LINEAR ALGEBRA

Second Edition

KENNETH HOFFMAN
Professor of Mathematics
Massachusetts Institute of Technology

RAY KUNZE
Professor of Mathematics
University of California, Irvine

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*

Current printing (last digit):

10 9 8 7 6

Library of Congress Catalog Card No. 75-142120
Printed in the United States of America

Preface

Our original purpose in writing this book was to provide a text for the undergraduate linear algebra course at the Massachusetts Institute of Technology. This course was designed for mathematics majors at the junior level, although three-fourths of the students were drawn from other scientific and technological disciplines and ranged from freshmen through graduate students. This description of the M.I.T. audience for the text remains generally accurate today. The ten years since the first edition have seen the proliferation of linear algebra courses throughout the country and have afforded one of the authors the opportunity to teach the basic material to a variety of groups at Brandeis University, Washington University (St. Louis), and the University of California (Irvine).

Our principal aim in revising *Linear Algebra* has been to increase the variety of courses which can easily be taught from it. On one hand, we have structured the chapters, especially the more difficult ones, so that there are several natural stopping points along the way, allowing the instructor in a one-quarter or one-semester course to exercise a considerable amount of choice in the subject matter. On the other hand, we have increased the amount of material in the text, so that it can be used for a rather comprehensive one-year course in linear algebra and even as a reference book for mathematicians.

The major changes have been in our treatments of canonical forms and inner product spaces. In Chapter 6 we no longer begin with the general spatial theory which underlies the theory of canonical forms. We first handle characteristic values in relation to triangulation and diagonalization theorems and then build our way up to the general theory. We have split Chapter 8 so that the basic material on inner product spaces and unitary diagonalization is followed by a Chapter 9 which treats sesqui-linear forms and the more sophisticated properties of normal operators, including normal operators on real inner product spaces.

We have also made a number of small changes and improvements from the first edition. But the basic philosophy behind the text is unchanged.

We have made no particular concession to the fact that the majority of the students may not be primarily interested in mathematics. For we believe a mathematics course should not give science, engineering, or social science students a hodgepodge of techniques, but should provide them with an understanding of basic mathematical concepts.

On the other hand, we have been keenly aware of the wide range of backgrounds which the students may possess and, in particular, of the fact that the students have had very little experience with abstract mathematical reasoning. For this reason, we have avoided the introduction of too many abstract ideas at the very beginning of the book. In addition, we have included an Appendix which presents such basic ideas as set, function, and equivalence relation. We have found it most profitable not to dwell on these ideas independently, but to advise the students to read the Appendix when these ideas arise.

Throughout the book we have included a great variety of examples of the important concepts which occur. The study of such examples is of fundamental importance and tends to minimize the number of students who can repeat definition, theorem, proof in logical order without grasping the meaning of the abstract concepts. The book also contains a wide variety of graded exercises (about six hundred), ranging from routine applications to ones which will extend the very best students. These exercises are intended to be an important part of the text.

Chapter 1 deals with systems of linear equations and their solution by means of elementary row operations on matrices. It has been our practice to spend about six lectures on this material. It provides the student with some picture of the origins of linear algebra and with the computational technique necessary to understand examples of the more abstract ideas occurring in the later chapters. Chapter 2 deals with vector spaces, subspaces, bases, and dimension. Chapter 3 treats linear transformations, their algebra, their representation by matrices, as well as isomorphism, linear functionals, and dual spaces. Chapter 4 defines the algebra of polynomials over a field, the ideals in that algebra, and the prime factorization of a polynomial. It also deals with roots, Taylor's formula, and the Lagrange interpolation formula. Chapter 5 develops determinants of square matrices, the determinant being viewed as an alternating n-linear function of the rows of a matrix, and then proceeds to multilinear functions on modules as well as the Grassman ring. The material on modules places the concept of determinant in a wider and more comprehensive setting than is usually found in elementary textbooks. Chapters 6 and 7 contain a discussion of the concepts which are basic to the analysis of a single linear transformation on a finite-dimensional vector space; the analysis of characteristic (eigen) values, triangulable and diagonalizable transformations; the concepts of the diagonalizable and nilpotent parts of a more general transformation, and the rational and Jordan canonical forms. The primary and cyclic decomposition theorems play a central role, the latter being arrived at through the study of admissible subspaces. Chapter 7 includes a discussion of matrices over a polynomial domain, the computation of invariant factors and elementary divisors of a matrix, and the development of the Smith canonical form. The chapter ends with a discussion of semi-simple operators, to round out the analysis of a single operator. Chapter 8 treats finite-dimensional inner product spaces in some detail. It covers the basic geometry, relating orthogonalization to the idea of 'best approximation to a vector' and leading to the concepts of the orthogonal projection of a vector onto a subspace and the orthogonal complement of a subspace. The chapter treats unitary operators and culminates in the diagonalization of self-adjoint and normal operators. Chapter 9 introduces sesqui-linear forms, relates them to positive and self-adjoint operators on an inner product space, moves on to the spectral theory of normal operators and then to more sophisticated results concerning normal operators on real or complex inner product spaces. Chapter 10 discusses bilinear forms, emphasizing canonical forms for symmetric and skew-symmetric forms, as well as groups preserving non-degenerate forms, especially the orthogonal, unitary, pseudo-orthogonal and Lorentz groups.

We feel that any course which uses this text should cover Chapters 1, 2, and 3 thoroughly, possibly excluding Sections 3.6 and 3.7 which deal with the double dual and the transpose of a linear transformation. Chapters 4 and 5, on polynomials and determinants, may be treated with varying degrees of thoroughness. In fact, polynomial ideals and basic properties of determinants may be covered quite sketchily without serious damage to the flow of the logic in the text; however, our inclination is to deal with these chapters carefully (except the results on modules), because the material illustrates so well the basic ideas of linear algebra. An elementary course may now be concluded nicely with the first four sections of Chapter 6, together with (the new) Chapter 8. If the rational and Jordan forms are to be included, a more extensive coverage of Chapter 6 is necessary.

Our indebtedness remains to those who contributed to the first edition, especially to Professors Harry Furstenberg, Louis Howard, Daniel Kan, Edward Thorp, to Mrs. Judith Bowers, Mrs. Betty Ann (Sargent) Rose and Miss Phyllis Ruby. In addition, we would like to thank the many students and colleagues whose perceptive comments led to this revision, and the staff of Prentice-Hall for their patience in dealing with two authors caught in the throes of academic administration. Lastly, special thanks are due to Mrs. Sophia Koulouras for both her skill and her tireless efforts in typing the revised manuscript.

K. M. H. / R. A. K.

Contents

3. Linear Transformations

3.1. Linear Transformations

We shall now introduce linear transformations, the objects which we shall study in most of the remainder of this book. The reader may find it helpful to read (or reread) the discussion of functions in the Appendix, since we shall freely use the terminology of that discussion.

Definition. *Let* V *and* W *be vector spaces over the field* F. *A* **linear transformation from** V **into** W *is a function* T *from* V *into* W *such that*

$$T(c\alpha + \beta) = c(T\alpha) + T\beta$$

for all α *and* β *in* V *and all scalars* c *in* F.

EXAMPLE 1. If V is any vector space, the identity transformation I, defined by $I\alpha = \alpha$, is a linear transformation from V into V. The **zero transformation** 0, defined by $0\alpha = 0$, is a linear transformation from V into V.

EXAMPLE 2. Let F be a field and let V be the space of polynomial functions f from F into F, given by

$$f(x) = c_0 + c_1x + \cdots + c_kx^k.$$

Let

$$(Df)(x) = c_1 + 2c_2x + \cdots + kc_kx^{k-1}.$$

Then D is a linear transformation from V into V—the differentiation transformation.

EXAMPLE 3. Let A be a fixed $m \times n$ matrix with entries in the field F. The function T defined by $T(X) = AX$ is a linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$. The function U defined by $U(\alpha) = \alpha A$ is a linear transformation from F^m into F^n.

EXAMPLE 4. Let P be a fixed $m \times m$ matrix with entries in the field F and let Q be a fixed $n \times n$ matrix over F. Define a function T from the space $F^{m\times n}$ into itself by $T(A) = PAQ$. Then T is a linear transformation from $F^{m\times n}$ into $F^{m\times n}$, because

$$\begin{aligned} T(cA + B) &= P(cA + B)Q \\ &= (cPA + PB)Q \\ &= cPAQ + PBQ \\ &= cT(A) + T(B). \end{aligned}$$

EXAMPLE 5. Let R be the field of real numbers and let V be the space of all functions from R into R which are *continuous*. Define T by

$$(Tf)(x) = \int_0^x f(t)\, dt.$$

Then T is a linear transformation from V into V. The function Tf is not only continuous but has a continuous first derivative. The linearity of integration is one of its fundamental properties.

The reader should have no difficulty in verifying that the transformations defined in Examples 1, 2, 3, and 5 are linear transformations. We shall expand our list of examples considerably as we learn more about linear transformations.

It is important to note that if T is a linear transformation from V into W, then $T(0) = 0$; one can see this from the definition because

$$T(0) = T(0 + 0) = T(0) + T(0).$$

This point is often confusing to the person who is studying linear algebra for the first time, since he probably has been exposed to a slightly different use of the term 'linear function.' A brief comment should clear up the confusion. Suppose V is the vector space R^1. A linear transformation from V into V is then a particular type of real-valued function on the real line R. In a calculus course, one would probably call such a function linear if its graph is a straight line. A linear transformation from R^1 into R^1, according to our definition, will be a function from R into R, the graph of which is a straight line *passing through the origin.*

In addition to the property $T(0) = 0$, let us point out another property of the general linear transformation T. Such a transformation 'preserves' linear combinations; that is, if $\alpha_1, \ldots, \alpha_n$ are vectors in V and $c_1, \ldots, c_n$ are scalars, then

$$T(c_1\alpha_1 + \cdots + c_n\alpha_n) = c_1(T\alpha_1) + \cdots + c_n(T\alpha_n).$$

This follows readily from the definition. For example,

$$\begin{aligned} T(c_1\alpha_1 + c_2\alpha_2) &= c_1(T\alpha_1) + T(c_2\alpha_2) \\ &= c_1(T\alpha_1) + c_2(T\alpha_2). \end{aligned}$$

Theorem 1. *Let V be a finite-dimensional vector space over the field F and let $\{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V. Let W be a vector space over the same field F and let $\beta_1, \ldots, \beta_n$ be any vectors in W. Then there is precisely one linear transformation T from V into W such that*

$$T\alpha_j = \beta_j, \qquad j = 1, \ldots, n.$$

Proof. To prove there is some linear transformation T with $T\alpha_j = \beta_j$ we proceed as follows. Given α in V, there is a unique n-tuple $(x_1, \ldots, x_n)$ such that

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n.$$

For this vector α we define

$$T\alpha = x_1\beta_1 + \cdots + x_n\beta_n.$$

Then T is a well-defined rule for associating with each vector α in V a vector $T\alpha$ in W. From the definition it is clear that $T\alpha_j = \beta_j$ for each j. To see that T is linear, let

$$\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$$

be in V and let c be any scalar. Now

$$c\alpha + \beta = (cx_1 + y_1)\alpha_1 + \cdots + (cx_n + y_n)\alpha_n$$

and so by definition

$$T(c\alpha + \beta) = (cx_1 + y_1)\beta_1 + \cdots + (cx_n + y_n)\beta_n.$$

On the other hand,

$$\begin{aligned} c(T\alpha) + T\beta &= c \sum_{i=1}^{n} x_i\beta_i + \sum_{i=1}^{n} y_i\beta_i \\ &= \sum_{i=1}^{n} (cx_i + y_i)\beta_i \end{aligned}$$

and thus

$$T(c\alpha + \beta) = c(T\alpha) + T\beta.$$

If U is a linear transformation from V into W with $U\alpha_j = \beta_j$, $j = 1, \ldots, n$, then for the vector $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ we have

$$\begin{aligned} U\alpha &= U\left(\sum_{i=1}^{n} x_i\alpha_i\right) \\ &= \sum_{i=1}^{n} x_i(U\alpha_i) \\ &= \sum_{i=1}^{n} x_i\beta_i \end{aligned}$$

so that U is exactly the rule T which we defined above. This shows that the linear transformation T with $T\alpha_j = \beta_j$ is unique. ∎

Theorem 1 is quite elementary; however, it is so basic that we have stated it formally. The concept of function is very general. If V and W are (non-zero) vector spaces, there is a multitude of functions from V into W. Theorem 1 helps to underscore the fact that the functions which are linear are extremely special.

Example 6. The vectors

$$\alpha_1 = (1, 2)$$
$$\alpha_2 = (3, 4)$$

are linearly independent and therefore form a basis for R^2. According to Theorem 1, there is a unique linear transformation from R^2 into R^3 such that

$$T\alpha_1 = (3, 2, 1)$$
$$T\alpha_2 = (6, 5, 4).$$

If so, we must be able to find $T(\epsilon_1)$. We find scalars c_1, c_2 such that $\epsilon_1 = c_1\alpha_1 + c_2\alpha_2$ and then we know that $T\epsilon_1 = c_1T\alpha_1 + c_2T\alpha_2$. If $(1, 0) = c_1(1, 2) + c_2(3, 4)$ then $c_1 = -2$ and $c_2 = 1$. Thus

$$\begin{aligned} T(1, 0) &= -2(3, 2, 1) + (6, 5, 4) \\ &= (0, 1, 2). \end{aligned}$$

Example 7. Let T be a linear transformation from the m-tuple space F^m into the n-tuple space F^n. Theorem 1 tells us that T is uniquely determined by the sequence of vectors $\beta_1, \ldots, \beta_m$ where

$$\beta_i = T\epsilon_i, \qquad i = 1, \ldots, m.$$

In short, T is uniquely determined by the images of the standard basis vectors. The determination is

$$\alpha = (x_1, \ldots, x_m)$$
$$T\alpha = x_1\beta_1 + \cdots + x_m\beta_m.$$

If B is the $m \times n$ matrix which has row vectors $\beta_1, \ldots, \beta_m$, this says that

$$T\alpha = \alpha B.$$

In other words, if $\beta_i = (B_{i1}, \ldots, B_{in})$, then

$$T(x_1, \ldots, x_n) = [x_1 \cdots x_m] \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mn} \end{bmatrix}.$$

This is a very explicit description of the linear transformation. In Section 3.4 we shall make a serious study of the relationship between linear transformations and matrices. We shall not pursue the particular description $T\alpha = \alpha B$ because it has the matrix B on the right of the vector α, and that can lead to some confusion. The point of this example is to show that we can give an explicit and reasonably simple description of all linear transformations from F^m into F^n.

If T is a linear transformation from V into W, then the range of T is not only a subset of W; it is a subspace of W. Let R_T be the range of T, that is, the set of all vectors β in W such that $\beta = T\alpha$ for some α in V. Let β_1 and β_2 be in R_T and let c be a scalar. There are vectors α_1 and α_2 in V such that $T\alpha_1 = \beta_1$ and $T\alpha_2 = \beta_2$. Since T is linear

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c\beta_1 + \beta_2, \end{aligned}$$

which shows that $c\beta_1 + \beta_2$ is also in R_T.

Another interesting subspace associated with the linear transformation T is the set N consisting of the vectors α in V such that $T\alpha = 0$. It is a subspace of V because

(a) $T(0) = 0$, so that N is non-empty;
(b) if $T\alpha_1 = T\alpha_2 = 0$, then

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c0 + 0 \\ &= 0 \end{aligned}$$

so that $c\alpha_1 + \alpha_2$ is in N.

