MULTIVARIATE STATISTICS
A Vector Space Approach

MORRIS L. EATON

MULTIVARIATE STATISTICS
A Vector Space Approach

MORRIS L. EATON
Department of Theoretical Statistics
University of Minnesota, Minneapolis

JOHN WILEY & SONS
New York Chichester Brisbane Toronto Singapore

Copyright © 1983 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data:

Eaton, Morris L.
Multivariate statistics.

(Wiley series in probability and mathematical statistics. Probability and mathematical statistics, ISSN 0271-6232)
Includes bibliographical references and index.
1. Multivariate analysis. 2. Vector spaces.
I. Title. II. Series.
QA278.E373 1983 519.5'35 83-1215
ISBN 0-471-02776-6

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

The purpose of this book is to present a version of multivariate statistical theory in which vector space and invariance methods replace, to a large extent, more traditional multivariate methods. The book is a text. Over the past ten years, various versions have been used for graduate multivariate courses at the University of Chicago, the University of Copenhagen, and the University of Minnesota. Designed for a one year lecture course or for independent study, the book contains a full complement of problems and problem solutions.

My interest in using vector space methods in multivariate analysis was aroused by William Kruskal's success with such methods in univariate linear model theory. In the late 1960s, I had the privilege of teaching from Kruskal's lecture notes where a coordinate free (vector space) approach to univariate analysis of variance was developed. (Unfortunately, Kruskal's notes have not been published.) This approach provided an elegant unification of linear model theory together with many useful geometric insights. In addition, I found the pedagogical advantages of the approach far outweighed the extra effort needed to develop the vector space machinery. Extending the vector space approach to multivariate situations became a goal, which is realized here. Basic material on vector spaces, random vectors, the normal distribution, and linear models take up most of the first half of this book.

Invariance (group theoretic) arguments have long been an important research tool in multivariate analysis as well as in other areas of statistics. In fact, invariance considerations shed light on most multivariate hypothesis testing, estimation, and distribution theory problems. When coupled with vector space methods, invariance provides an important complement to the traditional distribution theory–likelihood approach to multivariate analysis. Applications of invariance to multivariate problems occur throughout the second half of this book.

A brief summary of the contents and flavor of the ten chapters herein follows. In Chapter 1, the elements of vector space theory are presented. Since my approach to the subject is geometric rather than algebraic, there is an emphasis on inner product spaces where the notions of length, angle, and orthogonal projection make sense. Geometric topics of particular importance in multivariate analysis include singular value decompositions and angles between subspaces. Random vectors taking values in inner product spaces is the general topic of Chapter 2. Here, induced distributions, means, covariances, and independence are introduced in the inner product space setting. These results are then used to establish many traditional properties of the multivariate normal distribution in Chapter 3. In Chapter 4, a theory of linear models is given that applies directly to multivariate problems. This development, suggested by Kruskal's treatment of univariate linear models, contains results that identify all the linear models to which the Gauss–Markov Theorem applies.

Chapter 5 contains some standard matrix factorizations and some elementary Jacobians that are used in later chapters. In Chapter 6, the theory of invariant integrals (measures) is outlined. The many examples here were chosen to illustrate the theory and prepare the reader for the statistical applications to follow. A host of statistical applications of invariance, ranging from the invariance of likelihood methods to the use of invariance in deriving distributions and establishing independence, are given in Chapter 7. Invariance arguments are used throughout the remainder of the book.

The last three chapters are devoted to a discussion of some traditional and not so traditional problems in multivariate analysis. Here, I have stressed the connections between classical likelihood methods, linear model considerations, and invariance arguments. In Chapter 8, the Wishart distribution is defined via its representation in terms of normal random vectors. This representation, rather than the form of the Wishart density, is used to derive properties of the Wishart distribution. Chapter 9 begins with a thorough discussion of the multivariate analysis of variance (MANOVA) model. Variations on the MANOVA model including multivariate linear models with structured covariances are the main topic of the rest of Chapter 9. An invariance argument that leads to the relationship between canonical correlations and angles between subspaces is the lead topic in Chapter 10. After a discussion of some distribution theory, the chapter closes with the connection between testing for independence and testing in multivariate regression models.

