake, William, 203

block diagonal matrix, 255

© Springer International Publishing 2015

box in **R**n, 323

Cauchy, Augustin-Louis, 171, 317

Cauchy–Schwarz Inequality, 172

Cayley, Arthur, 262 Cayley–Hamilton Theorem

on complex vector space, 261 on real vector space, 284

change of basis, 298

change of variables in integral, 328

characteristic polynomial

on complex vector space, 261 on real vector space, 283

characteristic value, 134 Christina, Queen of Sweden, 1 closed under addition, 18

closed under scalar multiplication, 18

column rank of a matrix, 111 commutativity, 3, 7, 12, 25, 56, 75,

79, 144, 212

complex conjugate, 118

complex number, 2

Complex Spectral Theorem, 218 complex vector space, 13 complexiﬁcation

of a vector space, 276 of an operator, 277

conjugate symmetry, 166

337

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, DOI 10.1007/978-3-319-11080-6

**338** Index

conjugate transpose of a matrix, 207

coordinate, 6

cube root of an operator, 224 cubic formula, 124

degree of a polynomial, 31 derivative, 327

Descartes, René, 1 determinant

of a matrix, 314

of an operator, 307 diagonal matrix, 155

diagonal of a square matrix, 147 diagonalizable, 156

differentiable, 327

differentiation linear map, 53, 56,

59, 61, 62, 69, 72, 78,

144, 190, 248, 294

dimension, 44

of a sum of subspaces, 47 direct sum, 21, 42, 93

of a subspace and its orthogonal complement, 194

of null T n and range T n, 243

distributive property, 3, 12, 16, 56,

79

Division Algorithm for

Polynomials, 121 division of complex numbers, 4 dot product, 164

double dual space, 116 dual

of a basis, 102

of a linear map, 103 of a vector space, 101

eigenspace, 155

eigenvalue of an operator, 134 eigenvector, 134

Euclid, 275

Euclidean inner product, 166

factor of a polynomial, 122 Fibonacci, 131

Fibonacci sequence, 161

ﬁeld, 10

ﬁnite-dimensional vector space, 30

*Flatland*,6

Fundamental Theorem of Algebra, 124

Fundamental Theorem of Linear Maps, 63

Gauss, Carl Friedrich, 51 generalized eigenspace, 245

generalized eigenvector, 245

geometric multiplicity, 255

Gram, Jørgen, 182

Gram–Schmidt Procedure, 182 graph of a linear map, 98

Halmos, Paul, 27

Hamilton, William, 262

harmonic function, 179

Hermitian, 209

homogeneity, 52 homogeneous system of linear

equations, 65, 90

Hypatia, 241

identity map, 52, 56

identity matrix, 296

image, 62

imaginary part, 118

inﬁnite-dimensional vector space, 31

inhomogeneous system of linear equations, 66, 90

injective, 60

inner product, 166

inner product space, 167 integral, 327

invariant subspace, 132

of T 0, 110

of T \*, 208

of vector, 84

Index **339**

inverse

of a linear map, 80 of a matrix, 296

invertible linear map, 80 invertible matrix, 296

isometry, 228, 292, 321 isomorphic vector spaces, 82 isomorphism, 82

Jordan basis, 273

Jordan Form, 273

Jordan, Camille, 272

kernel, 59

Khayyám, Omar, 117

Laplacian, 179 length of list, 5

Leonardo of Pisa, 131 linear combination, 28

Linear Dependence Lemma, 34 linear functional, 101, 187

linear map, 52

linear span, 29

linear subspace, 18

linear transformation, 52

linearly dependent, 33

linearly independent, 32

list, 5

of vectors, 28

Lovelace, Ada, 295

matrix, 70

multiplication, 75 of linear map, 70

of nilpotent operator, 249 of operator, 146

of product of linear maps, 75, 297

minimal polynomial, 263, 279

minimizing distance, 198

monic polynomial, 262 multiplication, ***see*** product multiplicity of an eigenvalue, 254

Newton, Isaac, 203

nilpotent operator, 248, 271

nonsingular matrix, 296

norm, 164, 168

normal operator, 212, 287

null space, 59

of powers of an operator, 242 of T 0, 106

of T \*, 207

one-to-one, 60

onto, 62

operator, 86 orthogonal

complement, 193

operator, 229

projection, 195

vectors, 169 orthonormal

basis, 181

list, 180

parallel afﬁne subsets, 94 Parallelogram Equality, 174

permutation, 311

photo credits, 333

point, 13

polar coordinates, 329

Polar Decomposition, 233

polynomial, 30

positive operator, 225

positive semideﬁnite operator, 227

**340** Index

product

of complex numbers, 2 of linear maps, 55

of matrices, 75

of polynomials, 144

of scalar and linear map, 55 of scalar and vector, 12

of scalar and vector in **F**n, 10

of vector spaces, 91 Pythagorean Theorem, 170

quotient

map, 97

operator, 137

space, 95

range, 61

of powers of an operator, 251 of T 0, 107

of T \*, 207

rank of a matrix, 112

Raphael, 275

real part, 118

Real Spectral Theorem, 221 real vector space, 13 restriction operator, 137

Riesz Representation Theorem, 188

Riesz, Frigyes, 187

row rank of a matrix, 111

scalar, 4

scalar multiplication, 10, 12 in quotient space, 96

of linear maps, 55 of matrices, 73

Schmidt, Erhard, 182 *School of Athens*, 275 Schur’s Theorem, 186

Schur, Issai, 186

Schwarz, Hermann, 171

self-adjoint operator, 209 sign of a permutation, 312 signum, 313

singular matrix, 296

Singular Value Decomposition, 237

singular values, 236

span, 29

spans, 30

Spectral Theorem, 218, 221

spherical coordinates, 330 square root of an operator, 223,

225, 233, 259

standard basis, 39

subspace, 18

subtraction of complex numbers, 4 sum, ***see*** addition

Supreme Court, 174

surjective, 62

trace

of a matrix, 300

of an operator, 299 transpose of a matrix, 109, 207

Triangle Inequality, 173

tuple, 5

unitary operator, 229

upper-triangular matrix, 147, 256

vector, 8, 13

vector space, 12

volume, 324

zero of a polynomial, 122