Definition. *Let* V *and* W *be vector spaces over the field* F *and let* T *be a linear transformation from* V *into* W. *The* **null space** *of* T *is the set of all vectors* α *in* V *such that* Tα = 0.

If V *is finite-dimensional, the* **rank** *of* T *is the dimension of the range of* T *and the* **nullity** *of* T *is the dimension of the null space of* T.

The following is one of the most important results in linear algebra.

Theorem 2. *Let* V *and* W *be vector spaces over the field* F *and let* T *be a linear transformation from* V *into* W. *Suppose that* V *is finite-dimensional. Then*

$$\text{rank}\ (T) + \text{nullity}\ (T) = \dim V.$$

Proof. Let $\{\alpha_1, \ldots, \alpha_k\}$ be a basis for N, the null space of T. There are vectors $\alpha_{k+1}, \ldots, \alpha_n$ in V such that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. We shall now prove that $\{T\alpha_{k+1}, \ldots, T\alpha_n\}$ is a basis for the range of T. The vectors $T\alpha_1, \ldots, T\alpha_n$ certainly span the range of T, and since $T\alpha_j = 0$, for $j \leq k$, we see that $T\alpha_{k+1}, \ldots, T\alpha_n$ span the range. To see that these vectors are independent, suppose we have scalars c_i such that

$$\sum_{i=k+1}^{n} c_i(T\alpha_i) = 0.$$

This says that

$$T\left(\sum_{i=k+1}^{n} c_i\alpha_i\right) = 0$$

and accordingly the vector $\alpha = \sum_{i=k+1}^{n} c_i\alpha_i$ is in the null space of T. Since $\alpha_1, \ldots, \alpha_k$ form a basis for N, there must be scalars $b_1, \ldots, b_k$ such that

$$\alpha = \sum_{i=1}^{k} b_i\alpha_i.$$

Thus

$$\sum_{i=1}^{k} b_i\alpha_i - \sum_{j=k+1}^{n} c_j\alpha_j = 0$$

and since $\alpha_1, \ldots, \alpha_n$ are linearly independent we must have

$$b_1 = \cdots = b_k = c_{k+1} = \cdots = c_n = 0.$$

If r is the rank of T, the fact that $T\alpha_{k+1}, \ldots, T\alpha_n$ form a basis for the range of T tells us that $r = n - k$. Since k is the nullity of T and n is the dimension of V, we are done. ∎

Theorem 3. *If* A *is an* m × n *matrix with entries in the field* F, *then*

$$\textit{row rank}\ (A) = \textit{column rank}\ (A).$$

Proof. Let T be the linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$ defined by $T(X) = AX$. The null space of T is the solution space for the system $AX = 0$, i.e., the set of all column matrices X such that $AX = 0$. The range of T is the set of all $m \times 1$ column matrices Y such that $AX = Y$ has a solution for X. If $A_1, \ldots, A_n$ are the columns of A, then

$$AX = x_1A_1 + \cdots + x_nA_n$$

so that the range of T is the subspace spanned by the columns of A. In other words, the range of T is the column space of A. Therefore,

$$\text{rank}\ (T) = \text{column rank}\ (A).$$

Theorem 2 tells us that if S is the solution space for the system $AX = 0$, then

$$\dim S + \text{column rank}\ (A) = n.$$

We now refer to Example 15 of Chapter 2. Our deliberations there showed that, if r is the dimension of the row space of A, then the solution space S has a basis consisting of $n - r$ vectors:

$$\dim S = n - \text{row rank}\ (A).$$

It is now apparent that

$$\text{row rank}\ (A) = \text{column rank}\ (A). \quad ∎$$

The proof of Theorem 3 which we have just given depends upon explicit calculations concerning systems of linear equations. There is a more conceptual proof which does not rely on such calculations. We shall give such a proof in Section 3.7.

Exercises

1. Which of the following functions T from R^2 into R^2 are linear transformations?

(a) $T(x_1, x_2) = (1 + x_1, x_2)$;
(b) $T(x_1, x_2) = (x_2, x_1)$;
(c) $T(x_1, x_2) = (x_1^2, x_2)$;
(d) $T(x_1, x_2) = (\sin x_1, x_2)$;
(e) $T(x_1, x_2) = (x_1 - x_2, 0)$.

2. Find the range, rank, null space, and nullity for the zero transformation and the identity transformation on a finite-dimensional space V.

3. Describe the range and the null space for the differentiation transformation of Example 2. Do the same for the integration transformation of Example 5.

4. Is there a linear transformation T from R^3 into R^2 such that $T(1, -1, 1) = (1, 0)$ and $T(1, 1, 1) = (0, 1)$?

5. If

$$\begin{aligned} \alpha_1 &= (1, -1), & \beta_1 &= (1, 0) \\ \alpha_2 &= (2, -1), & \beta_2 &= (0, 1) \\ \alpha_3 &= (-3, 2), & \beta_3 &= (1, 1) \end{aligned}$$

is there a linear transformation T from R^2 into R^2 such that $T\alpha_i = \beta_i$ for $i = 1, 2$ and 3?

6. Describe explicitly (as in Exercises 1 and 2) the linear transformation T from F^2 into F^2 such that $T\epsilon_1 = (a, b)$, $T\epsilon_2 = (c, d)$.

7. Let F be a subfield of the complex numbers and let T be the function from F^3 into F^3 defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

(a) Verify that T is a linear transformation.

(b) If (a, b, c) is a vector in F^3, what are the conditions on a, b, and c that the vector be in the range of T? What is the rank of T?

(c) What are the conditions on a, b, and c that (a, b, c) be in the null space of T? What is the nullity of T?

8. Describe explicitly a linear transformation from R^3 into R^3 which has as its range the subspace spanned by $(1, 0, -1)$ and $(1, 2, 2)$.

9. Let V be the vector space of all $n \times n$ matrices over the field F, and let B be a fixed $n \times n$ matrix. If

$$T(A) = AB - BA$$

verify that T is a linear transformation from V into V.

10. Let V be the set of all complex numbers regarded as a vector space over the field of *real* numbers (usual operations). Find a function from V into V which is a linear transformation on the above vector space, but which is not a linear transformation on C^1, i.e., which is not complex linear.

11. Let V be the space of $n \times 1$ matrices over F and let W be the space of $m \times 1$ matrices over F. Let A be a fixed $m \times n$ matrix over F and let T be the linear transformation from V into W defined by $T(X) = AX$. Prove that T is the zero transformation if and only if A is the zero matrix.

12. Let V be an n-dimensional vector space over the field F and let T be a linear transformation from V into V such that the range and null space of T are identical. Prove that n is even. (Can you give an example of such a linear transformation T?)

13. Let V be a vector space and T a linear transformation from V into V. Prove that the following two statements about T are equivalent.

(a) The intersection of the range of T and the null space of T is the zero subspace of V.

(b) If $T(T\alpha) = 0$, then $T\alpha = 0$.

3.2. The Algebra of Linear Transformations

In the study of linear transformations from V into W, it is of fundamental importance that the set of these transformations inherits a natural vector space structure. The set of linear transformations from a space V into itself has even more algebraic structure, because ordinary composition of functions provides a 'multiplication' of such transformations. We shall explore these ideas in this section.

Theorem 4. *Let* V *and* W *be vector spaces over the field* F. *Let* T *and* U *be linear transformations from* V *into* W. *The function* (T + U) *defined by*

$$(T + U)(\alpha) = T\alpha + U\alpha$$

is a linear transformation from V *into* W. *If* c *is any element of* F, *the function* (cT) *defined by*

$$(cT)(\alpha) = c(T\alpha)$$

is a linear transformation from V *into* W. *The set of all linear transformations from* V *into* W, *together with the addition and scalar multiplication defined above, is a vector space over the field* F.

Proof. Suppose T and U are linear transformations from V into W and that we define $(T + U)$ as above. Then

$$\begin{aligned}(T + U)(c\alpha + \beta) &= T(c\alpha + \beta) + U(c\alpha + \beta)\\ &= c(T\alpha) + T\beta + c(U\alpha) + U\beta\\ &= c(T\alpha + U\alpha) + (T\beta + U\beta)\\ &= c(T + U)(\alpha) + (T + U)(\beta)\end{aligned}$$

which shows that $(T + U)$ is a linear transformation. Similarly,

$$\begin{aligned}(cT)(d\alpha + \beta) &= c[T(d\alpha + \beta)]\\ &= c[d(T\alpha) + T\beta]\\ &= cd(T\alpha) + c(T\beta)\\ &= d[c(T\alpha)] + c(T\beta)\\ &= d[(cT)\alpha] + (cT)\beta\end{aligned}$$

which shows that (cT) is a linear transformation.

To verify that the set of linear transformations of V into W (together with these operations) is a vector space, one must directly check each of the conditions on the vector addition and scalar multiplication. We leave the bulk of this to the reader, and content ourselves with this comment: The zero vector in this space will be the zero transformation, which sends every vector of V into the zero vector in W; each of the properties of the two operations follows from the corresponding property of the operations in the space W. ▌

We should perhaps mention another way of looking at this theorem. If one defines sum and scalar multiple as we did above, then the set of *all* functions from V into W becomes a vector space over the field F. This has nothing to do with the fact that V is a vector space, only that V is a non-empty set. When V is a vector space we can define a linear transformation from V into W, and Theorem 4 says that the linear transformations are a subspace of the space of all functions from V into W.

We shall denote the space of linear transformations from V into W by $L(V, W)$. We remind the reader that $L(V, W)$ is defined only when V and W are vector spaces over the same field.

Theorem 5. *Let* V *be an* n*-dimensional vector space over the field* F, *and let* W *be an* m*-dimensional vector space over* F. *Then the space* L(V, W) *is finite-dimensional and has dimension* mn.

Proof. Let

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \text{and} \quad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$$

be ordered bases for V and W, respectively. For each pair of integers (p, q) with $1 \leq p \leq m$ and $1 \leq q \leq n$, we define a linear transformation $E^{p,q}$ from V into W by

$$\begin{aligned}E^{p,q}(\alpha_i) &= \begin{cases} 0, & \text{if } i \neq q \\ \beta_p, & \text{if } i = q \end{cases}\\ &= \delta_{iq}\beta_p.\end{aligned}$$

According to Theorem 1, there is a unique linear transformation from V into W satisfying these conditions. The claim is that the mn transformations $E^{p,q}$ form a basis for $L(V, W)$.

Let T be a linear transformation from V into W. For each j, $1 \leq j \leq n$, let $A_{ij}, \ldots, A_{mj}$ be the coordinates of the vector $T\alpha_j$ in the ordered basis $\mathcal{B}'$, i.e.,

$$T\alpha_j = \sum_{p=1}^{m} A_{pj}\beta_p. \tag{3-1}$$

We wish to show that

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq}E^{p,q}. \tag{3-2}$$

Let U be the linear transformation in the right-hand member of (3-2). Then for each j

$$\begin{aligned} U\alpha_j &= \sum_p \sum_q A_{pq}E^{p,q}(\alpha_j) \\ &= \sum_p \sum_q A_{pq}\delta_{jq}\beta_p \\ &= \sum_{p=1}^{m} A_{pj}\beta_p \\ &= T\alpha_j \end{aligned}$$

and consequently $U = T$. Now (3-2) shows that the $E^{p,q}$ span $L(V, W)$; we must prove that they are independent. But this is clear from what we did above; for, if the transformation

$$U = \sum_p \sum_q A_{pq}E^{p,q}$$

is the zero transformation, then $U\alpha_j = 0$ for each j, so

$$\sum_{p=1}^{m} A_{pj}\beta_p = 0$$

and the independence of the β_p implies that $A_{pj} = 0$ for every p and j. ∎

Theorem 6. *Let* V, W, *and* Z *be vector spaces over the field* F. *Let* T *be a linear transformation from* V *into* W *and* U *a linear transformation from* W *into* Z. *Then the composed function* UT *defined by* $(UT)(\alpha) = U(T(\alpha))$ *is a linear transformation from* V *into* Z.

Proof.

$$\begin{aligned} (UT)(c\alpha + \beta) &= U[T(c\alpha + \beta)] \\ &= U(cT\alpha + T\beta) \\ &= c[U(T\alpha)] + U(T\beta) \\ &= c(UT)(\alpha) + (UT)(\beta). \end{aligned}$$ ∎

In what follows, we shall be primarily concerned with linear transformation of a vector space into itself. Since we would so often have to write 'T is a linear transformation from V into V,' we shall replace this with 'T is a linear operator on V.'

Definition. *If* V *is a vector space over the field* F, *a* **linear operator on** V *is a linear transformation from* V *into* V.

In the case of Theorem 6 when $V = W = Z$, so that U and T are linear operators on the space V, we see that the composition UT is again a linear operator on V. Thus the space $L(V, V)$ has a 'multiplication' defined on it by composition. In this case the operator TU is also defined, and one should note that in general $UT \neq TU$, i.e., $UT - TU \neq 0$. We should take special note of the fact that if T is a linear operator on V then we can compose T with T. We shall use the notation $T^2 = TT$, and in general $T^n = T \cdots T$ (n times) for $n = 1, 2, 3, \ldots$. We define $T^0 = I$ if $T \neq 0$.

Lemma. *Let* V *be a vector space over the field* F; *let* U, T_1 *and* T_2 *be linear operators on* V; *let* c *be an element of* F.

(a) $IU = UI = U$;
(b) $U(T_1 + T_2) = UT_1 + UT_2$; $(T_1 + T_2)U = T_1U + T_2U$;
(c) $c(UT_1) = (cU)T_1 = U(cT_1)$.

Proof. (a) This property of the identity function is obvious. We have stated it here merely for emphasis.

(b)
$$\begin{aligned}[U(T_1 + T_2)](\alpha) &= U[(T_1 + T_2)(\alpha)] \\ &= U(T_1\alpha + T_2\alpha) \\ &= U(T_1\alpha) + U(T_2\alpha) \\ &= (UT_1)(\alpha) + (UT_2)(\alpha)\end{aligned}$$

so that $U(T_1 + T_2) = UT_1 + UT_2$. Also

$$\begin{aligned}[(T_1 + T_2)U](\alpha) &= (T_1 + T_2)(U\alpha) \\ &= T_1(U\alpha) + T_2(U\alpha) \\ &= (T_1U)(\alpha) + (T_2U)(\alpha)\end{aligned}$$

so that $(T_1 + T_2)U = T_1U + T_2U$. (The reader may note that the proofs of these two distributive laws do not use the fact that T_1 and T_2 are linear, and the proof of the second one does not use the fact that U is linear either.)

(c) We leave the proof of part (c) to the reader. ∎

The contents of this lemma and a portion of Theorem 5 tell us that the vector space $L(V, V)$, together with the composition operation, is what is known as a linear algebra with identity. We shall discuss this in Chapter 4.

Example 8. If A is an $m \times n$ matrix with entries in F, we have the linear transformation T defined by $T(X) = AX$, from $F^{n\times 1}$ into $F^{m\times 1}$. If B is a $p \times m$ matrix, we have the linear transformation U from $F^{m\times 1}$ into $F^{p\times 1}$ defined by $U(Y) = BY$. The composition UT is easily described:

$$\begin{aligned}(UT)(X) &= U(T(X)) \\ &= U(AX) \\ &= B(AX) \\ &= (BA)X.\end{aligned}$$

Thus UT is 'left multiplication by the product matrix BA.'

Example 9. Let F be a field and V the vector space of all polynomial functions from F into F. Let D be the differentiation operator defined in Example 2, and let T be the linear operator 'multiplication by x':

$$(Tf)(x) = xf(x).$$

Then $DT \neq TD$. In fact, the reader should find it easy to verify that $DT - TD = I$, the identity operator.

