Further Mathematics for Economic Analysis

$$\pi = \pi(\mathbf{v}, p, \mathbf{q}) = pF(\mathbf{v}) - q_1 v_1 - \dots$$

Knut Sydsæter
Peter Hammond
Atle Seierstad
Arne Strøm

FT Prentice Hall
FINANCIAL TIMES

FURTHER MATHEMATICS FOR ECONOMIC ANALYSIS

KNUT SYDSÆTER
PETER HAMMOND
ATLE SEIERSTAD
ARNE STRØM

FT Prentice Hall
FINANCIAL TIMES

An imprint of Pearson Education
Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

Visit us on the World Wide Web at:
www.pearsoned.co.uk

First published 2005

ISBN 0 273 65576 0

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

10 9 8 7 6 5 4 3 2 1
09 08 07 06 05

Typeset in TEX by the authors.
Printed and bound by Ashford Colour Press Ltd., Gosport.

The publisher's policy is to use paper manufactured from sustainable forests.

PREFACE

The economic world is a misty region.
The first explorers used unaided vision.
Mathematics is the lantern by which what was before
dimly visible now looms up in firm, bold outlines.
The old phantasmagoria[1] *disappear.*
We see better. We also see further.
—Irving Fisher (1892)

This book is intended for advanced undergraduate and graduate students of economics whose mathematical requirements go beyond the material usually taught in undergraduate courses. In particular, it presents most of the mathematical tools required for typical graduate courses in economic theory—both micro and macro. The volume has many references to Sydsæter and Hammond's *Essential Mathematics for Economic Analysis* (Pearson Education, 2002) [generally referred to as EMEA throughout this successor volume], but that book is by no means a prerequisite. Indeed, this volume is designed to be accessible to anybody who has had a basic training in mathematical analysis and linear algebra at the level often encountered in courses taught to economics undergraduates. Like EMEA, the treatment here is deliberately quite rigorous, but rigour is not emphasized for its own sake.

An important aspect of the book is its systematic treatment of the calculus of variations, optimal control theory, and dynamic programming. Recent years may have seen control theory lose some of its prominence in economics, but it is still useful in several areas, notably resource economics and industrial organization. Furthermore, in our view the existing economics literature has paid too little attention to some of the subtler issues that frequently arise, especially when the time horizon is infinite.

Some early chapters review and extend elementary matrix algebra, multivariable calculus, and static optimization. Other chapters present multiple integration, as well as ordinary difference and differential equations, including systems of equations. There is a chapter on elementary topology in $\mathbb{R}^n$ and separation theorems. In the final chapter we discuss set-valued functions ("correspondences") and the fixed point theorems that economists most often use.

As the title suggests, this is a mathematics book with the material arranged to allow pro-

[1] "Phantasmagoria" is a term invented in 1802 to describe an exhibition of optical illusions produced by means of a magic lantern.

gressive learning of mathematical topics. If the student acquires some economic insight and intuition at the same time, so much the better. At times, we do emphasize economics not only to motivate a mathematical topic, but also to help acquire mathematical intuition. Obviously, our economic discussions will be more easily understood by students who already have a certain rudimentary understanding of economics, especially of what economics should be about.

In particular, this is not a book about economics or even about mathematical economics. As one reviewer of EMEA put it: "Mathematics is the star of the show". We expect students to learn economic theory systematically in other courses, based on other books or articles. We will have succeeded if they can concentrate on the economics in these courses, having mastered beforehand the relevant mathematical tools we present.

Almost every section includes worked examples and problems for students to solve as exercises. Many of the problems are quite easy in order to build the students' confidence in absorbing the material, but there are also a number of more challenging problems. Concise solutions to odd-numbered problems are suggested. Solutions to even-numbered problems will be available to instructors in a manual that can be downloaded from a restricted access part of an associated website. That website will also include other supplementary material, including exam type problems that instructors might find useful for assignments or even exams.