Throughout the book, I have assumed that the reader is familiar with the basic ideas of matrix and vector algebra in coordinate spaces and has some knowledge of measure and integration theory. As for statistical prerequisites, a solid first year graduate course in mathematical statistics should suffice. The book is probably best read and used as it was written—from

此为试读，需要完整PDF请访问：www.ertongbook.com

front to back. However, I have taught short (one quarter) courses on topics in MANOVA using the material in Chapters 1, 2, 3, 4, 8, and 9 as a basis.

It is very difficult to compare this text with others on multivariate analysis. Although there may be a moderate amount of overlap with other texts, the approach here is sufficiently different to make a direct comparison inappropriate. Upon reflection, my attraction to vector space and invariance methods was, I think, motivated by a desire for a more complete understanding of multivariate statistical models and techniques. Over the years, I have found vector space ideas and invariance arguments have served me well in this regard. There are many multivariate topics not even mentioned here. These include discrimination and classification, factor analysis, Bayesian multivariate analysis, asymptotic results and decision theory results. Discussions of these topics can be found in one or more of the books listed in the Bibliography.

As multivariate analysis is a relatively old subject within statistics, a bibliography of the subject is very large. For example, the entries in *A Bibliography of Multivariate Analysis* by T. W. Anderson, S. Das Gupta, and G. H. P. Styan, published in 1972, number over 6000. The condensed bibliography here contains a few of the important early papers plus a sample of some recent work that reflects my bias. A more balanced view of the subject as a whole can be obtained by perusing the bibliographies of the multivariate texts listed in the Bibliography.

My special thanks go to the staff of the Institute of Mathematical Statistics at the University of Copenhagen for support and encouragement. It was at their invitation that I spent the 1971–1972 academic year at the University of Copenhagen lecturing on multivariate analysis. These lectures led to *Multivariate Statistical Analysis*, which contains some of the ideas and the flavor of this book. Much of the work herein was completed during a second visit to Copenhagen in 1977–1978. Portions of the work have been supported by the National Science Foundation and the University of Minnesota. This generous support is gratefully acknowledged.

A number of people have read different versions of my manuscript and have made a host of constructive suggestions. Particular thanks go to Michael Meyer, whose good sense of pedagogy led to major revisions in a number of places. Others whose help I would like to acknowledge are Murray Clayton, Siu Chuen Ho, and Takeaki Kariya.

Most of the typing of the manuscript was done by Hanne Hansen. Her efforts are very much appreciated. For their typing of various corrections, addenda, changes, and so on, I would like to thank Melinda Hutson, Catherine Stepnes, and Victoria Wagner.

MORRIS L. EATON

Minneapolis, Minnesota
May 1983

Notation

$(V, (\cdot,\cdot))$	an inner product space, vector space V and inner product $(\cdot,\cdot)$
$\mathcal{L}(V, W)$	the vector space of linear transformations on V to W
$Gl(V)$	the group of nonsingular linear transformations on V to V
$\mathcal{O}(V)$	the orthogonal group of the inner product space $(V, (\cdot,\cdot))$
R^n	Euclidean coordinate space of all n-dimensional column vectors
$\mathcal{L}_{p,n}$	the linear space of all $n \times p$ real matrices
Gl_n	the group of $n \times n$ nonsingular matrices
$\mathcal{O}_n$	the group of $n \times n$ orthogonal matrices
$\mathcal{F}_{p,n}$	the space of $n \times p$ real matrices whose p columns form an orthonormal set in R^n
G_T^+	the group of lower triangular matrices with positive diagonal elements—dimension implied by context
G_U^+	the group of upper triangular matrices with positive diagonal elements—dimension implied by context
$\mathcal{S}_p^+$	the set of $p \times p$ real symmetric positive definite matrices
$A > 0$	the matrix or linear transformation A is positive definite
$A \geqslant 0$	A is positive semidefinite (non-negative definite)
det	determinant
tr	trace
$x \square y$	the outer product of the vectors x and y
$A \otimes B$	the Kronecker product of the linear transformations A and B
Δ_r	the right-hand modulus of a locally compact topological group
$\mathcal{L}(\cdot)$	the distributional law of "$\cdot$"
$N(\mu, \Sigma)$	the normal distribution with mean μ and covariance Σ on an inner product space
$W(\Sigma, p, n)$	the Wishart distribution with n degrees of freedom and $p \times p$ parameter matrix Σ

