A Brief Course in Linear Algebra

A BRIEF COURSE IN LINEAR ALGEBRA:

MATRICES AND MATRIX EQUATIONS FOR UNDERGRADUATE STUDENTS IN APPLIED MATHEMATICS, SCIENCE AND ENGINEERING

WHYE-TEONG ANG

BrownWalker Press
Irvine & Boca Raton

A Brief Course in Linear Algebra: Matrices and Matrix Equations for Undergraduate Students in Applied Mathematics, Science and Engineering

Copyright © 2019 Whye-Teong Ang. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.

BrownWalker Press • Universal Publishers, Inc.
Irvine, California & Boca Raton, Florida • USA
www.BrownWalkerPress.com
2019

ISBN: 978-1-62734-724-2 (pbk.)
ISBN: 978-1-62734-725-9 (ebk.)

Cover design by Ivan Popov

Publisher's Cataloging-in-Publication Data

Names: Ang, Whye-Teong, author.
Title: A brief course in linear algebra : matrices and matrix equations for undergraduate students in applied mathematics, science and engineering / Whye-Teong Ang.
Description: Irvine, CA : BrownWalker, 2019.
Identifiers: LCCN 2019939671 | ISBN 978-1-62734-724-2 (paperback) | ISBN 978-1-62734-725-9 (ebook)
Subjects: LCSH: Algebras, Linear. | Algebras, Linear--Study and teaching (Higher) | Matrices--Study and teaching (Higher) | Algebra. | Mathematics. | BISAC: MATHEMATICS / Algebra / Linear. | MATHEMATICS / Study & Teaching.
Classification: LCC QA184.2 .A54 2019 (print) | LCC QA184.2 (ebook) | DDC 512/.5--dc23.

Preface

This book contains a brief course in elementary linear algebra with an emphasis on solving systems of linear algebraic and ordinary differential equations. It is written for undergraduate students in the disciplines of applied mathematics, science and engineering. Basic knowledge of the arithmetic of complex numbers and exposure to elementary functions and calculus are assumed.

The book comprises six chapters.

Chapter 1 covers the basics of matrices and vectors, providing definitions and concepts needed in linear algebra studies in later chapters.

Chapter 2 is concerned with solving systems of linear algebraic equations. It shows how elementary row operations on an array of numbers can be used to reduce a given system of linear algebraic equations to a simpler but equivalent system that can be easily solved. The chapter also introduces the concept of linearly independent vectors and explains how the task of determining whether a given set of vectors is linearly independent or not can be formulated in terms of a homogeneous system of linear algebraic equations.

Chapter 3 looks at elementary matrices and matrix inverses. It shows how elementary row operations can be performed on an invertible square matrix to find its inverse matrix and explains how matrix invertibility is related to solving a system of linear algebraic equations. Formulae for some properties involving inverses of matrices are given in the chapter.

Chapter 4 begins with a formula defining the determinant of a square matrix, shows how elementary row operations can be performed on a square matrix to calculate its determinant, and derives alternative formulae for calculating the determinant. The relation between matrix determinant, matrix inverse and solutions of systems of linear algebraic equations is explained.

Chapter 5 deals with the matrix eigenproblem and the matrix diagonalization problem. The two related problems are of fundamental importance in linear algebra. The chapter explains how they can be applied to solve homogeneous systems of first order linear ordinary differential equations.

Chapter 6 gives a summary of the definition of terms and the main results in the earlier chapters.

The connections between the topics covered are carefully elucidated. Derivations or proofs are given for all the main results studied.

Problems are set at the end of each of the first five chapters to test the understanding of students and to provide further insights into the topics covered in the course.

Singapore W. T. ANG
10 June 2019

Contents

1 Basics of matrices and vectors **11**

1.1 Definition and notation 11

1.1.1 Null matrices 12

1.1.2 Submatrices 12

1.2 Square matrices 13

1.2.1 Upper triangular matrices 13

1.2.2 Lower triangular matrices 14

1.2.3 Diagonal matrices 15

1.2.4 Identity matrices 15

1.3 Arithmetic of matrices 15

1.3.1 Equality of matrices 15

1.3.2 Addition of matrices 16

1.3.3 Multiplication of a number to a matrix . . 17

1.3.4 Product of matrices 17

1.3.5 Powers of square matrices 21

1.4 Transpose of a matrix 23

1.5 Vectors 25

1.5.1 Ordered sets of numbers 25

1.5.2 Euclidean norm of a vector 26

1.5.3 Inner product of vectors 26

1.5.4 Orthogonal vectors 27

1.5.5 Arithmetic of vectors 27

1.5.6 Linear combination of vectors 29

1.5.7 Real vector spaces and subspaces 32

1.6 Problems 35

2 Linear algebraic equations and linearly independent vectors **39**

2.1 Linear algebraic equations 39

2.2 Solving systems of linear algebraic equations 41
2.2.1 Systems of linear algebraic equations 41
2.2.2 Solution of a system 43
2.2.3 Elimination of unknowns 44
2.2.4 Elementary row operations 47
2.2.5 General observations 58
2.3 Linearly independent vectors 61
2.3.1 Basic definition 61
2.3.2 Homogeneous systems with infinitely many solutions 67
2.4 Problems 69