The book is not intended to be studied in a steady progression from beginning to end. Some of the more challenging chapters start with a simple treatment where some technical aspects are played down, while the more complete theory is discussed later. Some of the material, including more challenging proofs, is in small print. Quite often those proofs rely on technical ideas that are only expounded in the last two chapters.

The author team consists of the two co-authors of EMEA, together with two other mathematicians in the Department of Economics at the University of Oslo.

Knut Sydsæter, Peter Hammond, Atle Seierstad, and Arne Strøm

Oslo and Stanford, December 2004

CONTENTS

Supporting resource

Visit www.pearsoned.co.uk/sydsaeter to find a valuable online resource

For instructors

- Complete, downloadable Instructor's Manual with exam problems

For more information please contact your local Pearson Education sales representative or visit www.pearsoned.co.uk/sydsaeter

TOPICS IN LINEAR ALGEBRA

I came to the position that mathematical analysis is not one of many ways of doing economic theory: It is the only way. Economic theory is mathematical analysis. Everything else is just pictures and talk.
—R. E. Lucas, Jr. (2001)

This chapter covers a few topics in linear algebra that are not always treated in standard mathematics courses for economics students. We assume that the reader has previously studied some basic concepts and results, which are briefly reviewed in Section 1.1. A fuller treatment including many practice problems, can be found in EMEA, or in many alternative textbooks.

Next we consider partitioned matrices. These are useful for computations involving large matrices, especially when they have a special structure.

In an economic model described by a linear system of equations, it is important to know when that system has a solution, and when the solution is unique. General conditions for existence and uniqueness are most easily stated using the concept of linear independence, along with the related concept of the rank of a matrix. These topics are treated in Sections 1.3 and 1.4.

The implications of these ideas for solving linear systems is the topic of Section 1.5.

This chapter also discusses eigenvalues, which are indispensable in several areas of mathematics of interest to economists—in particular, stability theory for difference and differential equations. Eigenvalues and the associated eigenvectors are also important in determining when a matrix can be "diagonalized", which greatly simplifies some calculations involving the matrix. The chapter concludes by looking at quadratic forms— first without linear constraints, then with them. Such quadratic forms are especially useful in deriving and checking second-order conditions for multivariable optimization.

1.1 Review of Basic Linear Algebra

An $m \times n$ matrix is a rectangular array with m rows and n columns:

$$\mathbf{A} = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1}$$

Here a_{ij} denotes the element in the ith row and the jth column.

If $\mathbf{A} = (a_{ij})_{m\times n}$, $\mathbf{B} = (b_{ij})_{m\times n}$, and α is a scalar, we define

$$\mathbf{A}+\mathbf{B} = (a_{ij}+b_{ij})_{m\times n}, \quad \alpha\mathbf{A} = (\alpha a_{ij})_{m\times n}, \quad \mathbf{A}-\mathbf{B} = \mathbf{A}+(-1)\mathbf{B} = (a_{ij}-b_{ij})_{m\times n} \tag{2}$$

Suppose that $\mathbf{A} = (a_{ij})_{m\times n}$ and that $\mathbf{B} = (b_{ij})_{n\times p}$. Then the product $\mathbf{C} = \mathbf{AB}$ is the $m \times p$ matrix $\mathbf{C} = (c_{ij})_{m\times p}$, whose element in the ith row and the jth column is the inner product (or dot product) of the ith row of $\mathbf{A}$ and the jth column of $\mathbf{B}$. That is,

$$c_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} + \cdots + a_{in}b_{nj} \tag{3}$$

It is important to note that the product $\mathbf{AB}$ is defined only if the number of columns in $\mathbf{A}$ is equal to the number of rows in $\mathbf{B}$.