Even though the 'multiplication' we have on $L(V, V)$ is not commutative, it is nicely related to the vector space operations of $L(V, V)$.

Example 10. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for a vector space V. Consider the linear operators $E^{p,q}$ which arose in the proof of Theorem 5:

$$E^{p,q}(\alpha_i) = \delta_{iq}\alpha_p.$$

These n^2 linear operators form a basis for the space of linear operators on V. What is $E^{p,q}E^{r,s}$? We have

$$\begin{aligned}(E^{p,q}E^{r,s})(\alpha_i) &= E^{p,q}(\delta_{is}\alpha_r)\\ &= \delta_{is}E^{p,q}(\alpha_r)\\ &= \delta_{is}\delta_{rq}\alpha_p.\end{aligned}$$

Therefore,

$$E^{p,q}E^{r,s} = \begin{cases}0, & \text{if} \quad r \neq q\\ E^{p,s}, & \text{if} \quad q = r.\end{cases}$$

Let T be a linear operator on V. We showed in the proof of Theorem 5 that if

$$\begin{aligned}A_j &= [T\alpha_j]_{\mathcal{B}}\\ A &= [A_1, \ldots, A_n]\end{aligned}$$

then

$$T = \sum_p \sum_q A_{pq}E^{p,q}.$$

If

$$U = \sum_r \sum_s B_{rs}E^{r,s}$$

is another linear operator on V, then the last lemma tells us that

$$\begin{aligned}TU &= (\sum_p \sum_q A_{pq}E^{p,q})(\sum_r \sum_s B_{rs}E^{r,s})\\ &= \sum_p \sum_q \sum_r \sum_s A_{pq}B_{rs}E^{p,q}E^{r,s}.\end{aligned}$$

As we have noted, the only terms which survive in this huge sum are the terms where $q = r$, and since $E^{p,r}E^{r,s} = E^{p,s}$, we have

$$\begin{aligned}TU &= \sum_p \sum_s (\sum_r A_{pr}B_{rs})E^{p,s}\\ &= \sum_p \sum_s (AB)_{ps}E^{p,s}.\end{aligned}$$

Thus, the effect of composing T and U is to multiply the matrices A and B.

In our discussion of algebraic operations with linear transformations we have not yet said anything about invertibility. One specific question of interest is this. For which linear operators T on the space V does there exist a linear operator T^{-1} such that $TT^{-1} = T^{-1}T = I$?

The function T from V into W is called **invertible** if there exists a function U from W into V such that UT is the identity function on V and TU is the identity function on W. If T is invertible, the function U is unique and is denoted by T^{-1}. (See Appendix.) Furthermore, T is invertible if and only if

1. T is 1:1, that is, $T\alpha = T\beta$ implies $\alpha = \beta$;
2. T is onto, that is, the range of T is (all of) W.

Theorem 7. *Let* V *and* W *be vector spaces over the field* F *and let* T *be a linear transformation from* V *into* W. *If* T *is invertible, then the inverse function* T^{-1} *is a linear transformation from* W *onto* V.

Proof. We repeat ourselves in order to underscore a point. When T is one-one and onto, there is a uniquely determined inverse function T^{-1} which maps W onto V such that $T^{-1}T$ is the identity function on V, and TT^{-1} is the identity function on W. What we are proving here is that if a linear function T is invertible, then the inverse T^{-1} is also linear.

Let β_1 and β_2 be vectors in W and let c be a scalar. We wish to show that

$$T^{-1}(c\beta_1 + \beta_2) = cT^{-1}\beta_1 + T^{-1}\beta_2.$$

Let $\alpha_i = T^{-1}\beta_i$, $i = 1, 2$, that is, let α_i be the unique vector in V such that $T\alpha_i = \beta_i$. Since T is linear,

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c\beta_1 + \beta_2. \end{aligned}$$

Thus $c\alpha_1 + \alpha_2$ is the unique vector in V which is sent by T into $c\beta_1 + \beta_2$, and so

$$\begin{aligned} T^{-1}(c\beta_1 + \beta_2) &= c\alpha_1 + \alpha_2 \\ &= c(T^{-1}\beta_1) + T^{-1}\beta_2 \end{aligned}$$

and T^{-1} is linear. ∎

Suppose that we have an invertible linear transformation T from V onto W and an invertible linear transformation U from W onto Z. Then UT is invertible and $(UT)^{-1} = T^{-1}U^{-1}$. That conclusion does not require the linearity nor does it involve checking separately that UT is 1:1 and onto. All it involves is verifying that $T^{-1}U^{-1}$ is both a left and a right inverse for UT.

If T is linear, then $T(\alpha - \beta) = T\alpha - T\beta$; hence, $T\alpha = T\beta$ if and only if $T(\alpha - \beta) = 0$. This simplifies enormously the verification that T is 1:1. Let us call a linear transformation T **non-singular** if $T\gamma = 0$ implies

$\gamma = 0$, i.e., if the null space of T is $\{0\}$. Evidently, T is 1:1 if and only if T is non-singular. The extension of this remark is that non-singular linear transformations are those which preserve linear independence.

Theorem 8. *Let* T *be a linear transformation from* V *into* W. *Then* T *is non-singular if and only if* T *carries each linearly independent subset of* V *onto a linearly independent subset of* W.

Proof. First suppose that T is non-singular. Let S be a linearly independent subset of V. If $\alpha_1, \ldots, \alpha_k$ are vectors in S, then the vectors $T\alpha_1, \ldots, T\alpha_k$ are linearly independent; for if

$$c_1(T\alpha_1) + \cdots + c_k(T\alpha_k) = 0$$

then

$$T(c_1\alpha_1 + \cdots + c_k\alpha_k) = 0$$

and since T is non-singular

$$c_1\alpha_1 + \cdots + c_k\alpha_k = 0$$

from which it follows that each $c_i = 0$ because S is an independent set. This argument shows that the image of S under T is independent.

Suppose that T carries independent subsets onto independent subsets. Let α be a non-zero vector in V. Then the set S consisting of the one vector α is independent. The image of S is the set consisting of the one vector $T\alpha$, and this set is independent. Therefore $T\alpha \neq 0$, because the set consisting of the zero vector alone is dependent. This shows that the null space of T is the zero subspace, i.e., T is non-singular. ∎

EXAMPLE 11. Let F be a subfield of the complex numbers (or a field of characteristic zero) and let V be the space of polynomial functions over F. Consider the differentiation operator D and the 'multiplication by x' operator T, from Example 9. Since D sends all constants into 0, D is singular; however, V is not finite dimensional, the range of D is all of V, and it is possible to define a right inverse for D. For example, if E is the indefinite integral operator:

$$E(c_0 + c_1x + \cdots + c_nx^n) = c_0x + \frac{1}{2}c_1x^2 + \cdots + \frac{1}{n+1}c_nx^{n+1}$$

then E is a linear operator on V and $DE = I$. On the other hand, $ED \neq I$ because ED sends the constants into 0. The operator T is in what we might call the reverse situation. If $xf(x) = 0$ for all x, then $f = 0$. Thus T is non-singular and it is possible to find a left inverse for T. For example if U is the operation 'remove the constant term and divide by x':

$$U(c_0 + c_1x + \cdots + c_nx^n) = c_1 + c_2x + \cdots + c_nx^{n-1}$$

then U is a linear operator on V and $UT = I$. But $TU \neq I$ since every function in the range of TU is in the range of T, which is the space of polynomial functions f such that $f(0) = 0$.

EXAMPLE 12. Let F be a field and let T be the linear operator on F^2 defined by

$$T(x_1, x_2) = (x_1 + x_2, x_1).$$

Then T is non-singular, because if $T(x_1, x_2) = 0$ we have

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 &= 0 \end{aligned}$$

so that $x_1 = x_2 = 0$. We also see that T is onto; for, let (z_1, z_2) be any vector in F^2. To show that (z_1, z_2) is in the range of T we must find scalars x_1 and x_2 such that

$$\begin{aligned} x_1 + x_2 &= z_1 \\ x_1 &= z_2 \end{aligned}$$

and the obvious solution is $x_1 = z_2$, $x_2 = z_1 - z_2$. This last computation gives us an explicit formula for T^{-1}, namely,

$$T^{-1}(z_1, z_2) = (z_2, z_1 - z_2).$$

We have seen in Example 11 that a linear transformation may be non-singular without being onto and may be onto without being non-singular. The present example illustrates an important case in which that cannot happen.

Theorem 9. *Let* V *and* W *be finite-dimensional vector spaces over the field* F *such that dim* V = *dim* W. *If* T *is a linear transformation from* V *into* W, *the following are equivalent:*

(i) T *is invertible.*
(ii) T *is non-singular.*
(iii) T *is onto, that is, the range of* T *is* W.

Proof. Let $n = \dim V = \dim W$. From Theorem 2 we know that

$$\text{rank } (T) + \text{nullity } (T) = n.$$

Now T is non-singular if and only if nullity $(T) = 0$, and (since $n = \dim W$) the range of T is W if and only if rank $(T) = n$. Since the rank plus the nullity is n, the nullity is 0 precisely when the rank is n. Therefore T is non-singular if and only if $T(V) = W$. So, if either condition (ii) or (iii) holds, the other is satisfied as well and T is invertible. ∎

We caution the reader not to apply Theorem 9 except in the presence of finite-dimensionality and with $\dim V = \dim W$. Under the hypotheses of Theorem 9, the conditions (i), (ii), and (iii) are also equivalent to these.

(iv) *If* $\{\alpha_1, \ldots, \alpha_n\}$ *is basis for* V, *then* $\{T\alpha_1, \ldots, T\alpha_n\}$ *is a basis for* W.

(v) *There is some basis* $\{\alpha_1, \ldots, \alpha_n\}$ *for* V *such that* $\{T\alpha_1, \ldots, T\alpha_n\}$ *is a basis for* W.

We shall give a proof of the equivalence of the five conditions which contains a different proof that (i), (ii), and (iii) are equivalent.

(i) $\to$ (ii). If T is invertible, T is non-singular. (ii) $\to$ (iii). Suppose T is non-singular. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. By Theorem 8, $\{T\alpha_1, \ldots, T\alpha_n\}$ is a linearly independent set of vectors in W, and since the dimension of W is also n, this set of vectors is a basis for W. Now let β be any vector in W. There are scalars $c_1, \ldots, c_n$ such that

$$\begin{aligned}\beta &= c_1(T\alpha_1) + \cdots + c_n(T\alpha_n)\\ &= T(c_1\alpha_1 + \cdots + c_n\alpha_n)\end{aligned}$$

which shows that β is in the range of T. (iii) $\to$ (iv). We now assume that T is onto. If $\{\alpha_1, \ldots, \alpha_n\}$ is any basis for V, the vectors $T\alpha_1, \ldots, T\alpha_n$ span the range of T, which is all of W by assumption. Since the dimension of W is n, these n vectors must be linearly independent, that is, must comprise a basis for W. (iv) $\to$ (v). This requires no comment. (v) $\to$ (i). Suppose there is some basis $\{\alpha_1, \ldots, \alpha_n\}$ for V such that $\{T\alpha_1, \ldots, T\alpha_n\}$ is a basis for W. Since the $T\alpha_i$ span W, it is clear that the range of T is all of W. If $\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n$ is in the null space of T, then

$$T(c_1\alpha_1 + \cdots + c_n\alpha_n) = 0$$

or

$$c_1(T\alpha_1) + \cdots + c_n(T\alpha_n) = 0$$

and since the $T\alpha_i$ are independent each $c_i = 0$, and thus $\alpha = 0$. We have shown that the range of T is W, and that T is non-singular, hence T is invertible.

The set of invertible linear operators on a space V, with the operation of composition, provides a nice example of what is known in algebra as a 'group.' Although we shall not have time to discuss groups in any detail, we shall at least give the definition.

Definition. *A* **group** *consists of the following.*

1. *A set* G;

2. *A rule (or operation) which associates with each pair of elements* x, y *in* G *an element* xy *in* G *in such a way that*

(a) x(yz) = (xy)z, *for all* x, y, *and* z *in* G *(associativity);*

(b) *there is an element* e *in* G *such that* ex = xe = x, *for every* x *in* G;

(c) *to each element* x *in* G *there corresponds an element* x^{-1} *in* G *such that* $xx^{-1} = x^{-1}x = e$.

We have seen that composition $(U, T) \to UT$ associates with each pair of invertible linear operators on a space V another invertible operator on V. Composition is an associative operation. The identity operator I

satisfies $IT = TI$ for each T, and for an invertible T there is (by Theorem 7) an invertible linear operator T^{-1} such that $TT^{-1} = T^{-1}T = I$. Thus the set of invertible linear operators on V, together with this operation, is a group. The set of invertible $n \times n$ matrices with matrix multiplication as the operation is another example of a group. A group is called **commutative** if it satisfies the condition $xy = yx$ for each x and y. The two examples we gave above are not commutative groups, in general. One often writes the operation in a commutative group as $(x, y) \rightarrow x + y$, rather than $(x, y) \rightarrow xy$, and then uses the symbol 0 for the 'identity' element e. The set of vectors in a vector space, together with the operation of vector addition, is a commutative group. A field can be described as a set with two operations, called addition and multiplication, which is a commutative group under addition, and in which the non-zero elements form a commutative group under multiplication, with the distributive law $x(y + z) = xy + xz$ holding.

Exercises

1. Let T and U be the linear operators on R^2 defined by

$$T(x_1, x_2) = (x_2, x_1) \quad \text{and} \quad U(x_1, x_2) = (x_1, 0).$$

(a) How would you describe T and U geometrically?

(b) Give rules like the ones defining T and U for each of the transformations $(U + T)$, UT, TU, T^2, U^2.

2. Let T be the (unique) linear operator on C^3 for which

$$T\epsilon_1 = (1, 0, i), \qquad T\epsilon_2 = (0, 1, 1), \qquad T\epsilon_3 = (i, 1, 0).$$

Is T invertible?

3. Let T be the linear operator on R^3 defined by

$$T(x_1, x_2, x_3) = (3x_1, x_1 - x_2, 2x_1 + x_2 + x_3).$$

Is T invertible? If so, find a rule for T^{-1} like the one which defines T.

4. For the linear operator T of Exercise 3, prove that

$$(T^2 - I)(T - 3I) = 0.$$

5. Let $C^{2\times 2}$ be the complex vector space of 2×2 matrices with complex entries. Let

$$B = \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix}$$

and let T be the linear operator on $C^{2\times 2}$ defined by $T(A) = BA$. What is the rank of T? Can you describe T^2?

6. Let T be a linear transformation from R^3 into R^2, and let U be a linear transformation from R^2 into R^3. Prove that the transformation UT is not invertible. Generalize the theorem.

7. Find two linear operators T and U on R^2 such that $TU = 0$ but $UT \neq 0$.

8. Let V be a vector space over the field F and T a linear operator on V. If $T^2 = 0$, what can you say about the relation of the range of T to the null space of T? Give an example of a linear operator T on R^2 such that $T^2 = 0$ but $T \neq 0$.

9. Let T be a linear operator on the finite-dimensional space V. Suppose there is a linear operator U on V such that $TU = I$. Prove that T is invertible and $U = T^{-1}$. Give an example which shows that this is false when V is not finite-dimensional. (*Hint:* Let $T = D$, the differentiation operator on the space of polynomial functions.)

10. Let A be an $m \times n$ matrix with entries in F and let T be the linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$ defined by $T(X) = AX$. Show that if $m < n$ it may happen that T is onto without being non-singular. Similarly, show that if $m > n$ we may have T non-singular but not onto.