Contents

CHAPTER 1

Vector Space Theory

In order to understand the structure and geometry of multivariate distributions and associated statistical problems, it is essential that we be able to distinguish those aspects of multivariate distributions that can be described without reference to a coordinate system and those that cannot. Finite dimensional vector space theory provides us with a framework in which it becomes relatively easy to distinguish between coordinate free and coordinate concepts. It is fair to say that the material presented in this chapter furnishes the language we use in the rest of this book to describe many of the geometric (coordinate free) and coordinate properties of multivariate probability models. The treatment of vector spaces here is far from complete, but those aspects of the theory that arise in later chapters are covered. Halmos (1958) has been followed quite closely in the first two sections of this chapter, and because of space limitations, proofs sometimes read "see Halmos (1958)."

The material in this chapter runs from the elementary notions of basis, dimension, linear transformation, and matrix to inner product space, orthogonal projection, and the spectral theorem for self-adjoint linear transformations. In particular, the linear space of linear transformations is studied in detail, and the chapter ends with a discussion of what is commonly known as the singular value decomposition theorem. Most of the vector spaces here are finite dimensional real vector spaces, although excursions into infinite dimensions occur via applications of the Cauchy–Schwarz Inequality. As might be expected, we introduce complex coordinate spaces in the discussion of determinants and eigenvalues.

Multilinear algebra and tensors are not covered systematically, although the outer product of vectors and the Kronecker product of linear transformations are covered. It was felt that the simplifications and generality obtained by introducing tensors were not worth the price in terms of added notation, vocabulary, and abstractness.

此为试读，需要完整PDF请访问：www.ertongbook.com

1.1. VECTOR SPACES

Let R denote the set of real numbers. Elements of R, called scalars, are denoted by $\alpha, \beta, \ldots$.

Definition 1.1. A set V, whose elements are called vectors, is called a real vector space if:

(I) to each pair of vectors $x, y \in V$, there is a vector $x + y \in V$, called the sum of x and y, and for all vectors in V,

(i) $x + y = y + x$.
(ii) $(x + y) + z = x + (y + z)$.
(iii) There exists a unique vector $0 \in V$ such that $x + 0 = x$ for all x.
(iv) For each $x \in V$, there is a unique vector $-x$ such that $x + (-x) = 0$.

(II) For each $\alpha \in R$ and $x \in V$, there is a vector denoted by $\alpha x \in V$, called the product of α and x, and for all scalars and vectors,

(i) $\alpha(\beta x) = (\alpha\beta)x$.
(ii) $1x = x$.
(iii) $(\alpha + \beta)x = \alpha x + \beta x$.
(iv) $\alpha(x + y) = \alpha x + \alpha y$.

In II(iii), $(\alpha + \beta)x$ means the sum of the two scalars, α and β, times x, while $\alpha x + \beta x$ means the sum of the two vectors, αx and βx. This multiple use of the plus sign should not cause any confusion. The reason for calling V a real vector space is that multiplication of vectors by real numbers is permitted.

A classical example of a real vector space is the set R^n of all ordered n-tuples of real numbers. An element of R^n, say x, is represented as

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad x_i \in R, \quad i = 1, \ldots, n,$$

and x_i is called the ith coordinate of x. The vector $x + y$ has ith coordinate $x_i + y_i$ and αx, $\alpha \in R$, is the vector with coordinates αx_i, $i = 1, \ldots, n$. With

$0 \in R^n$ representing the vector of all zeroes, it is routine to check that R^n is a real vector space. Vectors in the coordinate space R^n are always represented by a column of n real numbers as indicated above. For typographical convenience, a vector is often written as a row and appears as $x' = (x_1, \ldots, x_n)$. The prime denotes the *transpose* of the vector $x \in R^n$.

The following example provides a method of constructing real vector spaces and yields the space R^n as a special case.

◆ **Example 1.1.** Let $\mathfrak{X}$ be a set. The set V is the collection of all the real-valued functions defined on $\mathfrak{X}$. For any two elements $x_1, x_2 \in V$, define $x_1 + x_2$ as the function on $\mathfrak{X}$ whose value at t is $x_1(t) + x_2(t)$. Also, if $\alpha \in R$ and $x \in V$, αx is the function on $\mathfrak{X}$ given by $(\alpha x)(t) \equiv \alpha x(t)$. The symbol $0 \in V$ is the zero function. It is easy to verify that V is a real vector space with these definitions of addition and scalar multiplication. When $\mathfrak{X} = \{1, 2, \ldots, n\}$, then V is just the real vector space R^n and $x \in R^n$ has as its ith coordinate the value of x at $i \in \mathfrak{X}$. Every vector space discussed in the sequel is either V (for some set $\mathfrak{X}$) or a linear subspace (to be defined in a moment) of some V. ◆

Before defining the dimension of a vector space, we need to discuss linear dependence and independence. The treatment here follows Halmos (1958, Sections 5–9). Let V be a real vector space.