If $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are matrices whose dimensions are such that the given operations are defined, then the basic properties of matrix multiplication are:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \qquad \textbf{(associative law)} \tag{4}$$

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC} \qquad \textbf{(left distributive law)} \tag{5}$$

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC} \qquad \textbf{(right distributive law)} \tag{6}$$

If $\mathbf{A}$ and $\mathbf{B}$ are matrices, it is possible for $\mathbf{AB}$ to be defined even if $\mathbf{BA}$ is not. Moreover, even if $\mathbf{AB}$ and $\mathbf{BA}$ are both defined, $\mathbf{AB}$ is not necessarily equal to $\mathbf{BA}$. Matrix multiplication is *not* commutative. In fact,

$$\mathbf{AB} \neq \mathbf{BA}, \text{ except in special cases} \tag{7}$$

$$\mathbf{AB} = \mathbf{0} \text{ does not imply that } \mathbf{A} \text{ or } \mathbf{B} \text{ is } \mathbf{0} \tag{8}$$

$$\mathbf{AB} = \mathbf{AC} \text{ and } \mathbf{A} \neq \mathbf{0} \text{ do not imply that } \mathbf{B} = \mathbf{C} \tag{9}$$

By using matrix multiplication, one can write a general system of linear equations in a very concise way. Specifically, the system

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad \text{can be written as} \qquad \mathbf{Ax} = \mathbf{b}$$

if we define $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.

A matrix is **square** if it has an equal number of rows and columns. If **A** is a square matrix and n is a positive integer, we define the nth power of **A** in the obvious way:

$$\mathbf{A}^n = \underbrace{\mathbf{A}\mathbf{A}\cdots\mathbf{A}}_{n \text{ factors}} \tag{10}$$

For **diagonal matrices** it is particularly easy to compute powers:

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix} \implies \mathbf{D}^n = \begin{pmatrix} d_1^n & 0 & \dots & 0 \\ 0 & d_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^n \end{pmatrix} \tag{11}$$

The **identity matrix** of order n, denoted by $\mathbf{I}_n$ (or often just by **I**), is the $n \times n$ matrix having ones along the main diagonal and zeros elsewhere:

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}_{n\times n} \qquad \textbf{(identity matrix)} \tag{12}$$

If **A** is any $m \times n$ matrix, then $\mathbf{A}\mathbf{I}_n = \mathbf{A} = \mathbf{I}_m\mathbf{A}$. In particular,

$$\mathbf{A}\mathbf{I}_n = \mathbf{I}_n\mathbf{A} = \mathbf{A} \quad \text{for every } n \times n \text{ matrix } \mathbf{A} \tag{13}$$

If $\mathbf{A} = (a_{ij})_{m\times n}$ is any matrix, the **transpose** of **A** is defined as $\mathbf{A}' = (a_{ji})_{n\times m}$. The subscripts i and j are interchanged because every row of **A** becomes a column of $\mathbf{A}'$, and every column of **A** becomes a row of $\mathbf{A}'$.

The following rules apply to matrix transposition:

$$\text{(i) } (\mathbf{A}')' = \mathbf{A} \quad \text{(ii) } (\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}' \quad \text{(iii) } (\alpha\mathbf{A})' = \alpha\mathbf{A}' \quad \text{(iv) } (\mathbf{A}\mathbf{B})' = \mathbf{B}'\mathbf{A}' \tag{14}$$

A square matrix is called **symmetric** if $\mathbf{A} = \mathbf{A}'$.

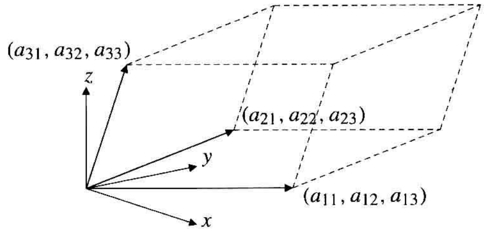
Determinants and Matrix Inverses

Recall that the determinants $|\mathbf{A}|$ of 2×2 and 3×3 matrices are defined by

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{cases} a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ \qquad - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{cases}$$