11. Let V be a finite-dimensional vector space and let T be a linear operator on V. Suppose that rank (T^2) = rank (T). Prove that the range and null space of T are disjoint, i.e., have only the zero vector in common.

12. Let p, m, and n be positive integers and F a field. Let V be the space of $m \times n$ matrices over F and W the space of $p \times n$ matrices over F. Let B be a fixed $p \times m$ matrix and let T be the linear transformation from V into W defined by $T(A) = BA$. Prove that T is invertible if and only if $p = m$ and B is an invertible $m \times m$ matrix.

3.3. Isomorphism

If V and W are vector spaces over the field F, any one-one linear transformation T of V onto W is called an **isomorphism of V onto W**. If there exists an isomorphism of V onto W, we say that V is **isomorphic** to W.

Note that V is trivially isomorphic to V, the identity operator being an isomorphism of V onto V. Also, if V is isomorphic to W via an isomorphism T, then W is isomorphic to V, because T^{-1} is an isomorphism of W onto V. The reader should find it easy to verify that if V is isomorphic to W and W is isomorphic to Z, then V is isomorphic to Z. Briefly, isomorphism is an equivalence relation on the class of vector spaces. If there exists an isomorphism of V onto W, we may sometimes say that V and W are isomorphic, rather than V is isomorphic to W. This will cause no confusion because V is isomorphic to W if and only if W is isomorphic to V.

Theorem 10. *Every* n*-dimensional vector space over the field* F *is isomorphic to the space* F^n.

Proof. Let V be an n-dimensional space over the field F and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V. We define a function T

from V into F^n, as follows: If α is in V, let $T\alpha$ be the n-tuple $(x_1, \ldots, x_n)$ of coordinates of α relative to the ordered basis $\mathcal{B}$, i.e., the n-tuple such that

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n.$$

In our discussion of coordinates in Chapter 2, we verified that this T is linear, one-one, and maps V onto F^n. ▌

For many purposes one often regards isomorphic vector spaces as being 'the same,' although the vectors and operations in the spaces may be quite different, that is, one often identifies isomorphic spaces. We shall not attempt a lengthy discussion of this idea at present but shall let the understanding of isomorphism and the sense in which isomorphic spaces are 'the same' grow as we continue our study of vector spaces.

We shall make a few brief comments. Suppose T is an isomorphism of V onto W. If S is a subset of V, then Theorem 8 tells us that S is linearly independent if and only if the set $T(S)$ in W is independent. Thus in deciding whether S is independent it doesn't matter whether we look at S or $T(S)$. From this one sees that an isomorphism is 'dimension preserving,' that is, any finite-dimensional subspace of V has the same dimension as its image under T. Here is a very simple illustration of this idea. Suppose A is an $m \times n$ matrix over the field F. We have really given two definitions of the solution space of the matrix A. The first is the set of all n-tuples $(x_1, \ldots, x_n)$ in F^n which satisfy each of the equations in the system $AX = 0$. The second is the set of all $n \times 1$ column matrices X such that $AX = 0$. The first solution space is thus a subspace of F^n and the second is a subspace of the space of all $n \times 1$ matrices over F. Now there is a completely obvious isomorphism between F^n and $F^{n\times 1}$, namely,

$$(x_1, \ldots, x_n) \to \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Under this isomorphism, the first solution space of A is carried onto the second solution space. These spaces have the same dimension, and so if we want to prove a theorem about the dimension of the solution space, it is immaterial which space we choose to discuss. In fact, the reader would probably not balk if we chose to identify F^n and the space of $n \times 1$ matrices. We may do this when it is convenient, and when it is not convenient we shall not.

Exercises

1. Let V be the set of complex numbers and let F be the field of real numbers. With the usual operations, V is a vector space over F. Describe explicitly an isomorphism of this space onto R^2.

2. Let V be a vector space over the field of complex numbers, and suppose there is an isomorphism T of V onto C^3. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be vectors in V such that

$$T\alpha_1 = (1, 0, i), \qquad T\alpha_2 = (-2, 1 + i, 0),$$
$$T\alpha_3 = (-1, 1, 1), \qquad T\alpha_4 = (\sqrt{2}, i, 3).$$

(a) Is α_1 in the subspace spanned by α_2 and α_3?

(b) Let W_1 be the subspace spanned by α_1 and α_2, and let W_2 be the subspace spanned by α_3 and α_4. What is the intersection of W_1 and W_2?

(c) Find a basis for the subspace of V spanned by the four vectors α_j.

3. Let W be the set of all 2×2 complex Hermitian matrices, that is, the set of 2×2 complex matrices A such that $A_{ij} = \overline{A_{ji}}$ (the bar denoting complex conjugation). As we pointed out in Example 6 of Chapter 2, W is a vector space over the field of *real* numbers, under the usual operations. Verify that

$$(x, y, z, t) \rightarrow \begin{bmatrix} t + x & y + iz \\ y - iz & t - x \end{bmatrix}$$

is an isomorphism of R^4 onto W.

4. Show that $F^{m \times n}$ is isomorphic to F^{mn}.

5. Let V be the set of complex numbers regarded as a vector space over the field of real numbers (Exercise 1). We define a function T from V into the space of 2×2 real matrices, as follows. If $z = x + iy$ with x and y real numbers, then

$$T(z) = \begin{bmatrix} x + 7y & 5y \\ -10y & x - 7y \end{bmatrix}.$$

(a) Verify that T is a one-one (real) linear transformation of V into the space of 2×2 real matrices.

(b) Verify that $T(z_1 z_2) = T(z_1)T(z_2)$.

(c) How would you describe the range of T?

6. Let V and W be finite-dimensional vector spaces over the field F. Prove that V and W are isomorphic if and only if $\dim V = \dim W$.

7. Let V and W be vector spaces over the field F and let U be an isomorphism of V onto W. Prove that $T \rightarrow UTU^{-1}$ is an isomorphism of $L(V, V)$ onto $L(W, W)$.

3.4. Representation of Transformations by Matrices

Let V be an n-dimensional vector space over the field F and let W be an m-dimensional vector space over F. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V and $\mathcal{B}' = \{\beta_1, \ldots, \beta_m\}$ an ordered basis for W. If T is any linear transformation from V into W, then T is determined by its action on the vectors α_j. Each of the n vectors $T\alpha_j$ is uniquely expressible as a linear combination

$$T\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i \tag{3-3}$$

of the β_i, the scalars $A_{1j}, \ldots, A_{mj}$ being the coordinates of $T\alpha_j$ in the ordered basis $\mathcal{B}'$. Accordingly, the transformation T is determined by the mn scalars A_{ij} via the formulas (3-3). The $m \times n$ matrix A defined by $A(i, j) = A_{ij}$ is called **the matrix of T relative to the pair of ordered bases $\mathcal{B}$ and $\mathcal{B}'$**. Our immediate task is to understand explicitly how the matrix A determines the linear transformation T.

If $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ is a vector in V, then

$$\begin{aligned} T\alpha &= T\left(\sum_{j=1}^{n} x_j\alpha_j\right) \\ &= \sum_{j=1}^{n} x_j(T\alpha_j) \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} A_{ij}\beta_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right)\beta_i. \end{aligned}$$

If X is the coordinate matrix of α in the ordered basis $\mathcal{B}$, then the computation above shows that AX is the coordinate matrix of the vector $T\alpha$ in the ordered basis $\mathcal{B}'$, because the scalar

$$\sum_{j=1}^{n} A_{ij}x_j$$

is the entry in the ith row of the column matrix AX. Let us also observe that if A is any $m \times n$ matrix over the field F, then

$$T\left(\sum_{j=1}^{n} x_j\alpha_j\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right)\beta_i \tag{3-4}$$

defines a linear transformation T from V into W, the matrix of which is A, relative to $\mathcal{B}$, $\mathcal{B}'$. We summarize formally:

Theorem 11. *Let* V *be an* n*-dimensional vector space over the field* F *and* W *an* m*-dimensional vector space over* F. *Let* $\mathcal{B}$ *be an ordered basis for* V *and* $\mathcal{B}'$ *an ordered basis for* W. *For each linear transformation* T *from* V *into* W, *there is an* m $\times$ n *matrix* A *with entries in* F *such that*

$$[T\alpha]_{\mathcal{B}'} = A[\alpha]_{\mathcal{B}}$$

for every vector α *in* V. *Furthermore,* T $\to$ A *is a one-one correspondence between the set of all linear transformations from* V *into* W *and the set of all* m $\times$ n *matrices over the field* F.

The matrix A which is associated with T in Theorem 11 is called the **matrix of T relative to the ordered bases** $\mathcal{B}$, $\mathcal{B}'$. Note that Equation (3-3) says that A is the matrix whose columns $A_1, \ldots, A_n$ are given by

$$A_j = [T\alpha_j]_{\mathcal{B}'}, \qquad j = 1, \ldots, n.$$

If U is another linear transformation from V into W and $B = [B_1, \ldots, B_n]$ is the matrix of U relative to the ordered bases $\mathcal{B}$, $\mathcal{B}'$ then $cA + B$ is the matrix of $cT + U$ relative to $\mathcal{B}$, $\mathcal{B}'$. That is clear because

$$\begin{aligned} cA_j + B_j &= c[T\alpha_j]_{\mathcal{B}'} + [U\alpha_j]_{\mathcal{B}'} \\ &= [cT\alpha_j + U\alpha_j]_{\mathcal{B}'} \\ &= [(cT + U)\alpha_j]_{\mathcal{B}'}. \end{aligned}$$

Theorem 12. *Let* V *be an* n*-dimensional vector space over the field* F *and let* W *be an* m*-dimensional vector space over* F. *For each pair of ordered bases* $\mathcal{B}$, $\mathcal{B}'$ *for* V *and* W *respectively, the function which assigns to a linear transformation* T *its matrix relative to* $\mathcal{B}$, $\mathcal{B}'$ *is an isomorphism between the space* L(V, W) *and the space of all* m × n *matrices over the field* F.

Proof. We observed above that the function in question is linear, and as stated in Theorem 11, this function is one-one and maps $L(V, W)$ onto the set of $m \times n$ matrices. ∎

We shall be particularly interested in the representation by matrices of linear transformations of a space into itself, i.e., linear operators on a space V. In this case it is most convenient to use the same ordered basis in each case, that is, to take $\mathcal{B} = \mathcal{B}'$. We shall then call the representing matrix simply the **matrix of T relative to the ordered basis** $\mathcal{B}$. Since this concept will be so important to us, we shall review its definition. If T is a linear operator on the finite-dimensional vector space V and $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V, the matrix of T relative to $\mathcal{B}$ (or, the matrix of T in the ordered basis $\mathcal{B}$) is the $n \times n$ matrix A whose entries A_{ij} are defined by the equations

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i, \qquad j = 1, \ldots, n. \tag{3-5}$$

One must always remember that this matrix representing T depends upon the ordered basis $\mathcal{B}$, and that there is a representing matrix for T in each ordered basis for V. (For transformations of one space into another the matrix depends upon two ordered bases, one for V and one for W.) In order that we shall not forget this dependence, we shall use the notation

$$[T]_{\mathcal{B}}$$

for the matrix of the linear operator T in the ordered basis $\mathcal{B}$. The manner in which this matrix and the ordered basis describe T is that for each α in V

$$[T\alpha]_{\mathcal{B}} = [T]_{\mathcal{B}}[\alpha]_{\mathcal{B}}.$$

EXAMPLE 13. Let V be the space of $n \times 1$ column matrices over the field F; let W be the space of $m \times 1$ matrices over F; and let A be a fixed $m \times n$ matrix over F. Let T be the linear transformation of V into W defined by $T(X) = AX$. Let $\mathcal{B}$ be the ordered basis for V analogous to the standard basis in F^n, i.e., the ith vector in $\mathcal{B}$ in the $n \times 1$ matrix X_i with a 1 in row i and all other entries 0. Let $\mathcal{B}'$ be the corresponding ordered basis for W, i.e., the jth vector in $\mathcal{B}'$ is the $m \times 1$ matrix Y_j with a 1 in row j and all other entries 0. Then the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$ is the matrix A itself. This is clear because the matrix AX_j is the jth column of A.

Example 14. Let F be a field and let T be the operator on F^2 defined by

$$T(x_1, x_2) = (x_1, 0).$$

It is easy to see that T is a linear operator on F^2. Let $\mathcal{B}$ be the standard ordered basis for F^2, $\mathcal{B} = \{\epsilon_1, \epsilon_2\}$. Now

$$\begin{aligned} T\epsilon_1 &= T(1, 0) = (1, 0) = 1\epsilon_1 + 0\epsilon_2 \\ T\epsilon_2 &= T(0, 1) = (0, 0) = 0\epsilon_1 + 0\epsilon_2 \end{aligned}$$

so the matrix of T in the ordered basis $\mathcal{B}$ is

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Example 15. Let V be the space of all polynomial functions from R into R of the form

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3$$

that is, the space of polynomial functions of degree three or less. The differentiation operator D of Example 2 maps V into V, since D is 'degree decreasing.' Let $\mathcal{B}$ be the ordered basis for V consisting of the four functions f_1, f_2, f_3, f_4 defined by $f_j(x) = x^{j-1}$. Then

$$\begin{aligned} (Df_1)(x) &= 0, & Df_1 &= 0f_1 + 0f_2 + 0f_3 + 0f_4 \\ (Df_2)(x) &= 1, & Df_2 &= 1f_1 + 0f_2 + 0f_3 + 0f_4 \\ (Df_3)(x) &= 2x, & Df_3 &= 0f_1 + 2f_2 + 0f_3 + 0f_4 \\ (Df_4)(x) &= 3x^2, & Df_4 &= 0f_1 + 0f_2 + 3f_3 + 0f_4 \end{aligned}$$

so that the matrix of D in the ordered basis $\mathcal{B}$ is

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have seen what happens to representing matrices when transformations are added, namely, that the matrices add. We should now like to ask what happens when we compose transformations. More specifically, let V, W, and Z be vector spaces over the field F of respective dimensions n, m, and p. Let T be a linear transformation from V into W and U a linear transformation from W into Z. Suppose we have ordered bases

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}, \qquad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\}, \qquad \mathcal{B}'' = \{\gamma_1, \ldots, \gamma_p\}$$

for the respective spaces V, W, and Z. Let A be the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$ and let B be the matrix of U relative to the pair $\mathcal{B}'$, $\mathcal{B}''$. It is then easy to see that the matrix C of the transformation UT relative to the pair $\mathcal{B}$, $\mathcal{B}''$ is the product of B and A; for, if α is any vector in V

$$\begin{aligned}[][T\alpha]_{\mathcal{B}'} &= A[\alpha]_{\mathcal{B}} \\ [U(T\alpha)]_{\mathcal{B}''} &= B[T\alpha]_{\mathcal{B}'}\end{aligned}$$

and so

$$[(UT)(\alpha)]_{\mathcal{B}''} = BA[\alpha]_{\mathcal{B}}$$

and hence, by the definition and uniqueness of the representing matrix, we must have $C = BA$. One can also see this by carrying out the computation

$$\begin{aligned}(UT)(\alpha_j) &= U(T\alpha_j) \\ &= U\left(\sum_{k=1}^{m} A_{kj}\beta_k\right) \\ &= \sum_{k=1}^{m} A_{kj}(U\beta_k) \\ &= \sum_{k=1}^{m} A_{kj} \sum_{i=1}^{p} B_{ik}\gamma_i \\ &= \sum_{i=1}^{p} \left(\sum_{k=1}^{m} B_{ik}A_{kj}\right)\gamma_i\end{aligned}$$

so that we must have

$$C_{ij} = \sum_{k=1}^{m} B_{ik}A_{kj}. \tag{3-6}$$

We motivated the definition (3-6) of matrix multiplication via operations on the rows of a matrix. One sees here that a very strong motivation for the definition is to be found in composing linear transformations. Let us summarize formally.