Definition 1.2. A finite set of vectors $\{x_i | i = 1, \ldots, k\}$ is *linearly dependent* if there exist real numbers $\alpha_1, \ldots, \alpha_k$, not all zero, such that $\Sigma \alpha_i x_i = 0$. Otherwise, $\{x_i | i = 1, \ldots, k\}$ is *linearly independent*.

A brief word about summation notation. Ordinarily, we do not indicate indices of summation on a summation sign when the range of summation is clear from the context. For example, in Definition 1.2, the index i was specified to range between 1 and k before the summation on i appeared; hence, no range was indicated on the summation sign.

An arbitrary subset $S \subseteq V$ is linearly independent if every finite subset of S is linearly independent. Otherwise, S is linearly dependent.

Definition 1.3. A *basis* for a vector space V is a linearly independent set S such that every vector in V is a linear combination of elements of S. V is *finite dimensional* if it has a finite set S that is a basis.

◆ **Example 1.2.** Take $V = R^n$ and let $\varepsilon_i' = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the one occurs as the ith coordinate of ε_i, $i = 1, \ldots, n$. For $x \in R^n$, it is clear that $x = \Sigma x_i \varepsilon_i$ where x_i is the ith coordinate of x. Thus every vector in R^n is a linear combination of $\varepsilon_1, \ldots, \varepsilon_n$. To show that $\{\varepsilon_i | i = 1, \ldots, n\}$ is a linearly independent set, suppose $\Sigma \alpha_i \varepsilon_i = 0$ for some scalars α_i, $i = 1, \ldots, n$. Then $x = \Sigma \alpha_i \varepsilon_i = 0$ has α_i as its ith coordinate, so $\alpha_i = 0$, $i = 1, \ldots, n$. Thus $\{\varepsilon_i | i = 1, \ldots, n\}$ is a basis for R^n and R^n is finite dimensional. The basis $\{\varepsilon_i | i = 1, \ldots, n\}$ is called the *standard basis* for R^n. ◆

Let V be a finite dimensional real vector space. The basic properties of linearly independent sets and bases are:

(i) If $\{x_1, \ldots, x_m\}$ is a linearly independent set in V, then there exist vectors $x_{m+1}, \ldots, x_{m+k}$ such that $\{x_1, \ldots, x_{m+k}\}$ is a basis for V.
(ii) All bases for V have the same number of elements. The *dimension* of V is defined to be the number of elements in any basis.
(iii) Every set of $n + 1$ vectors in an n-dimensional vector space is linearly dependent.

Proofs of the above assertions can be found in Halmos (1958, Sections 5–8). The dimension of a finite dimensional vector space is denoted by $\dim(V)$. If $\{x_1, \ldots, x_n\}$ is a basis for V, then every $x \in V$ is a unique linear combination of $\{x_1, \ldots, x_n\}$—say $x = \Sigma \alpha_i x_i$. That every x can be so expressed follows from the definition of a basis and the uniqueness follows from the linear independence of $\{x_1, \ldots, x_n\}$. The numbers $\alpha_1, \ldots, \alpha_n$ are called the *coordinates* of x in the basis $\{x_1, \ldots, x_n\}$. Clearly, the coordinates of x depend on the order in which we write the basis. Thus by a basis we always mean an ordered basis.

We now introduce the notion of a subspace of a vector space.

Definition 1.4. A nonempty subset $M \subseteq V$ is a *subspace* (or *linear manifold*) of V if, for each $x, y \in M$ and $\alpha, \beta \in R$, $\alpha x + \beta y \in M$.

A subspace M of a real vector space V is easily shown to satisfy the vector space axioms (with addition and scalar multiplication inherited from V), so subspaces are real vector spaces. It is not difficult to verify the following assertions (Halmos, 1958, Sections 10–12):

(i) The intersection of subspaces is a subspace.
(ii) If M is a subspace of a finite dimensional vector space V, then $\dim(M) \leqslant \dim(V)$.