Determinants of order 2 and 3 have a geometric interpretation which is shown and explained in Fig. 1 for the case $n = 3$.



$$\begin{array}{l}\text{the volume of the} \\ \text{“box” spanned by} \\ \text{the three vectors}\end{array} = \pm \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Figure 1

For a general $n \times n$ matrix $\mathbf{A} = \{a_{ij}\}$, the determinant $|\mathbf{A}|$ can be defined recursively. In fact,

$$|\mathbf{A}| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{ij}A_{ij} + \cdots + a_{in}A_{in} \tag{15}$$

where the *cofactors* A_{ij} are determinants of $(n-1) \times (n-1)$ matrices given by

$$A_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & \dots & a_{i,j-1} & \boxed{a_{ij}} & a_{i,j+1} & \dots & a_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix} \tag{16}$$

Here lines have been drawn through row i and column j, which are to be deleted from the matrix $\mathbf{A}$ to produce A_{ij}. Formula (15) gives the *cofactor expansion of* $|\mathbf{A}|$ *along the* i*th row*.

In general,

$$\begin{aligned} a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} &= |\mathbf{A}| \\ a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} &= 0 \qquad (k \neq i) \end{aligned} \tag{17}$$

$$\begin{aligned} a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} &= |\mathbf{A}| \\ a_{1j}A_{1k} + a_{2j}A_{2k} + \cdots + a_{nj}A_{nk} &= 0 \qquad (k \neq j) \end{aligned} \tag{18}$$

This result says that an expansion of a determinant along row i in terms of the cofactors of row k vanishes when $k \neq i$, and is equal to $|\mathbf{A}|$ if $k = i$. Likewise, an expansion along column j in terms of the cofactors of column k vanishes when $k \neq j$, and is equal to $|\mathbf{A}|$ if $k = j$.

The following rules for manipulating determinants are often useful:

If two rows (or two columns) of $\mathbf{A}$ are interchanged, the determinant changes sign but its absolute value remains unchanged. (19)

If all the elements in a single row (or column) of $\mathbf{A}$ are multiplied by a number c, the determinant is multiplied by c. (20)

If two of the rows (or columns) of $\mathbf{A}$ are proportional, then $|\mathbf{A}| = 0$. (21)

The value of $|\mathbf{A}|$ remains unchanged if a multiple of one row (or one column) is added to another row (or column). (22)

Furthermore,

$$|\mathbf{A}'| = |\mathbf{A}|, \quad \text{where } \mathbf{A}' \text{ is the transpose of } \mathbf{A} \tag{23}$$

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}| \tag{24}$$

$$|\mathbf{A}+\mathbf{B}| \neq |\mathbf{A}| + |\mathbf{B}| \quad \text{(usually)} \tag{25}$$

The *inverse* $\mathbf{A}^{-1}$ of an $n \times n$ matrix $\mathbf{A}$ has the following properties:

$$\mathbf{B} = \mathbf{A}^{-1} \iff \mathbf{AB} = \mathbf{I}_n \iff \mathbf{BA} = \mathbf{I}_n \tag{26}$$

$$\mathbf{A}^{-1} \text{ exists} \iff |\mathbf{A}| \neq 0 \tag{27}$$

If $\mathbf{A} = (a_{ij})_{n\times n}$ and $|\mathbf{A}| \neq 0$, the unique inverse of $\mathbf{A}$ is given by

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\operatorname{adj}(\mathbf{A}), \quad \text{where} \quad \operatorname{adj}(\mathbf{A}) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{28}$$

with A_{ij}, the *cofactor* of the element a_{ij}, given by (16). Note carefully the order of the indices in the *adjoint matrix*, adj(**A**) with the column number preceding the row number. The matrix $(A_{ij})_{n\times n}$ is called the *cofactor matrix*, whose transpose is the adjoint matrix.