Theorem 13. *Let* V, W, *and* Z *be finite-dimensional vector spaces over the field* F; *let* T *be a linear transformation from* V *into* W *and* U *a linear transformation from* W *into* Z. *If* $\mathcal{B}$, $\mathcal{B}'$, *and* $\mathcal{B}''$ *are ordered bases for the spaces* V, W, *and* Z, *respectively, if* A *is the matrix of* T *relative to the pair* $\mathcal{B}$, $\mathcal{B}'$, *and* B *is the matrix of* U *relative to the pair* $\mathcal{B}'$, $\mathcal{B}''$, *then the matrix of the composition* UT *relative to the pair* $\mathcal{B}$, $\mathcal{B}''$ *is the product matrix* C = BA.

We remark that Theorem 13 gives a proof that matrix multiplication is associative—a proof which requires no calculations and is independent of the proof we gave in Chapter 1. We should also point out that we proved a special case of Theorem 13 in Example 12.

It is important to note that if T and U are linear operators on a space V and we are representing by a single ordered basis $\mathcal{B}$, then Theorem 13 assumes the simple form $[UT]_{\mathcal{B}} = [U]_{\mathcal{B}}[T]_{\mathcal{B}}$. Thus in this case, the correspondence which $\mathfrak{B}$ determines between operators and matrices is not only a vector space isomorphism but also preserves products. A simple consequence of this is that the linear operator T is invertible if and only if $[T]_{\mathfrak{B}}$ is an invertible matrix. For, the identity operator I is represented by the identity matrix in any ordered basis, and thus

$$UT = TU = I$$

is equivalent to

$$[U]_{\mathfrak{B}}[T]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[U]_{\mathfrak{B}} = I.$$

Of course, when T is invertible

$$[T^{-1}]_{\mathfrak{B}} = [T]_{\mathfrak{B}}^{-1}.$$

Now we should like to inquire what happens to representing matrices when the ordered basis is changed. For the sake of simplicity, we shall consider this question only for linear operators on a space V, so that we can use a single ordered basis. The specific question is this. Let T be a linear operator on the finite-dimensional space V, and let

$$\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \text{and} \quad \mathfrak{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

be two ordered bases for V. How are the matrices $[T]_{\mathfrak{B}}$ and $[T]_{\mathfrak{B}'}$ related? As we observed in Chapter 2, there is a unique (invertible) $n \times n$ matrix P such that

$$[\alpha]_{\mathfrak{B}} = P[\alpha]_{\mathfrak{B}'} \tag{3-7}$$

for every vector α in V. It is the matrix $P = [P_1, \ldots, P_n]$ where $P_j = [\alpha'_j]_{\mathfrak{B}}$. By definition

$$[T\alpha]_{\mathfrak{B}} = [T]_{\mathfrak{B}}[\alpha]_{\mathfrak{B}}. \tag{3-8}$$

Applying (3-7) to the vector $T\alpha$, we have

$$[T\alpha]_{\mathfrak{B}} = P[T\alpha]_{\mathfrak{B}'}. \tag{3-9}$$

Combining (3-7), (3-8), and (3-9), we obtain

$$[T]_{\mathfrak{B}}P[\alpha]_{\mathfrak{B}'} = P[T\alpha]_{\mathfrak{B}'}$$

or

$$P^{-1}[T]_{\mathfrak{B}}P[\alpha]_{\mathfrak{B}'} = [T\alpha]_{\mathfrak{B}'}$$

and so it must be that

$$[T]_{\mathfrak{B}'} = P^{-1}[T]_{\mathfrak{B}}P. \tag{3-10}$$

This answers our question.

Before stating this result formally, let us observe the following. There is a unique linear operator U which carries $\mathfrak{B}$ onto $\mathfrak{B}'$, defined by

$$U\alpha_j = \alpha'_j, \qquad j = 1, \ldots, n.$$

This operator U is invertible since it carries a basis for V onto a basis for

V. The matrix P (above) is precisely the matrix of the operator U in the ordered basis $\mathcal{B}$. For, P is defined by

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i$$

and since $U\alpha_j = \alpha_j'$, this equation can be written

$$U\alpha_j = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

So $P = [U]_{\mathcal{B}}$, by definition.

Theorem 14. *Let* V *be a finite-dimensional vector space over the field* F, *and let*

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\} \quad \textit{and} \quad \mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$$

be ordered bases for V. *Suppose* T *is a linear operator on* V. *If* $P = [P_1, \ldots, P_n]$ *is the* $n \times n$ *matrix with columns* $P_j = [\alpha_j']_{\mathcal{B}}$, *then*

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P.$$

Alternatively, if U *is the invertible operator on* V *defined by* $U\alpha_j = \alpha_j'$, $j = 1, \ldots, n$, *then*

$$[T]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}}.$$

EXAMPLE 16. Let T be the linear operator on R^2 defined by $T(x_1, x_2) = (x_1, 0)$. In Example 14 we showed that the matrix of T in the standard ordered basis $\mathcal{B} = \{\epsilon_1, \epsilon_2\}$ is

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suppose $\mathcal{B}'$ is the ordered basis for R^2 consisting of the vectors $\epsilon_1' = (1, 1)$, $\epsilon_2' = (2, 1)$. Then

$$\begin{aligned} \epsilon_1' &= \epsilon_1 + \epsilon_2 \\ \epsilon_2' &= 2\epsilon_1 + \epsilon_2 \end{aligned}$$

so that P is the matrix

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

By a short computation

$$P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Thus

$$\begin{aligned} [T]_{\mathcal{B}'} &= P^{-1}[T]_{\mathcal{B}}P \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}. \end{aligned}$$

We can easily check that this is correct because

$$T\epsilon_1' = (1, 0) = -\epsilon_1' + \epsilon_2'$$
$$T\epsilon_2' = (2, 0) = -2\epsilon_1' + 2\epsilon_2'.$$

EXAMPLE 17. Let V be the space of polynomial functions from R into R which have 'degree' less than or equal to 3. As in Example 15, let D be the differentiation operator on V, and let

$$\mathcal{B} = \{f_1, f_2, f_3, f_4\}$$

be the ordered basis for V defined by $f_i(x) = x^{i-1}$. Let t be a real number and define $g_i(x) = (x + t)^{i-1}$, that is

$$\begin{aligned} g_1 &= f_1 \\ g_2 &= tf_1 + f_2 \\ g_3 &= t^2f_1 + 2tf_2 + f_3 \\ g_4 &= t^3f_1 + 3t^2f_2 + 3tf_3 + f_4. \end{aligned}$$

Since the matrix

$$P = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is easily seen to be invertible with

$$P^{-1} = \begin{bmatrix} 1 & -t & t^2 & -t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

it follows that $\mathcal{B}' = \{g_1, g_2, g_3, g_4\}$ is an ordered basis for V. In Example 15, we found that the matrix of D in the ordered basis $\mathcal{B}$ is

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix of D in the ordered basis $\mathcal{B}'$ is thus

$$\begin{aligned} P^{-1}[D]_{\mathcal{B}}P &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -t & t^2 & t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 2 & 6t \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Thus D is represented by the same matrix in the ordered bases $\mathcal{B}$ and $\mathcal{B}'$. Of course, one can see this somewhat more directly since

$$\begin{aligned} Dg_1 &= 0 \\ Dg_2 &= g_1 \\ Dg_3 &= 2g_2 \\ Dg_4 &= 3g_3. \end{aligned}$$

This example illustrates a good point. If one knows the matrix of a linear operator in some ordered basis $\mathcal{B}$ and wishes to find the matrix in another ordered basis $\mathcal{B}'$, it is often most convenient to perform the coordinate change using the invertible matrix P; however, it may be a much simpler task to find the representing matrix by a direct appeal to its definition.

Definition. *Let* A *and* B *be* n $\times$ n *(square) matrices over the field* F. *We say that* B **is similar to** A **over** F *if there is an invertible* n $\times$ n *matrix* P *over* F *such that* B = P^{-1}AP.

According to Theorem 14, we have the following: If V is an n-dimensional vector space over F and $\mathcal{B}$ and $\mathcal{B}'$ are two ordered bases for V, then for each linear operator T on V the matrix $B = [T]_{\mathcal{B}'}$ is similar to the matrix $A = [T]_{\mathcal{B}}$. The argument also goes in the other direction. Suppose A and B are $n \times n$ matrices and that B is similar to A. Let V be any n-dimensional space over F and let $\mathcal{B}$ be an ordered basis for V. Let T be the linear operator on V which is represented in the basis $\mathcal{B}$ by A. If $B = P^{-1}AP$, let $\mathcal{B}'$ be the ordered basis for V obtained from $\mathcal{B}$ by P, i.e.,

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

Then the matrix of T in the ordered basis $\mathcal{B}'$ will be B.

Thus the statement that B is similar to A means that on each n-dimensional space over F the matrices A and B represent the same linear transformation in two (possibly) different ordered bases.

Note that each $n \times n$ matrix A is similar to itself, using $P = I$; if B is similar to A, then A is similar to B, for $B = P^{-1}AP$ implies that $A = (P^{-1})^{-1}BP^{-1}$; if B is similar to A and C is similar to B, then C is similar to A, for $B = P^{-1}AP$ and $C = Q^{-1}BQ$ imply that $C = (PQ)^{-1}A(PQ)$. Thus, similarity is an equivalence relation on the set of $n \times n$ matrices over the field F. Also note that the only matrix similar to the identity matrix I is I itself, and that the only matrix similar to the zero matrix is the zero matrix itself.

Exercises

1. Let T be the linear operator on C^2 defined by $T(x_1, x_2) = (x_1, 0)$. Let $\mathcal{B}$ be the standard ordered basis for C^2 and let $\mathcal{B}' = \{\alpha_1, \alpha_2\}$ be the ordered basis defined by $\alpha_1 = (1, i)$, $\alpha_2 = (-i, 2)$.

(a) What is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?
(b) What is the matrix of T relative to the pair $\mathcal{B}'$, $\mathcal{B}$?
(c) What is the matrix of T in the ordered basis $\mathcal{B}'$?
(d) What is the matrix of T in the ordered basis $\{\alpha_2, \alpha_1\}$?

2. Let T be the linear transformation from R^3 into R^2 defined by

$$T(x_1, x_2, x_3) = (x_1 + x_2, 2x_3 - x_1).$$

(a) If $\mathcal{B}$ is the standard ordered basis for R^3 and $\mathcal{B}'$ is the standard ordered basis for R^2, what is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?

(b) If $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ and $\mathcal{B}' = \{\beta_1, \beta_2\}$, where

$$\alpha_1 = (1, 0, -1), \quad \alpha_2 = (1, 1, 1), \quad \alpha_3 = (1, 0, 0), \quad \beta_1 = (0, 1), \quad \beta_2 = (1, 0)$$

what is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$?

3. Let T be a linear operator on F^n, let A be the matrix of T in the standard ordered basis for F^n, and let W be the subspace of F^n spanned by the column vectors of A. What does W have to do with T?

4. Let V be a two-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. If T is a linear operator on V and

$$[T]_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

prove that $T^2 - (a + d)T + (ad - bc)I = 0$.

5. Let T be the linear operator on R^3, the matrix of which in the standard ordered basis is

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

Find a basis for the range of T and a basis for the null space of T.

6. Let T be the linear operator on R^2 defined by

$$T(x_1, x_2) = (-x_2, x_1).$$

(a) What is the matrix of T in the standard ordered basis for R^2?

(b) What is the matrix of T in the ordered basis $\mathcal{B} = \{\alpha_1, \alpha_2\}$, where $\alpha_1 = (1, 2)$ and $\alpha_2 = (1, -1)$?

(c) Prove that for every real number c the operator $(T - cI)$ is invertible.

(d) Prove that if $\mathcal{B}$ is any ordered basis for R^2 and $[T]_{\mathcal{B}} = A$, then $A_{12}A_{21} \neq 0$.

7. Let T be the linear operator on R^3 defined by

$$T(x_1, x_2, x_3) = (3x_1 + x_3, -2x_1 + x_2, -x_1 + 2x_2 + 4x_3).$$

(a) What is the matrix of T in the standard ordered basis for R^3?

(b) What is the matrix of T in the ordered basis

$$\{\alpha_1, \alpha_2, \alpha_3\}$$

where $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (-1, 2, 1)$, and $\alpha_3 = (2, 1, 1)$?

(c) Prove that T is invertible and give a rule for T^{-1} like the one which defines T.

8. Let θ be a real number. Prove that the following two matrices are similar over the field of complex numbers:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

(*Hint:* Let T be the linear operator on C^2 which is represented by the first matrix in the standard ordered basis. Then find vectors α_1 and α_2 such that $T\alpha_1 = e^{i\theta}\alpha_1$, $T\alpha_2 = e^{-i\theta}\alpha_2$, and $\{\alpha_1, \alpha_2\}$ is a basis.)

9. Let V be a finite-dimensional vector space over the field F and let S and T be linear operators on V. We ask: When do there exist ordered bases $\mathcal{B}$ and $\mathcal{B}'$ for V such that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$? Prove that such bases exist if and only if there is an invertible linear operator U on V such that $T = USU^{-1}$. (*Outline of proof:* If $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$, let U be the operator which carries $\mathcal{B}$ onto $\mathcal{B}'$ and show that $S = UTU^{-1}$. Conversely, if $T = USU^{-1}$ for some invertible U, let $\mathcal{B}$ be any ordered basis for V and let $\mathcal{B}'$ be its image under U. Then show that $[S]_{\mathcal{B}} = [T]_{\mathcal{B}'}$.)

10. We have seen that the linear operator T on R^2 defined by $T(x_1, x_2) = (x_1, 0)$ is represented in the standard ordered basis by the matrix

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

This operator satisfies $T^2 = T$. Prove that if S is a linear operator on R^2 such that $S^2 = S$, then $S = 0$, or $S = I$, or there is an ordered basis $\mathcal{B}$ for R^2 such that $[S]_{\mathcal{B}} = A$ (above).

11. Let W be the space of all $n \times 1$ column matrices over a field F. If A is an $n \times n$ matrix over F, then A defines a linear operator L_A on W through left multiplication: $L_A(X) = AX$. Prove that every linear operator on W is left multiplication by some $n \times n$ matrix, i.e., is L_A for some A.

Now suppose V is an n-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis for V. For each α in V, define $U\alpha = [\alpha]_{\mathcal{B}}$. Prove that U is an isomorphism of V onto W. If T is a linear operator on V, then UTU^{-1} is a linear operator on W. Accordingly, UTU^{-1} is left multiplication by some $n \times n$ matrix A. What is A?

12. Let V be an n-dimensional vector space over the field F, and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V.

(a) According to Theorem 1, there is a unique linear operator T on V such that

$$T\alpha_j = \alpha_{j+1}, \qquad j = 1, \ldots, n-1, \qquad T\alpha_n = 0.$$

What is the matrix A of T in the ordered basis $\mathcal{B}$?

(b) Prove that $T^n = 0$ but $T^{n-1} \neq 0$.

(c) Let S be any linear operator on V such that $S^n = 0$ but $S^{n-1} \neq 0$. Prove that there is an ordered basis $\mathcal{B}'$ for V such that the matrix of S in the ordered basis $\mathcal{B}'$ is the matrix A of part (a).

(d) Prove that if M and N are $n \times n$ matrices over F such that $M^n = N^n = 0$ but $M^{n-1} \neq 0 \neq N^{n-1}$, then M and N are similar.

