Matrix Theory and Linear Algebra

I. N. HERSTEIN

DAVID J. WINTER

Matrix Theory and Linear Algebra

I. N. HERSTEIN
University of Chicago

DAVID J. WINTER
University of Michigan, Ann Arbor

Macmillan Publishing Company
NEW YORK
Collier Macmillan Publishers
LONDON

PRINTED IN THE UNITED STATES OF AMERICA

Macmillan Publishing Company
866 Third Avenue, New York, New York 10022

Collier Macmillan Canada, Inc.

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Herstein, I. N.
Matrix theory and linear algebra.

Includes index.
1. Matrices. 2. Algebras, Linear. I. Winter, David J. II. Title.
QA188.H47 1988 512.9'434 87-11669
ISBN 0-02-353951-8

Printing: 1 2 3 4 5 6 7 8

Year: 8 9 0 1 2 3 4 5 6 7

To the memory of A. A. Albert

I. N. H.

To Alison

D. J. W.

Preface

Matrix theory and linear algebra is a subject whose material can, and is, taught at a variety of levels of sophistication. These levels depend on the particular needs of students learning the subject matter. For many students it suffices to know matrix theory as a computational tool. For them, a course that stresses manipulation with matrices, a few of the basic theorems, and some applications fits the bill. For others, especially students in engineering, chemistry, physics, and economics, quite a bit more is required. Not only must they acquire a solid control of calculating with matrices, but they also need some of the deeper results, many of the techniques, and a sufficient knowledge and familiarity with the theoretical aspects of the subject. This will allow them to adapt arguments or extend results to the particular problems they are considering. Finally, there are the students of mathematics. Regardless of what branch of mathematics they go into, a thorough mastery not only of matrix theory, but also of linear spaces and transformations, is a must.

We have endeavored to write a book that can be of use to all these groups of students, so that by picking and choosing the material, one can make up a course that would satisfy the needs of each of these groups.

During the writing, we were confronted by a dilemma. On the one hand, we wanted to keep our presentation as simple as possible, and on the other, we wanted to cover the subject fully and without compromise, even when the going got tough. To solve this dilemma, we decided to prepare two versions of the book, this version for the more experienced or ambitious student, and another, *A Primer on Linear Algebra*, for students who desire to get a first look at linear algebra but do not want to go into it in as great a depth the first time around. These two versions are almost identical in many respects, but there are some important differences. Whereas in *A Primer on Linear Algebra* we excluded some advanced topics and simplified some others, in this version we go into more depth for some topics treated in both books (determinants, Markov processes, incidence models, differential equations, least squares methods) and included some others (triangulation of matrices with real entries, the Jordan canonical form). Also, there are some harder exercises, and the grading of the exercises presupposes a somewhat more experienced student.

Toward the end of the preface we lay out some possible programs of study at these various levels. At the same time, the other material is there for them to look through or study if they desire.

Our approach is to start slowly, setting out at the level of 2×2 matrices. These matrices have the great advantage that everything about them is open to the eye—that students can get their hands on the material, experiment with it, see what any specific theorem says about them. Furthermore, all this can be done by performing some simple calculations.

However, in treating the 2×2 matrices, we try to handle them as we would the general $n \times n$ case. In this microcosm of the larger matrix world, virtually every concept that will arise for $n \times n$ matrices, general vector spaces, and linear transformations makes its appearance. This appearance is usually in a form ready for extension to the general situation. Probably the only exception to this is the theory of determinants, for which the 2×2 case is far too simplistic.

With the background acquired in playing with the 2×2 matrices in this general manner, the results for the most general case, as they unfold, are not as surprising, mystifying, or mysterious to the students as they might otherwise be. After all, these results are almost old friends whose acquaintance we made in our earlier 2×2 incarnation. So this simplified context serves both as a laboratory and as a motivation for what is to come.

From the fairly concrete world of the $n \times n$ matrices we pass to the more abstract realm of vector spaces and linear transformations. Here the basic strategy is to prove that an n-dimensional vector space is isomorphic to the space of n-tuples. With this isomorphism established, the whole corpus of concepts and results that we had obtained in the context of n-tuples and $n \times n$ matrices is readily transferred to the setting of arbitrary vector spaces and linear transformations. Moreover, this transfer is accomplished with little or no need of proof. Because of the nature of isomorphism, it is enough merely to cite the proof or result obtained earlier for n-tuples or $n \times n$ matrices.