In particular, for 2×2 matrices,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{if} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0 \tag{29}$$

The following are important rules for inverses (when the relevant inverses exist):

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad (\mathbf{A}')^{-1} = (\mathbf{A}^{-1})', \quad (c\mathbf{A})^{-1} = c^{-1}\mathbf{A}^{-1} \tag{30}$$

Cramer's Rule

A linear system of n equations and n unknowns,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{31}$$

has a unique solution if and only if $|\mathbf{A}| = |(a_{ij})_{n\times n}| \neq 0$. The solution is then

$$x_j = \frac{|\mathbf{A}_j|}{|\mathbf{A}|}, \quad j = 1, 2, \ldots, n \tag{32}$$

where the determinant

$$|\mathbf{A}_j| = \begin{vmatrix} a_{11} & \ldots & a_{1,j-1} & b_1 & a_{1,j+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,j-1} & b_2 & a_{2,j+1} & \ldots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,j-1} & b_n & a_{n,j+1} & \ldots & a_{nn} \end{vmatrix} \tag{33}$$

is obtained by replacing the jth column of $|\mathbf{A}|$ by the column whose components are $b_1, b_2, \ldots, b_n$.

If the right-hand side of the equation system (31) consists only of zeros, so that it can be written in matrix form as $\mathbf{Ax} = \mathbf{0}$, the system is called **homogeneous**. A homogeneous system will always have the **trivial solution** $x_1 = x_2 = \cdots = x_n = 0$. The following result is useful.

$$\mathbf{Ax} = \mathbf{0} \quad \text{has nontrivial solutions} \quad \Longleftrightarrow \quad |\mathbf{A}| = 0 \tag{34}$$

Vectors

Recall that an ***n*-vector** is an ordered n-tuple of numbers. It is often convenient to regard the rows and columns of a matrix as vectors, and an ***n*-vector** can be understood either as a $1 \times n$ matrix $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ (a *row vector*) or as an $n \times 1$ matrix $\mathbf{a}' = (a_1, a_2, \ldots, a_n)'$ (a *column vector*). The operations of addition, subtraction and multiplication by scalars of vectors are defined in the obvious way. The **dot product** (or **inner product**) of the n-vectors $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ is defined as

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i \tag{35}$$

If $\mathbf{a} = (a_1, \ldots, a_n)'$ and $\mathbf{b} = (b_1, \ldots, b_n)'$ both happen to be $n \times 1$ matrices, $\mathbf{a} \cdot \mathbf{b}$ is again given by (35). Then the transpose $\mathbf{a}'$ of $\mathbf{a}$ is a $1 \times n$ matrix, and the matrix product $\mathbf{a}'\mathbf{b}$ is a 1×1 matrix. In fact, $\mathbf{a}'\mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \mathbf{a} \cdot \mathbf{b}$.

Important properties of the dot product include these: If $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are n-vectors and α is a scalar, then

$$\text{(i) } \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}, \quad \text{(ii) } \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}, \quad \text{(iii) } (\alpha \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha \mathbf{b}) = \alpha(\mathbf{a} \cdot \mathbf{b}) \tag{36}$$

The **Euclidean norm** or **length** of the vector $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \tag{37}$$

Note that $\|\alpha \mathbf{a}\| = |\alpha| \|\mathbf{a}\|$ for all scalars and vectors.

The following useful inequalities hold:

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \cdot \|\mathbf{b}\| \qquad \textbf{(Cauchy–Schwarz inequality)} \tag{38}$$

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\| \qquad \textbf{(triangle inequality for norms)} \tag{39}$$

The **angle** θ between nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$ is defined by

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} \qquad (\theta \in [0, \pi]) \tag{40}$$

This definition makes sense because the Cauchy–Schwarz inequality implies that the right-hand side has absolute value ≤ 1. According to (40), $\cos\theta = 0$ iff $\mathbf{a} \cdot \mathbf{b} = 0$. Then $\theta = \pi/2 = 90°$.