13. Let V and W be finite-dimensional vector spaces over the field F and let T be a linear transformation from V into W. If

$$\mathcal{B} = \{\alpha_1, \dots, \alpha_n\} \quad \text{and} \quad \mathcal{B}' = \{\beta_1, \dots, \beta_m\}$$

are ordered bases for V and W, respectively, define the linear transformations $E^{p,q}$ as in the proof of Theorem 5: $E^{p,q}(\alpha_i) = \delta_{iq}\beta_p$. Then the $E^{p,q}$, $1 \leq p \leq m$, $1 \leq q \leq n$, form a basis for $L(V, W)$, and so

$$T = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{pq}E^{p,q}$$

for certain scalars A_{pq} (the coordinates of T in this basis for $L(V, W)$). Show that the matrix A with entries $A(p, q) = A_{pq}$ is precisely the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$.

3.5. Linear Functionals

If V is a vector space over the field F, a linear transformation f from V into the scalar field F is also called a **linear functional** on V. If we start from scratch, this means that f is a function from V into F such that

$$f(c\alpha + \beta) = cf(\alpha) + f(\beta)$$

for all vectors α and β in V and all scalars c in F. The concept of linear functional is important in the study of finite-dimensional spaces because it helps to organize and clarify the discussion of subspaces, linear equations, and coordinates.

EXAMPLE 18. Let F be a field and let $a_1, \dots, a_n$ be scalars in F. Define a function f on F^n by

$$f(x_1, \dots, x_n) = a_1x_1 + \cdots + a_nx_n.$$

Then f is a linear functional on F^n. It is the linear functional which is represented by the matrix $[a_1 \cdots a_n]$ relative to the standard ordered basis for F^n and the basis $\{1\}$ for F:

$$a_j = f(\epsilon_j), \qquad j = 1, \dots, n.$$

Every linear functional on F^n is of this form, for some scalars $a_1, \dots, a_n$. That is immediate from the definition of linear functional because we define $a_j = f(\epsilon_j)$ and use the linearity

$$\begin{aligned} f(x_1, \dots, x_n) &= f\left(\sum_j x_j\epsilon_j\right) \\ &= \sum_j x_jf(\epsilon_j) \\ &= \sum_j a_jx_j. \end{aligned}$$

EXAMPLE 19. Here is an important example of a linear functional. Let n be a positive integer and F a field. If A is an $n \times n$ matrix with entries in F, the **trace** of A is the scalar

$$\operatorname{tr} A = A_{11} + A_{22} + \cdots + A_{nn}.$$

The trace function is a linear functional on the matrix space $F^{n\times n}$ because

$$\begin{aligned}\operatorname{tr}(cA + B) &= \sum_{i=1}^{n} (cA_{ii} + B_{ii})\\ &= c\sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii}\\ &= c \operatorname{tr} A + \operatorname{tr} B.\end{aligned}$$

EXAMPLE 20. Let V be the space of all polynomial functions from the field F into itself. Let t be an element of F. If we define

$$L_t(p) = p(t)$$

then L_t is a linear functional on V. One usually describes this by saying that, for each t, 'evaluation at t' is a linear functional on the space of polynomial functions. Perhaps we should remark that the fact that the functions are polynomials plays no role in this example. Evaluation at t is a linear functional on the space of all functions from F into F.

EXAMPLE 21. This may be the most important linear functional in mathematics. Let $[a, b]$ be a closed interval on the real line and let $C([a, b])$ be the space of continuous real-valued functions on $[a, b]$. Then

$$L(g) = \int_a^b g(t)\, dt$$

defines a linear functional L on $C([a, b])$.

If V is a vector space, the collection of all linear functionals on V forms a vector space in a natural way. It is the space $L(V, F)$. We denote this space by V^* and call it the **dual space** of V:

$$V^* = L(V, F).$$

If V is finite-dimensional, we can obtain a rather explicit description of the dual space V^*. From Theorem 5 we know something about the space V^*, namely that

$$\dim V^* = \dim V.$$

Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. According to Theorem 1, there is (for each i) a unique linear functional f_i on V such that

$$f_i(\alpha_j) = \delta_{ij}. \tag{3-11}$$

In this way we obtain from $\mathcal{B}$ a set of n distinct linear functionals $f_1, \ldots, f_n$ on V. These functionals are also linearly independent. For, suppose

$$f = \sum_{i=1}^{n} c_i f_i. \tag{3-12}$$

Then

$$\begin{aligned} f(\alpha_j) &= \sum_{i=1}^{n} c_i f_i(\alpha_j) \\ &= \sum_{i=1}^{n} c_i \delta_{ij} \\ &= c_j. \end{aligned}$$

In particular, if f is the zero functional, $f(\alpha_j) = 0$ for each j and hence the scalars c_j are all 0. Now $f_1, \ldots, f_n$ are n linearly independent functionals, and since we know that V^* has dimension n, it must be that $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ is a basis for V^*. This basis is called the **dual basis** of $\mathcal{B}$.

Theorem 15. *Let* V *be a finite-dimensional vector space over the field* F, *and let* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *be a basis for* V. *Then there is a unique dual basis* $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ *for* V^* *such that* $f_i(\alpha_j) = \delta_{ij}$. *For each linear functional* f *on* V *we have*

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i \tag{3-13}$$

and for each vector α *in* V *we have*

$$\alpha = \sum_{i=1}^{n} f_i(\alpha) \alpha_i. \tag{3-14}$$

Proof. We have shown above that there is a unique basis which is 'dual' to $\mathcal{B}$. If f is a linear functional on V, then f is some linear combination (3-12) of the f_i, and as we observed after (3-12) the scalars c_j must be given by $c_j = f(\alpha_j)$. Similarly, if

$$\alpha = \sum_{i=1}^{n} x_i \alpha_i$$

is a vector in V, then

$$\begin{aligned} f_j(\alpha) &= \sum_{i=1}^{n} x_i f_j(\alpha_i) \\ &= \sum_{i=1}^{n} x_i \delta_{ij} \\ &= x_j \end{aligned}$$

so that the unique expression for α as a linear combination of the α_i is

$$\alpha = \sum_{i=1}^{n} f_i(\alpha) \alpha_i. \quad \blacksquare$$

Equation (3-14) provides us with a nice way of describing what the dual basis is. It says, if $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V and

$\mathcal{B}^* = \{f_1, \ldots, f_n\}$ is the dual basis, then f_i is precisely the function which assigns to each vector α in V the ith coordinate of α relative to the ordered basis $\mathcal{B}$. Thus we may also call the f_i the coordinate functions for $\mathcal{B}$. The formula (3-13), when combined with (3-14) tells us the following: If f is in V^*, and we let $f(\alpha_i) = a_i$, then when

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$$

we have

$$f(\alpha) = a_1x_1 + \cdots + a_nx_n. \tag{3-15}$$

In other words, if we choose an ordered basis $\mathcal{B}$ for V and describe each vector in V by its n-tuple of coordinates $(x_1, \ldots, x_n)$ relative to $\mathcal{B}$, then every linear functional on V has the form (3-15). This is the natural generalization of Example 18, which is the special case $V = F^n$ and $\mathcal{B} = \{\epsilon_1, \ldots, \epsilon_n\}$.

Example 22. Let V be the vector space of all polynomial functions from R into R which have degree less than or equal to 2. Let t_1, t_2, and t_3 be any three *distinct* real numbers, and let

$$L_i(p) = p(t_i).$$

Then L_1, L_2, and L_3 are linear functionals on V. These functionals are linearly independent; for, suppose

$$L = c_1L_1 + c_2L_2 + c_3L_3.$$

If $L = 0$, i.e., if $L(p) = 0$ for each p in V, then applying L to the particular polynomial 'functions' 1, x, x^2, we obtain

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ t_1c_1 + t_2c_2 + t_3c_3 &= 0 \\ t_1^2c_1 + t_2^2c_2 + t_3^2c_3 &= 0 \end{aligned}$$

From this it follows that $c_1 = c_2 = c_3 = 0$, because (as a short computation shows) the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ t_1 & t_2 & t_3 \\ t_1^2 & t_2^2 & t_3^2 \end{bmatrix}$$

is invertible when t_1, t_2, and t_3 are distinct. Now the L_i are independent, and since V has dimension 3, these functionals form a basis for V^*. What is the basis for V, of which this is the dual? Such a basis $\{p_1, p_2, p_3\}$ for V must satisfy

$$L_i(p_j) = \delta_{ij}$$

or

$$p_j(t_i) = \delta_{ij}.$$

These polynomial functions are rather easily seen to be

$$p_1(x) = \frac{(x - t_2)(x - t_3)}{(t_1 - t_2)(t_1 - t_3)}$$
$$p_2(x) = \frac{(x - t_1)(x - t_3)}{(t_2 - t_1)(t_2 - t_3)}$$
$$p_3(x) = \frac{(x - t_1)(x - t_2)}{(t_3 - t_1)(t_3 - t_2)}.$$

The basis $\{p_1, p_2, p_3\}$ for V is interesting, because according to (3-14) we have for each p in V

$$p = p(t_1)p_1 + p(t_2)p_2 + p(t_3)p_3.$$

Thus, if c_1, c_2, and c_3 are any real numbers, there is exactly one polynomial function p over R which has degree at most 2 and satisfies $p(t_j) = c_j$, $j = 1, 2, 3$. This polynomial function is $p = c_1p_1 + c_2p_2 + c_3p_3$.

Now let us discuss the relationship between linear functionals and subspaces. If f is a non-zero linear functional, then the rank of f is 1 because the range of f is a non-zero subspace of the scalar field and must (therefore) be the scalar field. If the underlying space V is finite-dimensional, the rank plus nullity theorem (Theorem 2) tells us that the null space N_f has dimension

$$\dim N_f = \dim V - 1.$$

In a vector space of dimension n, a subspace of dimension $n - 1$ is called a **hyperspace.** Such spaces are sometimes called hyperplanes or subspaces of codimension 1. Is every hyperspace the null space of a linear functional? The answer is easily seen to be yes. It is not much more difficult to show that each d-dimensional subspace of an n-dimensional space is the intersection of the null spaces of $(n - d)$ linear functionals (Theorem 16 below).

Definition. *If* V *is a vector space over the field* F *and* S *is a subset of* V, *the* **annihilator** *of* S *is the set* S^0 *of linear functionals* f *on* V *such that* $f(\alpha) = 0$ *for every* α *in* S.

It should be clear to the reader that S^0 is a subspace of V^*, whether S is a subspace of V or not. If S is the set consisting of the zero vector alone, then $S^0 = V^*$. If $S = V$, then S^0 is the zero subspace of V^*. (This is easy to see when V is finite-dimensional.)

Theorem 16. *Let* V *be a finite-dimensional vector space over the field* F, *and let* W *be a subspace of* V. *Then*

$$\dim W + \dim W^0 = \dim V.$$

Proof. Let k be the dimension of W and $\{\alpha_1, \ldots, \alpha_k\}$ a basis for W. Choose vectors $\alpha_{k+1}, \ldots, \alpha_n$ in V such that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. Let $\{f_1, \ldots, f_n\}$ be the basis for V^* which is dual to this basis for V.

The claim is that $\{f_{k+1}, \ldots, f_n\}$ is a basis for the annihilator W^0. Certainly f_i belongs to W^0 for $i \geq k+1$, because

$$f_i(\alpha_j) = \delta_{ij}$$

and $\delta_{ij} = 0$ if $i \geq k+1$ and $j \leq k$; from this it follows that, for $i \geq k+1$, $f_i(\alpha) = 0$ whenever α is a linear combination of $\alpha_1, \ldots, \alpha_k$. The functionals $f_{k+1}, \ldots, f_n$ are independent, so all we must show is that they span W^0. Suppose f is in V^*. Now

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i$$

so that if f is in W^0 we have $f(\alpha_i) = 0$ for $i \leq k$ and

$$f = \sum_{i=k+1}^{n} f(\alpha_i) f_i.$$

We have shown that if dim $W = k$ and dim $V = n$ then dim $W^0 = n - k$. ▮

Corollary. *If* W *is a* k*-dimensional subspace of an* n*-dimensional vector space* V, *then* W *is the intersection of* (n − k) *hyperspaces in* V.

Proof. This is a corollary of the proof of Theorem 16 rather than its statement. In the notation of the proof, W is exactly the set of vectors α such that $f_i(\alpha) = 0$, $i = k+1, \ldots, n$. In case $k = n - 1$, W is the null space of f_n. ▮

Corollary. *If* W_1 *and* W_2 *are subspaces of a finite-dimensional vector space, then* $W_1 = W_2$ *if and only if* $W_1^0 = W_2^0$.

Proof. If $W_1 = W_2$, then of course $W_1^0 = W_2^0$. If $W_1 \neq W_2$, then one of the two subspaces contains a vector which is not in the other. Suppose there is a vector α which is in W_2 but not in W_1. By the previous corollaries (or the proof of Theorem 16) there is a linear functional f such that $f(\beta) = 0$ for all β in W, but $f(\alpha) \neq 0$. Then f is in W_1^0 but not in W_2^0 and $W_1^0 \neq W_2^0$. ▮

In the next section we shall give different proofs for these two corollaries. The first corollary says that, if we select some ordered basis for the space, each k-dimensional subspace can be described by specifying $(n - k)$ homogeneous linear conditions on the coordinates relative to that basis.

Let us look briefly at systems of homogeneous linear equations from the point of view of linear functionals. Suppose we have a system of linear equations,

$$\begin{aligned} A_{11}x_1 + \cdots + A_{1n}x_n &= 0 \\ \vdots \qquad\qquad & \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n &= 0 \end{aligned}$$

for which we wish to find the solutions. If we let f_i, $i = 1, \ldots, m$, be the linear functional on F^n defined by

$$f_i(x_1, \ldots, x_n) = A_{i1}x_1 + \cdots + A_{in}x_n$$

then we are seeking the subspace of F^n of all α such that

$$f_i(\alpha) = 0, \qquad i = 1, \ldots, m.$$

In other words, we are seeking the subspace annihilated by $f_1, \ldots, f_m$. Row-reduction of the coefficient matrix provides us with a systematic method of finding this subspace. The n-tuple $(A_{i1}, \ldots, A_{in})$ gives the coordinates of the linear functional f_i relative to the basis which is dual to the standard basis for F^n. The row space of the coefficient matrix may thus be regarded as the space of linear functionals spanned by $f_1, \ldots, f_m$. The solution space is the subspace annihilated by this space of functionals.

Now one may look at the system of equations from the 'dual' point of view. That is, suppose that we are given m vectors in F^n

$$\alpha_i = (A_{i1}, \ldots, A_{in})$$

and we wish to find the annihilator of the subspace spanned by these vectors. Since a typical linear functional on F^n has the form

$$f(x_1, \ldots, x_n) = c_1x_1 + \cdots + c_nx_n$$

the condition that f be in this annihilator is that

$$\sum_{j=1}^{n} A_{ij}c_j = 0, \qquad i = 1, \ldots, m$$

that is, that $(c_1, \ldots, c_n)$ be a solution of the system $AX = 0$. From this point of view, row-reduction gives us a systematic method of finding the annihilator of the subspace spanned by a given finite set of vectors in F^n.