(iii) Given an m-dimensional subspace M of an n-dimensional vector space V, there is a basis $\{x_1,\dots,x_m,\dots,x_n\}$ for V such that $\{x_1,\dots,x_m\}$ is a basis for M.

Given any set $S \subseteq V$, span(S) is defined to be the intersection of all the subspaces that contain S—that is, span(S) is the smallest subspace that contains S. It is routine to show that span(S) is equal to the set of all linear combinations of elements of S. The subspace span(S) is often called the subspace spanned by the set S.

If M and N are subspaces of V, then span($M \cup N$) is the set of all vectors of the form $x + y$ where $x \in M$ and $y \in N$. The suggestive notation $M + N \equiv \{z | z = x + y, x \in M, y \in N\}$ is used for span($M \cup N$) when M and N are subspaces. Using the fact that a linearly independent set can be extended to a basis in a finite dimensional vector space, we have the following. Let V be finite dimensional and suppose M and N are subspaces of V.

(i) Let $m = \dim(M)$, $n = \dim(N)$, and $k = \dim(M \cap N)$. Then there exist vectors $x_1,\dots,x_k$, $y_{k+1},\dots,y_m$, and $z_{k+1},\dots,z_n$ such that $\{x_1,\dots,x_k\}$ is a basis for $M \cap N$, $\{x_1,\dots,x_k, y_{k+1},\dots,y_m\}$ is a basis for M, $\{x_1,\dots,x_k, z_{k+1},\dots,z_n\}$ is a basis for N, and $\{x_1,\dots,x_k, y_{k+1},\dots,y_m, z_{k+1},\dots,z_n\}$ is a basis for $M + N$. If $k = 0$, then $\{x_1,\dots,x_k\}$ is interpreted as the empty set.

(ii) $\dim(M + N) = \dim(M) + \dim(N) - \dim(M \cap N)$.

(iii) There exists a subspace $M_1 \subseteq V$ such that $M \cap M_1 = \{0\}$ and $M + M_1 = V$.

Definition 1.5. If M and N are subspaces of V that satisfy $M \cap N = \{0\}$ and $M + N = V$, then M and N are *complementary* subspaces.

The technique of decomposing a vector space into two (or more) complementary subspaces arises again and again in the sequel. The basic property of such a decomposition is given in the following proposition.

Proposition 1.1. Suppose M and N are complementary subspaces in V. Then each $x \in V$ has a unique representation $x = y + z$ with $y \in M$ and $z \in N$.

Proof. Since $M + N = V$, each $x \in V$ can be written $x = y_1 + z_1$ with $y_1 \in M$ and $z_1 \in N$. If $x = y_2 + z_2$ with $y_2 \in M$ and $z_2 \in N$, then $0 = x -$

$x = (y_1 - y_2) + (z_1 - z_2)$. Hence $(y_2 - y_1) = (z_1 - z_2)$ so $(y_2 - y_1) \in M \cap N = \{0\}$. Thus $y_1 = y_2$. Similarly, $z_1 = z_2$. □

The above proposition shows that we can decompose the vector space V into two vector spaces M and N and each x in V has a unique piece in M and in N. Thus x can be represented as (y, z) with $y \in M$ and $z \in N$. Also, note that if $x_1, x_2 \in V$ and have the representations (y_1, z_1), (y_2, z_2), then $\alpha x_1 + \beta x_2$ has the representation $(\alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2)$, for $\alpha, \beta \in R$. In other words the function that maps x into its decomposition (y, z) is linear. To make this a bit more precise, we now define the direct sum of two vector spaces.

Definition 1.6. Let V_1 and V_2 be two real vector spaces. The *direct sum* of V_1 and V_2, denoted by $V_1 \oplus V_2$, is the set of all ordered pairs $\{x, y\}$, $x \in V_1$, $y \in V_2$, with the linear operations defined by $\alpha_1\{x_1, y_1\} + \alpha_2\{x_2, y_2\} \equiv \{\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2\}$.