The vector spaces we treat in the book are only over the fields of real or complex numbers. While a little is lost in imposing this restriction, much is gained. For instance, our vector spaces can always be endowed with an inner product. Using this inner product, we can always decompose the space as a direct sum of any subspace and its orthogonal complement. This direct sum decomposition is then exploited to the hilt to obtain very simple and illuminating proofs of many of the theorems.

There is an attempt made to give some nice, characteristic applications of the material of the book. Some of these can be integrated into a course almost from the beginning, where we discuss 2×2 matrices.

Least squares methods are discussed to show how linear equations having no solutions can always be solved approximately in a very neat and efficient way. These methods are then used to show how to find functions that approximate given data.

Finally, in the last chapter, we discuss how to translate some of our methods into *linear algorithms*, that is, finite-numerical step-by-step versions of methods of linear algebra. The emphasis is on linear algorithms that can be used in writing computer programs for finding exact and approximate solutions of linear equations. We then illustrate how some of these algorithms are used in such a computer program, written in the programming language Pascal.

There are many exercises in the book. These are usually divided into categories entitled *numerical*, *more theoretical*: *easier*, *middle-level*, and *harder*. One even runs across some problems that are downright hard, which we put in the subcategory *very hard*. It goes without saying that the problems are an intrinsic part of any course. They are the best means for checking on one's understanding and mastery of the material. Included are exercises that are treated later in the text itself or that can be solved easily using later results. This gives you, the reader, a chance to try your own hand at developing important tools, and to compare your approach in an early context to our approach in the later context. An answer manual is available from the publisher.

We mentioned earlier that the book can serve as a textbook for several different levels. Of course, how this is done is up to the individual instructor. We present below some possible sample courses.

1. *One-term course emphasizing computational aspects.*

   Chapters 1 through 4, Chapters 5 and 6 with an emphasis on methods and a minimum of proofs. One might merely do determinants in the 3×3 case with a statement that the $n \times n$ case follows similar rules. Sections marked "optional" should be skipped. The sections in Chapter 11 entitled "Fibonacci Numbers" and "Equations of Curves" could be integrated into the course. Problems should primarily be all the numerical ones and a sampling of the easier and middle-level theoretical problems.

2. *One-term course for users of matrix theory and linear algebra in allied fields.*

   Chapters 1 through 6, with a possible deemphasis on the proofs of the properties of determinants and with an emphasis on computing with determinants. Some introduction to vector spaces and linear transformations would be desirable. Chapters 12 and 13, which deal with least squares methods and computing, could play an important role in the course. Each of the applications in Chapter 11 could be touched on. As for problems, again all the numerical ones, most of the middle-level ones, and a few of the harder ones should be appropriate for such a course.

3. *One-term course for mathematics majors.*

   Most of Chapter 1 done very quickly, with much left for the students to read on their own. All of Chapters 2 through 7, including some of the optional topics. Definitely some emphasis should be given to abstract vector spaces and linear transformations, as in Chapters 8 and 9, possibly skipping quotient spaces and invariant subspaces. The whole gamut of problems should be assignable to the students.

4. *Two-term course for users of matrix theory and linear algebra.*

   The entire book, but going easy on proofs for determinants, on the material on abstract vector spaces and linear transformations, totally omitting Chapter 10 on the Jordan canonical form and the discussion of differential equations in Chapter 11, plus a fairly thorough treatment of Chapters 12 and 13. The problems can be chosen from all parts of the problem sections.

5. *Two-term course for mathematics majors.*
The entire book, with perhaps less emphasis on Chapters 12 and 13.

We should like to thank the many people who have looked at the manuscript, commented on it, and made useful suggestions. We want to thank Bill Blair and Lynne Small for their extensive and valuable analysis of the book at its different stages, which had a very substantial effect on its final form. We should also like to thank Gary Ostedt, Bob Clark, and Elaine Wetterau of the Macmillan Publishing Company for their help in bringing this book into being. We should like to thank Lee Zukowski for the excellent typing job he did on the manuscript. And we should like to thank Pedro Sanchez for his valuable help with the computer program and the last chapter.