Example 23. Here are three linear functionals on R^4:

$$\begin{aligned} f_1(x_1, x_2, x_3, x_4) &= x_1 + 2x_2 + 2x_3 + x_4 \\ f_2(x_1, x_2, x_3, x_4) &= 2x_2 + x_4 \\ f_3(x_1, x_2, x_3, x_4) &= -2x_1 - 4x_3 + 3x_4. \end{aligned}$$

The subspace which they annihilate may be found explicitly by finding the row-reduced echelon form of the matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 2 & 0 & 1 \\ -2 & 0 & -4 & 3 \end{bmatrix}.$$

A short calculation, or a peek at Example 21 of Chapter 2, shows that

$$R = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Therefore, the linear functionals

$$\begin{aligned} g_1(x_1, x_2, x_3, x_4) &= x_1 + 2x_3 \\ g_2(x_1, x_2, x_3, x_4) &= x_2 \\ g_3(x_1, x_2, x_3, x_4) &= x_4 \end{aligned}$$

span the same subspace of $(R^4)^*$ and annihilate the same subspace of R^4 as do f_1, f_2, f_3. The subspace annihilated consists of the vectors with

$$\begin{aligned} x_1 &= -2x_3 \\ x_2 &= x_4 = 0. \end{aligned}$$

EXAMPLE 24. Let W be the subspace of R^5 which is spanned by the vectors

$$\begin{aligned} \alpha_1 &= (2, -2, 3, 4, -1), & \alpha_3 &= (0, 0, -1, -2, 3) \\ \alpha_2 &= (-1, 1, 2, 5, 2), & \alpha_4 &= (1, -1, 2, 3, 0). \end{aligned}$$

How does one describe W^0, the annihilator of W? Let us form the 4×5 matrix A with row vectors $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, and find the row-reduced echelon matrix R which is row-equivalent to A:

$$A = \begin{bmatrix} 2 & -2 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 2 \\ 0 & 0 & -1 & -2 & 3 \\ 1 & -1 & 2 & 3 & 0 \end{bmatrix} \longrightarrow R = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

If f is a linear functional on R^5:

$$f(x_1, \ldots, x_5) = \sum_{j=1}^{5} c_j x_j$$

then f is in W^0 if and only if $f(\alpha_i) = 0$, $i = 1, 2, 3, 4$, i.e., if and only if

$$\sum_{j=1}^{5} A_{ij} c_j = 0, \qquad 1 \le i \le 4.$$

This is equivalent to

$$\sum_{j=1}^{5} R_{ij} c_j = 0, \qquad 1 \le i \le 3$$

or

$$\begin{aligned} c_1 - c_2 - c_4 &= 0 \\ c_3 + 2c_4 &= 0 \\ c_5 &= 0. \end{aligned}$$

We obtain all such linear functionals f by assigning arbitrary values to c_2 and c_4, say $c_2 = a$ and $c_4 = b$, and then finding the corresponding $c_1 = a + b$, $c_3 = -2b$, $c_5 = 0$. So W^0 consists of all linear functionals f of the form

$$f(x_1, x_2, x_3, x_4, x_5) = (a + b)x_1 + ax_2 - 2bx_3 + bx_4.$$

The dimension of W^0 is 2 and a basis $\{f_1, f_2\}$ for W^0 can be found by first taking $a = 1$, $b = 0$ and then $a = 0$, $b = 1$:

$$f_1(x_1, \ldots, x_5) = x_1 + x_2$$
$$f_2(x_1, \ldots, x_5) = x_1 - 2x_3 + x_4.$$

The above general f in W^0 is $f = af_1 + bf_2$.

Exercises

1. In R^3, let $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (0, 1, -2)$, $\alpha_3 = (-1, -1, 0)$.

(a) If f is a linear functional on R^3 such that

$$f(\alpha_1) = 1, \qquad f(\alpha_2) = -1, \qquad f(\alpha_3) = 3,$$

and if $\alpha = (a, b, c)$, find $f(\alpha)$.

(b) Describe explicitly a linear functional f on R^3 such that

$$f(\alpha_1) = f(\alpha_2) = 0 \quad \text{but} \quad f(\alpha_3) \neq 0.$$

(c) Let f be any linear functional such that

$$f(\alpha_1) = f(\alpha_2) = 0 \quad \text{and} \quad f(\alpha_3) \neq 0.$$

If $\alpha = (2, 3, -1)$, show that $f(\alpha) \neq 0$.

2. Let $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$ be the basis for C^3 defined by

$$\alpha_1 = (1, 0, -1), \qquad \alpha_2 = (1, 1, 1), \qquad \alpha_3 = (2, 2, 0).$$

Find the dual basis of $\mathcal{B}$.

3. If A and B are $n \times n$ matrices over the field F, show that trace (AB) = trace (BA). Now show that similar matrices have the same trace.

4. Let V be the vector space of all polynomial functions p from R into R which have degree 2 or less:

$$p(x) = c_0 + c_1x + c_2x^2.$$

Define three linear functionals on V by

$$f_1(p) = \int_0^1 p(x)\,dx, \qquad f_2(p) = \int_0^2 p(x)\,dx, \qquad f_3(p) = \int_0^{-1} p(x)\,dx.$$

Show that $\{f_1, f_2, f_3\}$ is a basis for V^* by exhibiting the basis for V of which it is the dual.

5. If A and B are $n \times n$ complex matrices, show that $AB - BA = I$ is impossible.

6. Let m and n be positive integers and F a field. Let $f_1, \ldots, f_m$ be linear functionals on F^n. For α in F^n define

$$T\alpha = (f_1(\alpha), \ldots, f_m(\alpha)).$$

Show that T is a linear transformation from F^n into F^m. Then show that every linear transformation from F^n into F^m is of the above form, for some $f_1, \ldots, f_m$.

7. Let $\alpha_1 = (1, 0, -1, 2)$ and $\alpha_2 = (2, 3, 1, 1)$, and let W be the subspace of R^4 spanned by α_1 and α_2. Which linear functionals f:

$$f(x_1, x_2, x_3, x_4) = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$$

are in the annihilator of W?

8. Let W be the subspace of R^5 which is spanned by the vectors

$$\alpha_1 = \epsilon_1 + 2\epsilon_2 + \epsilon_3, \qquad \alpha_2 = \epsilon_2 + 3\epsilon_3 + 3\epsilon_4 + \epsilon_5$$
$$\alpha_3 = \epsilon_1 + 4\epsilon_2 + 6\epsilon_3 + 4\epsilon_4 + \epsilon_5.$$

Find a basis for W^0.

9. Let V be the vector space of all 2×2 matrices over the field of real numbers, and let

$$B = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}.$$

Let W be the subspace of V consisting of all A such that $AB = 0$. Let f be a linear functional on V which is in the annihilator of W. Suppose that $f(I) = 0$ and $f(C) = 3$, where I is the 2×2 identity matrix and

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Find $f(B)$.

10. Let F be a subfield of the complex numbers. We define n linear functionals on F^n $(n \geq 2)$ by

$$f_k(x_1, \ldots, x_n) = \sum_{j=1}^{n} (k - j)x_j, \qquad 1 \leq k \leq n.$$

What is the dimension of the subspace annihilated by $f_1, \ldots, f_n$?

11. Let W_1 and W_2 be subspaces of a finite-dimensional vector space V.

(a) Prove that $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$.

(b) Prove that $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

12. Let V be a finite-dimensional vector space over the field F and let W be a subspace of V. If f is a linear functional on W, prove that there is a linear functional g on V such that $g(\alpha) = f(\alpha)$ for each α in the subspace W.

13. Let F be a subfield of the field of complex numbers and let V be any vector space over F. Suppose that f and g are linear functionals on V such that the function h defined by $h(\alpha) = f(\alpha)g(\alpha)$ is also a linear functional on V. Prove that either $f = 0$ or $g = 0$.

14. Let F be a field of characteristic zero and let V be a finite-dimensional vector space over F. If $\alpha_1, \ldots, \alpha_m$ are finitely many vectors in V, each different from the zero vector, prove that there is a linear functional f on V such that

$$f(\alpha_i) \neq 0, \qquad i = 1, \ldots, m.$$

15. According to Exercise 3, similar matrices have the same trace. Thus we can define the trace of a linear operator on a finite-dimensional space to be the trace of any matrix which represents the operator in an ordered basis. This is well-defined since all such representing matrices for one operator are similar.

Now let V be the space of all 2×2 matrices over the field F and let P be a fixed 2×2 matrix. Let T be the linear operator on V defined by $T(A) = PA$. Prove that trace $(T) = 2$ trace (P).

16. Show that the trace functional on $n \times n$ matrices is unique in the following sense. If W is the space of $n \times n$ matrices over the field F and if f is a linear functional on W such that $f(AB) = f(BA)$ for each A and B in W, then f is a scalar multiple of the trace function. If, in addition, $f(I) = n$, then f is the trace function.

17. Let W be the space of $n \times n$ matrices over the field F, and let W_0 be the subspace spanned by the matrices C of the form $C = AB - BA$. Prove that W_0 is exactly the subspace of matrices which have trace zero. (*Hint:* What is the dimension of the space of matrices of trace zero? Use the matrix 'units,' i.e., matrices with exactly one non-zero entry, to construct enough linearly independent matrices of the form $AB - BA$.)

3.6. The Double Dual

One question about dual bases which we did not answer in the last section was whether every basis for V^* is the dual of some basis for V. One way to answer that question is to consider V^{**}, the dual space of V^*.

If α is a vector in V, then α induces a linear functional L_α on V^* defined by

$$L_\alpha(f) = f(\alpha), \qquad f \text{ in } V^*.$$

The fact that L_α is linear is just a reformulation of the definition of linear operations in V^*:

$$\begin{aligned} L_\alpha(cf + g) &= (cf + g)(\alpha) \\ &= (cf)(\alpha) + g(\alpha) \\ &= cf(\alpha) + g(\alpha) \\ &= cL_\alpha(f) + L_\alpha(g). \end{aligned}$$

If V is finite-dimensional and $\alpha \neq 0$, then $L_\alpha \neq 0$; in other words, there exists a linear functional f such that $f(\alpha) \neq 0$. The proof is very simple and was given in Section 3.5: Choose an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V such that $\alpha_1 = \alpha$ and let f be the linear functional which assigns to each vector in V its first coordinate in the ordered basis $\mathcal{B}$.

Theorem 17. *Let* V *be a finite-dimensional vector space over the field* F. *For each vector* α *in* V *define*

$$L_\alpha(f) = f(\alpha), \qquad f \text{ in } V^*.$$

The mapping $\alpha \to L_\alpha$ *is then an isomorphism of* V *onto* V^{**}.

Proof. We showed that for each α the function L_α is linear. Suppose α and β are in V and c is in F, and let $\gamma = c\alpha + \beta$. Then for each f in V^*

$$\begin{aligned} L_\gamma(f) &= f(\gamma) \\ &= f(c\alpha + \beta) \\ &= cf(\alpha) + f(\beta) \\ &= cL_\alpha(f) + L_\beta(f) \end{aligned}$$

and so

$$L_\gamma = cL_\alpha + L_\beta.$$

This shows that the mapping $\alpha \to L_\alpha$ is a linear transformation from V into V^{**}. This transformation is non-singular; for, according to the remarks above $L_\alpha = 0$ if and only if $\alpha = 0$. Now $\alpha \to L_\alpha$ is a non-singular linear transformation from V into V^{**}, and since

$$\dim V^{**} = \dim V^* = \dim V$$

Theorem 9 tells us that this transformation is invertible, and is therefore an isomorphism of V onto V^{**}. ▌

Corollary. *Let* V *be a finite-dimensional vector space over the field* F. *If* L *is a linear functional on the dual space* V* *of* V, *then there is a unique vector* α *in* V *such that*

$$L(f) = f(\alpha)$$

for every f *in* V*.

Corollary. *Let* V *be a finite-dimensional vector space over the field* F. *Each basis for* V* *is the dual of some basis for* V.

Proof. Let $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ be a basis for V^*. By Theorem 15, there is a basis $\{L_1, \ldots, L_n\}$ for V^{**} such that

$$L_i(f_j) = \delta_{ij}.$$

Using the corollary above, for each i there is a vector α_i in V such that

$$L_i(f) = f(\alpha_i)$$

for every f in V^*, i.e., such that $L_i = L_{\alpha_i}$. It follows immediately that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V and that $\mathcal{B}^*$ is the dual of this basis. ▌

In view of Theorem 17, we usually identify α with L_α and say that V 'is' the dual space of V^* or that the spaces V, V^* are naturally in duality with one another. Each is the dual space of the other. In the last corollary we have an illustration of how that can be useful. Here is a further illustration.

If E is a subset of V^*, then the annihilator E^0 is (technically) a subset of V^{**}. If we choose to identify V and V^{**} as in Theorem 17, then E^0 is a subspace of V, namely, the set of all α in V such that $f(\alpha) = 0$ for all f in E. In a corollary of Theorem 16 we noted that each subspace W is determined by its annihilator W^0. How is it determined? The answer is that W is the subspace annihilated by all f in W^0, that is, the intersection of the null spaces of all f's in W^0. In our present notation for annihilators, the answer may be phrased very simply: $W = (W^0)^0$.

Theorem 18. *If* S *is any subset of a finite-dimensional vector space* V, *then* $(S^0)^0$ *is the subspace spanned by* S.

Proof. Let W be the subspace spanned by S. Clearly $W^0 = S^0$. Therefore, what we are to prove is that $W = W^{00}$. We have given one proof. Here is another. By Theorem 16

$$\dim W + \dim W^0 = \dim V$$
$$\dim W^0 + \dim W^{00} = \dim V^*$$

and since $\dim V = \dim V^*$ we have

$$\dim W = \dim W^{00}.$$

Since W is a subspace of W^{00}, we see that $W = W^{00}$. ∎

The results of this section hold for arbitrary vector spaces; however, the proofs require the use of the so-called Axiom of Choice. We want to avoid becoming embroiled in a lengthy discussion of that axiom, so we shall not tackle annihilators for general vector spaces. But, there are two results about linear functionals on arbitrary vector spaces which are so fundamental that we should include them.

Let V be a vector space. We want to define hyperspaces in V. Unless V is finite-dimensional, we cannot do that with the dimension of the hyperspace. But, we can express the idea that a space N falls just one dimension short of filling out V, in the following way:

1. N is a proper subspace of V;
2. if W is a subspace of V which contains N, then either $W = N$ or $W = V$.

Conditions (1) and (2) together say that N is a proper subspace and there is no larger proper subspace, in short, N is a maximal proper subspace.

Definition. *If* V *is a vector space, a* **hyperspace** *in* V *is a maximal proper subspace of* V.

Theorem 19. *If* f *is a non-zero linear functional on the vector space* V, *then the null space of* f *is a hyperspace in* V. *Conversely, every hyperspace in* V *is the null space of a (not unique) non-zero linear functional on* V.

Proof. Let f be a non-zero linear functional on V and N_f its null space. Let α be a vector in V which is not in N_f, i.e., a vector such that $f(\alpha) \neq 0$. We shall show that every vector in V is in the subspace spanned by N_f and α. That subspace consists of all vectors

$$\gamma + c\alpha, \qquad \gamma \text{ in } N_f,\ c \text{ in } F.$$

Let β be in V. Define

$$c = \frac{f(\beta)}{f(\alpha)}$$

which makes sense because $f(\alpha) \neq 0$. Then the vector $\gamma = \beta - c\alpha$ is in N_f since

$$\begin{aligned} f(\gamma) &= f(\beta - c\alpha) \\ &= f(\beta) - cf(\alpha) \\ &= 0. \end{aligned}$$

So β is in the subspace spanned by N_f and α.