That $V_1 \oplus V_2$ is a real vector space with the above operations can easily be verified. Further, identifying V_1 with $\{\{x_1, 0\}|x \in V_1\} \equiv \tilde{V}_1$ and V_2 with $\{\{0, y\}|y \in V_2\} \equiv \tilde{V}_2$, we can think of V_1 and V_2 as complementary subspaces of $V_1 \oplus V_2$, since $\tilde{V}_1 + \tilde{V}_2 = V_1 \oplus V_2$ and $\tilde{V}_1 \cap \tilde{V}_2 = \{0, 0\}$, which is the zero element in $V_1 \oplus V_2$. The relation of the direct sum to our previous decomposition of a vector space should be clear.

◆ **Example 1.3.** Consider $V = R^n$, $n \geqslant 2$, and let p and q be positive integers such that $p + q = n$. Then R^p and R^q are both real vector spaces. Each element of R^n is a n-tuple of real numbers, and we can construct subspaces of R^n by setting some of these coordinates equal to zero. For example, consider $M = \{x \in R^n | x = \begin{pmatrix} y \\ 0 \end{pmatrix}$ with $y \in R^p, 0 \in R^q\}$ and $N = \{x \in R^n | x = \begin{pmatrix} 0 \\ z \end{pmatrix}$ with $0 \in R^p$ and $z \in R^q\}$. It is clear that $\dim(M) = p$, $\dim(N) = q$, $M \cap N = \{0\}$, and $M + N = R^n$. The identification of R^p with M and R^q with N shows that it is reasonable to write $R^p \oplus R^q = R^{p+q}$. ◆

1.2. LINEAR TRANSFORMATIONS

Linear transformations occupy a central position, both in vector space theory and in multivariate analysis. In this section, we discuss the basic properties of linear transforms, leaving the deeper results for consideration after the introduction of inner products. Let V and W be real vector spaces.

Definition 1.7. Any function A defined on V and taking values in W is called a *linear transformation* if $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$ for all $x_1, x_2 \in V$ and $\alpha_1, \alpha_2 \in R$.

Frequently, $A(x)$ is written Ax when there is no danger of confusion. Let $\mathcal{L}(V, W)$ be the set of all linear transformations on V to W. For two linear transformations A_1 and A_2 in $\mathcal{L}(V, W)$, $A_1 + A_2$ is defined by $(A_1 + A_2)(x) = A_1 x + A_2 x$ and $(\alpha A)(x) = \alpha A x$ for $\alpha \in R$. The zero linear transformation is denoted by 0. It should be clear that $\mathcal{L}(V, W)$ is a real vector space with these definitions of addition and scalar multiplication.

◆ **Example 1.4.** Suppose $\dim(V) = m$ and let $x_1, \ldots, x_m$ be a basis for V. Also, let $y_1, \ldots, y_m$ be arbitrary vectors in W. The claim is that there is a unique linear transformation A such that $Ax_i = y_i$, $i = 1, \ldots, m$. To see this, consider $x \in V$ and express x as a unique linear combination of the basis vectors, $x = \Sigma \alpha_i x_i$. Define A by

$$Ax = \sum_{i}^{n} \alpha_i A x_i = \sum_{1}^{n} \alpha_i y_i .$$

The linearity of A is easy to check. To show that A is unique, let B be another linear transformation with $Bx_i = y_i$, $i = 1, \ldots, n$. Then $(A - B)(x_i) = 0$ for $i = 1, \ldots, n$, and $(A - B)(x) = (A - B)(\Sigma \alpha_i x_i) = \Sigma \alpha_i (A - B)(x_i) = 0$ for all $x \in V$. Thus $A = B$. ◆

The above example illustrates a general principle—namely, a linear transformation is completely determined by its values on a basis. This principle is used often to construct linear transformations with specified properties. A modification of the construction in Example 1.4 yields a basis for $\mathcal{L}(V, W)$ when V and W are both finite dimensional. This basis is given in the proof of the following proposition.

Proposition 1.2. If $\dim(V) = m$ and $\dim(W) = n$, then $\dim(\mathcal{L}(V, W)) = mn$.

Proof. Let $x_1, \ldots, x_m$ be a basis for V and let $y_1, \ldots, y_n$ be a basis for W. Define a linear transformation A_{ji}, $i = 1, \ldots, m$ and $j = 1, \ldots, n$, by

$$A_{ji}(x_k) = \begin{cases} 0 & \text{if } k \neq i \\ y_j & \text{if } k = i \end{cases}.$$

此为试读,需要完整PDF请访问: www.ertongbook.com