By definition, **a** and **b** in $\mathbb{R}^n$ are **orthogonal** if their dot product is 0. In symbols:

$$\mathbf{a} \perp \mathbf{b} \iff \mathbf{a} \cdot \mathbf{b} = 0 \tag{41}$$

The **straight line** through two distinct points $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ in $\mathbb{R}^n$ is the set of all $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ such that

$$\mathbf{x} = t\mathbf{a} + (1 - t)\mathbf{b} \tag{42}$$

for some real number t.

The **hyperplane** in $\mathbb{R}^n$ that passes through the point $\mathbf{a} = (a_1, \ldots, a_n)$ and is orthogonal to the nonzero vector $\mathbf{p} = (p_1, \ldots, p_n)$, is the set of all points $\mathbf{x} = (x_1, \ldots, x_n)$ such that

$$\mathbf{p} \cdot (\mathbf{x} - \mathbf{a}) = 0 \tag{43}$$

1.2 Partitioned Matrices and Their Inverses

Many applications of linear algebra deal with matrices of high order. To see the structure of such matrices and to ease the computational burden in dealing with them, it is often helpful to partition the matrices into suitably chosen submatrices. The operation of subdividing a matrix into submatrices is called **partitioning**.

EXAMPLE 1 Consider the 3×5 matrix $\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 & 0 & 4 \\ 1 & 2 & 1 & 3 & 4 \\ 0 & 0 & 2 & 1 & 4 \end{pmatrix}$. The matrix **A** can be partitioned in a number of ways. For example,

$$\mathbf{A} = \left(\begin{array}{cc:ccc} 2 & 0 & 1 & 0 & 4 \\ 1 & 2 & 1 & 3 & 4 \\ \hdashline 0 & 0 & 2 & 1 & 4 \end{array}\right) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \tag{$*$}$$

where $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$, and $\mathbf{A}_{22}$ are submatrices of dimensions 2×2, 2×3, 1×2, and 1×3, respectively. This is useful because A_{21} is a zero matrix. Other less useful partitionings of **A** include the one where **A** is partitioned into three row vectors, and the one where **A** is partitioned into five column vectors.

Though Example 1 raises the possibility of partitioning a matrix into arbitrarily many submatrices, the rest of this section considers only partitionings into 2×2 arrays of submatrices as in $(*)$.

Operations on Partitioned Matrices

One can perform standard matrix operations on partitioned matrices, treating the submatrices as if they were ordinary matrix elements. This requires obeying the rules for sums, differences, and products.

Adding or subtracting partitioned matrices is simple. For example,

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} + \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \mathbf{A}_{12} + \mathbf{B}_{12} \\ \mathbf{A}_{21} + \mathbf{B}_{21} & \mathbf{A}_{22} + \mathbf{B}_{22} \end{pmatrix} \tag{1}$$

as long as the dimensions of $\mathbf{A}_{11}$ are those of $\mathbf{B}_{11}$, the dimensions of $\mathbf{A}_{12}$ are those of $\mathbf{B}_{12}$, and so on. The result follows directly from the definition of matrix addition. The rule for subtracting partitioned matrices is similar.

The rule for multiplying a partitioned matrix by a number is obvious. For example,

$$\alpha \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{A}_{11} & \alpha\mathbf{A}_{12} \\ \alpha\mathbf{A}_{21} & \alpha\mathbf{A}_{22} \end{pmatrix} \tag{2}$$

The following example shows how to multiply partitioned matrices.