Now let N be a hyperspace in V. Fix some vector α which is not in N. Since N is a maximal proper subspace, the subspace spanned by N and α is the entire space V. Therefore each vector β in V has the form

$$\beta = \gamma + c\alpha, \qquad \gamma \text{ in } N,\ c \text{ in } F.$$

The vector γ and the scalar c are uniquely determined by β. If we have also

$$\beta = \gamma' + c'\alpha, \qquad \gamma' \text{ in } N,\ c' \text{ in } F.$$

then

$$(c' - c)\alpha = \gamma - \gamma'.$$

If $c' - c \neq 0$, then α would be in N; hence, $c' = c$ and $\gamma' = \gamma$. Another way to phrase our conclusion is this: If β is in V, there is a unique scalar c such that $\beta - c\alpha$ is in N. Call that scalar $g(\beta)$. It is easy to see that g is a linear functional on V and that N is the null space of g. ∎

Lemma. *If* f *and* g *are linear functionals on a vector space* V, *then* g *is a scalar multiple of* f *if and only if the null space of* g *contains the null space of* f, *that is, if and only if* $f(\alpha) = 0$ *implies* $g(\alpha) = 0$.

Proof. If $f = 0$ then $g = 0$ as well and g is trivially a scalar multiple of f. Suppose $f \neq 0$ so that the null space N_f is a hyperspace in V. Choose some vector α in V with $f(\alpha) \neq 0$ and let

$$c = \frac{g(\alpha)}{f(\alpha)}.$$

The linear functional $h = g - cf$ is 0 on N_f, since both f and g are 0 there, and $h(\alpha) = g(\alpha) - cf(\alpha) = 0$. Thus h is 0 on the subspace spanned by N_f and α—and that subspace is V. We conclude that $h = 0$, i.e., that $g = cf$. ∎

Theorem 20. *Let* $g, f_1, \ldots, f_r$ *be linear functionals on a vector space* V *with respective null spaces* $N, N_1, \ldots, N_r$. *Then* g *is a linear combination of* $f_1, \ldots, f_r$ *if and only if* N *contains the intersection* $N_1 \cap \cdots \cap N_r$.

Proof. If $g = c_1f_1 + \cdots + c_rf_r$ and $f_i(\alpha) = 0$ for each i, then clearly $g(\alpha) = 0$. Therefore, N contains $N_1 \cap \cdots \cap N_r$.

We shall prove the converse (the 'if' half of the theorem) by induction on the number r. The preceding lemma handles the case $r = 1$. Suppose we know the result for $r = k - 1$, and let $f_1, \ldots, f_k$ be linear functionals with null spaces $N_1, \ldots, N_k$ such that $N_1 \cap \cdots \cap N_k$ is contained in N, the null space of g. Let $g', f'_1, \ldots, f'_{k-1}$ be the restrictions of $g, f_1, \ldots, f_{k-1}$ to the subspace N_k. Then $g', f'_1, \ldots, f'_{k-1}$ are linear functionals on the vector space N_k. Furthermore, if α is a vector in N_k and $f'_i(\alpha) = 0$, $i = 1, \ldots, k-1$, then α is in $N_1 \cap \cdots \cap N_k$ and so $g'(\alpha) = 0$. By the induction hypothesis (the case $r = k - 1$), there are scalars c_i such that

$$g' = c_1 f'_1 + \cdots + c_{k-1} f'_{k-1}.$$

Now let

$$h = g - \sum_{i=1}^{k-1} c_i f_i. \tag{3-16}$$

Then h is a linear functional on V and (3-16) tells us that $h(\alpha) = 0$ for every α in N_k. By the preceding lemma, h is a scalar multiple of f_k. If $h = c_k f_k$, then

$$g = \sum_{i=1}^{k} c_i f_i. \quad \blacksquare$$

Exercises

1. Let n be a positive integer and F a field. Let W be the set of all vectors $(x_1, \ldots, x_n)$ in F^n such that $x_1 + \cdots + x_n = 0$.

(a) Prove that W^0 consists of all linear functionals f of the form

$$f(x_1, \ldots, x_n) = c \sum_{j=1}^{n} x_j.$$

(b) Show that the dual space W^* of W can be 'naturally' identified with the linear functionals

$$f(x_1, \ldots, x_n) = c_1 x_1 + \cdots + c_n x_n$$

on F^n which satisfy $c_1 + \cdots + c_n = 0$.

2. Use Theorem 20 to prove the following. If W is a subspace of a finite-dimensional vector space V and if $\{g_1, \ldots, g_r\}$ is any basis for W^0, then

$$W = \bigcap_{i=1}^{r} N_{g_i}.$$

3. Let S be a set, F a field, and $V(S; F)$ the space of all functions from S into F:

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

Let W be any n-dimensional subspace of $V(S; F)$. Show that there exist points $x_1, \ldots, x_n$ in S and functions $f_1, \ldots, f_n$ in W such that $f_i(x_j) = \delta_{ij}$.

3.7. The Transpose of a Linear Transformation

Suppose that we have two vector spaces over the field F, V, and W, and a linear transformation T from V into W. Then T induces a linear transformation from W^* into V^*, as follows. Suppose g is a linear functional on W, and let

$$f(\alpha) = g(T\alpha) \tag{3-17}$$

for each α in V. Then (3-17) defines a function f from V into F, namely, the composition of T, a function from V into W, with g, a function from W into F. Since both T and g are linear, Theorem 6 tells us that f is also linear, i.e., f is a linear functional on V. Thus T provides us with a rule T^t which associates with each linear functional g on W a linear functional $f = T^t g$ on V, defined by (3-17). Note also that T^t is actually a linear transformation from W^* into V^*; for, if g_1 and g_2 are in W^* and c is a scalar

$$\begin{aligned}[T^t(cg_1 + g_2)](\alpha) &= (cg_1 + g_2)(T\alpha)\\ &= cg_1(T\alpha) + g_2(T\alpha)\\ &= c(T^t g_1)(\alpha) + (T^t g_2)(\alpha)\end{aligned}$$

so that $T^t(cg_1 + g_2) = cT^t g_1 + T^t g_2$. Let us summarize.

Theorem 21. *Let* V *and* W *be vector spaces over the field* F. *For each linear transformation* T *from* V *into* W, *there is a unique linear transformation* T^t *from* W^* *into* V^* *such that*

$$(\mathrm{T}^t\mathrm{g})(\alpha) = \mathrm{g}(\mathrm{T}\alpha)$$

for every g *in* W^* *and* α *in* V.

We shall call T^t the **transpose** of T. This transformation T^t is often called the adjoint of T; however, we shall not use this terminology.

Theorem 22. *Let* V *and* W *be vector spaces over the field* F, *and let* T *be a linear transformation from* V *into* W. *The null space of* T^t *is the annihilator of the range of* T. *If* V *and* W *are finite-dimensional, then*

(i) *rank* (T^t) = *rank* (T)
(ii) *the range of* T^t *is the annihilator of the null space of* T.

Proof. If g is in W^*, then by definition

$$(T^t g)(\alpha) = g(T\alpha)$$

for each α in V. The statement that g is in the null space of T^t means that $g(T\alpha) = 0$ for every α in V. Thus the null space of T^t is precisely the annihilator of the range of T.

Suppose that V and W are finite-dimensional, say $\dim V = n$ and $\dim W = m$. For (i): Let r be the rank of T, i.e., the dimension of the range of T. By Theorem 16, the annihilator of the range of T then has dimension $(m - r)$. By the first statement of this theorem, the nullity of T^t must be $(m - r)$. But then since T^t is a linear transformation on an m-dimensional space, the rank of T^t is $m - (m - r) = r$, and so T and T^t have the same rank. For (ii): Let N be the null space of T. Every functional in the range of T^t is in the annihilator of N; for, suppose $f = T^t g$ for some g in W^*; then, if α is in N

$$f(\alpha) = (T^t g)(\alpha) = g(T\alpha) = g(0) = 0.$$

Now the range of T^t is a subspace of the space N^0, and

$$\dim N^0 = n - \dim N = \operatorname{rank}(T) = \operatorname{rank}(T^t)$$

so that the range of T^t must be exactly N^0. ▌

Theorem 23. *Let V and W be finite-dimensional vector spaces over the field F. Let $\mathcal{B}$ be an ordered basis for V with dual basis $\mathcal{B}^*$, and let $\mathcal{B}'$ be an ordered basis for W with dual basis $\mathcal{B}'^*$. Let T be a linear transformation from V into W; let A be the matrix of T relative to $\mathcal{B}$, $\mathcal{B}'$ and let B be the matrix of T^t relative to $\mathcal{B}'^*$, $\mathcal{B}^*$. Then $B_{ij} = A_{ji}$.*

Proof. Let

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}, \qquad \mathcal{B}' = \{\beta_1, \ldots, \beta_m\},$$
$$\mathcal{B}^* = \{f_1, \ldots, f_n\}, \qquad \mathcal{B}'^* = \{g_1, \ldots, g_m\}.$$

By definition,

$$T\alpha_j = \sum_{i=1}^{m} A_{ij}\beta_i, \qquad j = 1, \ldots, n$$

$$T^t g_j = \sum_{i=1}^{n} B_{ij} f_i, \qquad j = 1, \ldots, m.$$

On the other hand,

$$\begin{aligned}(T^t g_j)(\alpha_i) &= g_j(T\alpha_i)\\ &= g_j\left(\sum_{k=1}^{m} A_{ki}\beta_k\right)\\ &= \sum_{k=1}^{m} A_{ki} g_j(\beta_k)\\ &= \sum_{k=1}^{m} A_{ki}\delta_{jk}\\ &= A_{ji}.\end{aligned}$$

For any linear functional f on V

$$f = \sum_{i=1}^{m} f(\alpha_i) f_i.$$

If we apply this formula to the functional $f = T^t g_j$ and use the fact that $(T^t g_j)(\alpha_i) = A_{ji}$, we have

$$T^t g_j = \sum_{i=1}^{n} A_{ji} f_i$$

from which it immediately follows that $B_{ij} = A_{ji}$. ▌

Definition. *If* A *is an* m × n *matrix over the field* F, *the* **transpose** *of* A *is the* n × m *matrix* A^t *defined by* $A^t_{ij} = A_{ji}$.

Theorem 23 thus states that if T is a linear transformation from V into W, the matrix of which in some pair of bases is A, then the transpose transformation T^t is represented in the dual pair of bases by the transpose matrix A^t.

Theorem 24. *Let* A *be any* m × n *matrix over the field* F. *Then the row rank of* A *is equal to the column rank of* A.

Proof. Let $\mathcal{B}$ be the standard ordered basis for F^n and $\mathcal{B}'$ the standard ordered basis for F^m. Let T be the linear transformation from F^n into F^m such that the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$ is A, i.e.,

$$T(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$$

where

$$y_i = \sum_{j=1}^{n} A_{ij}x_j.$$

The column rank of A is the rank of the transformation T, because the range of T consists of all m-tuples which are linear combinations of the column vectors of A.

Relative to the dual bases $\mathcal{B}'^*$ and $\mathcal{B}^*$, the transpose mapping T^t is represented by the matrix A^t. Since the columns of A^t are the rows of A, we see by the same reasoning that the row rank of A (the column rank of A^t) is equal to the rank of T^t. By Theorem 22, T and T^t have the same rank, and hence the row rank of A is equal to the column rank of A. ∎

Now we see that if A is an $m \times n$ matrix over F and T is the linear transformation from F^n into F^m defined above, then

$$\text{rank } (T) = \text{row rank } (A) = \text{column rank } (A)$$

and we shall call this number simply the **rank** of A.

EXAMPLE 25. This example will be of a general nature—more discussion than example. Let V be an n-dimensional vector space over the field F, and let T be a linear operator on V. Suppose $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V. The matrix of T in the ordered basis $\mathcal{B}$ is defined to be the $n \times n$ matrix A such that

$$T\alpha_j = \sum_{j=1}^{n} A_{ij}\alpha_i$$

in other words, A_{ij} is the ith coordinate of the vector $T\alpha_j$ in the ordered basis $\mathcal{B}$. If $\{f_1, \ldots, f_n\}$ is the dual basis of $\mathcal{B}$, this can be stated simply

$$A_{ij} = f_i(T\alpha_j).$$

Let us see what happens when we change basis. Suppose

$$\mathcal{B}' = \{\alpha'_1, \ldots, \alpha'_n\}$$

is another ordered basis for V, with dual basis $\{f'_1, \ldots, f'_n\}$. If B is the matrix of T in the ordered basis $\mathcal{B}'$, then

$$B_{ij} = f'_i(T\alpha'_j).$$

Let U be the invertible linear operator such that $U\alpha_j = \alpha'_j$. Then the transpose of U is given by $U^t f'_i = f_i$. It is easy to verify that since U is invertible, so is U^t and $(U^t)^{-1} = (U^{-1})^t$. Thus $f'_i = (U^{-1})^t f_i$, $i = 1, \ldots, n$. Therefore,

$$\begin{aligned} B_{ij} &= [(U^{-1})^t f_i](T\alpha'_j) \\ &= f_i(U^{-1}T\alpha'_j) \\ &= f_i(U^{-1}TU\alpha_j). \end{aligned}$$

Now what does this say? Well, $f_i(U^{-1}TU\alpha_j)$ is the i, j entry of the matrix of $U^{-1}TU$ in the ordered basis $\mathcal{B}$. Our computation above shows that this scalar is also the i, j entry of the matrix of T in the ordered basis $\mathcal{B}'$. In other words

$$\begin{aligned} [T]_{\mathcal{B}'} &= [U^{-1}TU]_{\mathcal{B}} \\ &= [U^{-1}]_{\mathcal{B}}[T]_{\mathcal{B}}[U]_{\mathcal{B}} \\ &= [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}} \end{aligned}$$

and this is precisely the change-of-basis formula which we derived earlier.

Exercises

1. Let F be a field and let f be the linear functional on F^2 defined by $f(x_1, x_2) = ax_1 + bx_2$. For each of the following linear operators T, let $g = T^t f$, and find $g(x_1, x_2)$.

(a) $T(x_1, x_2) = (x_1, 0)$;

(b) $T(x_1, x_2) = (-x_2, x_1)$;

(c) $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$.

2. Let V be the vector space of all polynomial functions over the field of real numbers. Let a and b be fixed real numbers and let f be the linear functional on V defined by

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

If D is the differentiation operator on V, what is $D^t f$?

3. Let V be the space of all $n \times n$ matrices over a field F and let B be a fixed $n \times n$ matrix. If T is the linear operator on V defined by $T(A) = AB - BA$, and if f is the trace function, what is $T^t f$?

4. Let V be a finite-dimensional vector space over the field F and let T be a linear operator on V. Let c be a scalar and suppose there is a non-zero vector α in V such that $T\alpha = c\alpha$. Prove that there is a non-zero linear functional f on V such that $T^t f = cf$.

5. Let A be an $m \times n$ matrix with *real* entries. Prove that $A = 0$ if and only if trace $(A^tA) = 0$.

6. Let n be a positive integer and let V be the space of all polynomial functions over the field of real numbers which have degree at most n, i.e., functions of the form

$$f(x) = c_0 + c_1x + \cdots + c_nx^n.$$

Let D be the differentiation operator on V. Find a basis for the null space of the transpose operator D^t.

7. Let V be a finite-dimensional vector space over the field F. Show that $T \rightarrow T^t$ is an isomorphism of $L(V, V)$ onto $L(V^*, V^*)$.

8. Let V be the vector space of $n \times n$ matrices over the field F.

(a) If B is a fixed $n \times n$ matrix, define a function f_B on V by $f_B(A) =$ trace (B^tA). Show that f_B is a linear functional on V.

(b) Show that every linear functional on V is of the above form, i.e., is f_B for some B.

(c) Show that $B \rightarrow f_B$ is an isomorphism of V onto V^*.