I. N. H.
D. J. W.

List of Symbols

$s \in S$	the element s is in the set S
$S \subset T$	the set S is contained in the set T
$S \cup T$	the union of the sets S and T
$S_1 \cup \cdots \cup S_n$	the union of the sets $S_1, \ldots, S_n$
$S \cap T$	the intersection of the sets S and T
$S_1 \cap \cdots \cap S_n$	the intersection of the sets $S_1, \ldots, S_n$
$\mathbb{R}$	the set of all real numbers, 2
$M_2(\mathbb{R})$	the set of all 2×2 matrices over $\mathbb{R}$, 2, 90
$A + B$	the sum of matrices A and B, 3, 91
$A - B$	the difference of matrices A and B, 3, 91
0	the zero matrix, 3, 91
$-A$	the negative of the matrix A, 4, 91
uA	the matrix obtained by multiplying the matrix A by the scalar u, 4, 92
AB	the product of matrices A and B, 4, 93
I	the identity matrix, 5, 92
A^{-1}	the inverse of the matrix A, when it exists, 5, 92
aI	the scalar matrix corresponding to the scalar a, 6, 93
$\sum_{r=1}^{n} a_r$	the summation of a_r from 1 to n, 10
$\sum_{r=1}^{m} \sum_{s=1}^{n} a_{rs}$	the double summation of a_{rs}, 10
A^m	the product of the matrix A with itself m times, 12, 97
A^0	the identity matrix, 12, 97
A^{-m}	the product of the matrix A^{-1} with itself m times, when it exists, 12, 97
$\operatorname{tr}(A)$	the trace of the matrix A, 16, 109
A'	the transpose of the matrix A, 18, 116

$\det(A)$, $\|A\|$	the determinant of the matrix A, 21, 215
fg	the product or composite of functions f, g, 30
f^{-1}	the inverse of a $1-1$ onto function f, 30
f^n	product of the function f with itself n times, 32
$\mathbb{R}^{(2)}$	the set of all vectors in the Cartesian plane, 35, 98
$v + w$	the sum of vectors in $\mathbb{R}^{(2)}$, 35
dv	the vector v in $\mathbb{R}^{(2)}$ multiplied by the scalar d, 35
Av	the vector obtained by applying $A \in M_2(\mathbb{R})$ to $v \in \mathbb{R}^{(2)}$, 36
$P_A(x)$	the characteristic polynomial of the matrix A, 39, 266
$\alpha = a + bi$	the complex number, α, with real part a and pure imaginery part bi, 46
$\mathbb{C}$	the set of all complex numbers, 46
$\bar{\alpha} = a - bi$	the conjugate of the complex number $\alpha = a + bi$, 49
$\|\alpha\|$	the absolute value of the complex number α, 49
$M_2(\mathbb{C})$	the set of all 2×2 matrices over $\mathbb{C}$, 52, 90
(v, w)	the inner product of column vectors v, w, 55, 124
A^*	the Hermitian adjoint of A, 57, 116
$\\|v\\|$	the length of vector v, 58, 127
Add $(r, s; u)$	the operation of adding u times row s to row r, 8, 257
Interchange (r, s)	the operation of interchanging rows r and s, 80, 257
Multiply $(r; u)$	the operation of multiplying row r by the scalar u, 80, 257
F	the set of all scalars: $F = \mathbb{R}$ or $F = \mathbb{C}$, 91
E_{rs}	the matrix whose (r, s) entry is 1 and all of whose other entries are 0, 94
$M_n(F)$	the set of all $n \times n$ matrices over F, 90
$F^{(n)}$	the set of all column vectors with n coordinates from F, 98
$v + w$	the sum of vectors in $F^{(n)}$, 99
0	the zero vector in $F^{(n)}$, 99
$-v$	the negative of the vector v in $F^{(n)}$, 99
tv	the vector v multiplied by the scalar t, 99
Av	the vector obtained by applying the matrix or transformation $A \in M_n(F)$ to the vector $v \in F^{(n)}$, 101
$v^{\perp}$	the set of all vectors orthogonal to v, 125
e_s	the vector whose r entry is 0 if $r \neq s$ and 1 if $r = s$, 131
$A \sim B$	the matrices A and B are similar, 154
$\mathrm{cl}(A)$	the set of all matrices B such that $A \sim B$, 155
$\dim(V)$	the dimension of V, 167, 310
$V + W$	the sum of subspaces, V, W, 170
$V \oplus W$	the direct sum of subspaces V, W, 172, 225
$W^{\perp}$	the subspace of vectors orthogonal to W, 173, 325