EXAMPLE 2 Let $\mathbf{A}$ be the 3×5 matrix $(*)$ in Example 1, and let $\mathbf{B}$ be the 5×4 matrix

$$\mathbf{B} = \left(\begin{array}{cc|cc} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 5 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$$

with the indicated partitioning. The product $\mathbf{AB}$ is defined, and the ordinary rules of matrix multiplication applied to the entire matrices yield

$$\mathbf{AB} = \begin{pmatrix} 2 & 0 & 5 & 6 \\ 1 & 2 & 6 & 15 \\ 0 & 0 & 3 & 4 \end{pmatrix}$$

Consider next how to take advantage of the partitioning of the two matrices to compute the product $\mathbf{AB}$. Simply multiply the partitioned matrices as if the submatrices were ordinary matrix elements to obtain

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 4 & 2 \\ 2 & 11 \end{pmatrix} + \begin{pmatrix} 1 & 4 \\ 4 & 4 \end{pmatrix} \\ (0 \quad 0) + (0 \quad 0) & (0 \quad 0) + (3 \quad 4) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 5 & 6 \\ 1 & 2 & 6 & 15 \\ 0 & 0 & 3 & 4 \end{pmatrix}$$

Note that the two matrices $\mathbf{A}$ and $\mathbf{B}$ were partitioned with dimensions chosen so that all the needed products of submatrices are well defined.

The method suggested by Example 2 is valid in general. It is not difficult to formulate and prove the general result, though the notation becomes cumbersome. If you work through Problem 1 in detail, the general idea should become clear enough.

Multiplying matrices using partitioning is particularly convenient if the matrices have a special structure and involve simple submatrices (like identity or zero matrices).

EXAMPLE 3 Consider the problem of computing powers of the following matrix with the indicated partitioning:

$$\mathbf{M} = \left(\begin{array}{cc:cc} 1/3 & 1/2 & 1/6 & 0 \\ 1/2 & 1/3 & 0 & 1/6 \\ \hdashline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

Then

$$\mathbf{M}^2 = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{P}^2 & (\mathbf{P}+\mathbf{I})\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \quad \text{and} \quad \mathbf{M}^3 = \begin{pmatrix} \mathbf{P}^3 & (\mathbf{P}^2+\mathbf{P}+\mathbf{I})\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

In general, for all natural numbers n, it can be shown by induction that

$$\mathbf{M}^n = \begin{pmatrix} \mathbf{P}^n & (\mathbf{P}^{n-1}+\cdots+\mathbf{P}^2+\mathbf{P}+\mathbf{I})\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

Inverses by Partitioning

Inverting large square matrices is often made much easier using partitioning. Consider an $n \times n$ matrix $\mathbf{A}$ which has an inverse. Assume that $\mathbf{A}$ is partitioned as follows:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \qquad \text{where } \mathbf{A}_{11} \text{ is a } k \times k \text{ matrix with an inverse} \tag{3}$$

Hence $\mathbf{A}_{12}$ is a $k \times (n-k)$ matrix, $\mathbf{A}_{21}$ is $(n-k) \times k$, while $\mathbf{A}_{22}$ is an $(n-k) \times (n-k)$ matrix. Since $\mathbf{A}$ has an inverse, there exists an $n \times n$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}_n$. Partitioning $\mathbf{B}$ in the same way as $\mathbf{A}$ yields

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$$

The equality $\mathbf{AB} = \mathbf{I}_n$ implies the following four matrix equations for determining $\mathbf{B}_{11}$, $\mathbf{B}_{12}$, $\mathbf{B}_{21}$, and $\mathbf{B}_{22}$:

$$\text{(i)} \quad \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} = \mathbf{I}_k \qquad \text{(ii)} \quad \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} = \mathbf{0}_{k\times(n-k)}$$

$$\text{(iii)} \quad \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} = \mathbf{0}_{(n-k)\times k} \qquad \text{(iv)} \quad \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} = \mathbf{I}_{n-k}$$

where the subscripts attached to $\mathbf{I}$ and $\mathbf{0}$ indicate the dimensions of these matrices. Because $\mathbf{A}_{11}$ has an inverse, (ii) gives $\mathbf{B}_{12} = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}$. Inserting this into (iv) gives $(-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} + \mathbf{A}_{22})\mathbf{B}_{22} = \mathbf{I}_{n-k}$, and so $\mathbf{B}_{22} = (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}$. Next, solve