$n(A)$	the nullity of the matrix A, 181
$\langle v_1, \ldots, v_r \rangle$	the subspace spanned by the vectors $v_1, \ldots, v_r$, 134
$r(A)$	the rank of the matrix A, 181
$q_A(x)$	the minimum polynomial of the matrix A, 184
V_a	the set of characteristic vectors of A associated with a, 194
M_{rs}	the (r, s) minor of the matrix A, 215
A_{rs}	the (r, s) cofactor of the matrix A, 246
$A^{@}$	the classical adjoint of the matrix A, 253
$A(r, s; q)$	the elementary matrix corresponding to Add $(r, s; q)$, 257
$M(r; q)$	the elementary matrix corresponding to Multiply $(r; q)$, 257
$I(r, s)$	the elementary matrix corresponding to Interchange (r, s), 257
$v + w$	the sum of vectors v, w in a vector space V, 293
0	the zero vector in a vector space V, 293
av	the vector v multiplied by the scalar a in a vector space V, 186
$\operatorname{Ker} \phi$	the kernel of the homomorphism ϕ from V to W, 305
$V \cong W$	there exists an isomorphism from V to W, 305
(v, w)	the inner product of elements v, w in an inner product space, 314
$\|v\|$	the length of a vector in an inner product space, 314
$T_1 + T_2$	the sum of linear transformations T_1, T_2, 334
aT	the linear transformation obtained by multiplying the linear transformation T by the scalar a, 335
0	the zero linear transformation, 335
$-T$	the negative of the linear transformation T, 335
$L(V)$	the set of all linear transformations of V over F, 335
$T_1 T_2$	the product of linear transformation T_1, T_2, 337
I	the identity linear transformation, 338
T^{-1}	the inverse of the linear transformation T, when it exists, 339
$m(T)$	the matrix of a linear transformation in a given basis, 342
$\operatorname{tr}(T)$	the trace of the linear transformation T, 346
$\det(T)$	the determinant of the linear transformation T, 346
$P_T(x)$	the characteristic polynomial of the linear transformation T, 346
$[v, w]$	the inner product $(\phi(v), \phi(w))$ corresponding to a given isomorphism ϕ, 352
T^*	the Hermitian adjoint of the linear transformation T, 354
$v + W$	the set of all vectors $v + w$ with w in W, 357
V/W	the quotient space of V by W, 357
$\tilde{T}$	the linear transformation on a subspace W of V induced by the linear transformation T of V when W is invariant under T, 366
$\hat{T}$	the linear transformation on the quotient space V/W induced by the linear transformation T on V when W is invariant under T, 367

$V_0(T)$	the generalized nullspace of the linear transformation T of V, 380
$V_*(T)$	the intersection of $T^e(V)$ over all positive integers e, 380
$V_a(T)$	the generalized characteristic space of T at a, 381
e^A	the exponential of an $n \times n$ real or complex matrix A, 394
$x'(t)$	the derivative of the vector function $x(t)$, 400
$S^{(k)}$	the kth state in a Markov process, 416
$\mathrm{Proj}_W(y)$	the projection of the vector y on the subspace W, 440
A^-	the approximate inverse of the $m \times n$ matrix A, 444
$\langle v, w \rangle_P$	the inner product $(P(v), P(w))$ corresponding to a given invertible matrix P, 455
$\|v\|_P$	the length of v, given the inner product $(v, w)_P$, 455
$\tilde{A}$	the weighted approximate inverse of an $m \times n$ matrix A, given invertible weighting matrices $P \in M_m(\mathbb{R})$, $Q \in M_n(\mathbb{R})$, 456

MATRIX THEORY AND LINEAR ALGEBRA

Contents

3 The $n \times n$ Matrices 90

4 More on $n \times n$ Matrices 164

5 Determinants 214

6 Rectangular Matrices. More on Determinants 273