3 Inverse of a square matrix 73
3.1 Elementary row operations and elementary matrices 73
3.2 Invertible square matrices 79
3.2.1 Inverse of a matrix 79
3.2.2 Inverses of elementary matrices 80
3.2.3 Finding the inverse of a matrix by row operations 83
3.2.4 Invertible matrices as products of elementary matrices 88
3.2.5 Some properties involving inverses of matrices 91
3.3 Invertible matrices and systems of linear algebraic equations 93
3.4 Elementary column row operations 94
3.5 Problems 96

4 Determinant of a square matrix 99
4.1 Starting definition for determinant 99
4.2 Determinant and elementary row and column operations 104
4.2.1 Effects of elementary row operations on determinant 104
4.2.2 Effects of elementary column operations on determinant 106
4.2.3 Evaluating determinant by row operations . 109
4.3 Some properties of determinant 115
4.3.1 Product of square matrices 115
4.3.2 Inverse of a square matrix 117
4.3.3 Transpose of a square matrix 117

4.4 Alternative equivalent definition for determinant . 119
4.5 Determinant in terms of determinants of submatrices 119
4.6 Cofactor, adjugate and inverse 125
4.7 Determinant and systems of linear algebraic equations 128
4.8 Problems 132

5 Eigenvalues and eigenvectors, diagonalization and differential equations 135
5.1 Matrix eigenproblem 135
5.2 Eigenvalues and eigenvectors 138
5.3 Linearly independent eigenvectors 146
5.4 Matrix diagonalization problem 149
5.5 Diagonalization matrices 152
5.6 Real symmetric matrices 157
5.6.1 Eigenvalues of a real symmetric matrix . . 157
5.6.2 Eigenvectors of a real symmetric matrix . . 157
5.6.3 Diagonalizing real symmetric matrices . . . 158
5.7 Solutions of homogeneous linear ordinary differential equations with constant coefficients 168
5.7.1 Systems of first order homogeneous linear ordinary differential equations with constant coefficients 168
5.7.2 Solutions of first order systems with diagonalizable coefficient matrices 169
5.7.3 Solutions of first order systems with nondiagonalizable coefficient matrices 176
5.8 Problems 185

6 Summary 189
6.1 Definitions and concepts 189
6.2 Main results 194

Chapter 1

Basics of matrices and vectors

1.1 Definition and notation

A matrix of order $M \times N$, that is, an $M \times N$ matrix, is a collection (set) of MN numbers arranged in M rows and N columns. For clarity, a pair of brackets is used to enclose the numbers in a matrix.

Examples:

1. (10) $(1 \times 1$ matrix)

2. $\begin{pmatrix} 2 & -9 \\ 0 & 1 \end{pmatrix}$ $(2 \times 2$ matrix)

3. $\begin{pmatrix} 1 & -1 & 2 \\ 2 & 5 & 2 \\ 10 & 2 & -10 \\ 1 & 1 & -9 \\ \sqrt{2} & 1 & \pi \end{pmatrix}$ $(5 \times 3$ matrix)

4. $\begin{pmatrix} 3 & 6 & 4 & 9 \\ 1/10 & 1 & 1/2 & 1 \\ 2 & 2 & 4 & 5 \end{pmatrix}$ $(3 \times 4$ matrix).

We refer to the numbers in a matrix as the elements of the matrix. The elements of a matrix may be real or complex numbers.

We use bold capital letters such as **A**, **B**, **M** and **N** to represent matrices.

Example:
If **A** and **B** are used to represent respectively the 5×3 and 3×4 matrices in the examples given above, then we write:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 5 & 2 \\ 10 & 2 & -10 \\ 1 & 1 & -9 \\ \sqrt{2} & 1 & \pi \end{pmatrix} \text{ and } \mathbf{B} = \begin{pmatrix} 3 & 6 & 4 & 9 \\ 1/10 & 1 & 1/2 & 1 \\ 2 & 2 & 4 & 5 \end{pmatrix}.$$

If we denote the element in the i-th row and the j-th column of an $M \times N$ matrix **A** by a_{ij} then we can write:

$$\mathbf{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & \cdots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \cdots & \cdots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \cdots & \cdots & a_{MN} \end{pmatrix}.$$

1.1.1 Null matrices

If all the elements in a matrix are zero, the matrix is a null (or zero) matrix. The $M \times N$ null matrix is denoted by $\mathbf{O}_{M \times N}$. If the order of a null matrix is implicitly understood in a discussion, the null matrix may be simply written as **O**.

1.1.2 Submatrices

A matrix formed by deleting away selected rows and/or columns of a larger matrix **A** is called a submatrix of **A**.

A matrix can be partitioned and expressed in terms of its submatrices.

Examples:

1. If we delete away the first row, the third row and the last two columns of the 4×4 matrix $\mathbf{A} = \begin{pmatrix} 1 & 3 & 8 & 2 \\ 2 & 3 & 6 & 4 \\ 3 & 0 & 7 & 0 \\ 1 & 0 & 5 & 9 \end{pmatrix}$, we obtain $\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$ which is a 2×2 submatrix of $\mathbf{A}$.

2. If we define $\mathbf{P} = \begin{pmatrix} 1 & 6 & 5 & 9 \\ 3 & 1 & 8 & 1 \\ 2 & 7 & 4 & 6 \end{pmatrix}$, $\mathbf{R} = \begin{pmatrix} 1 & 6 \\ 3 & 1 \\ 2 & 7 \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} 5 & 9 \\ 8 & 1 \\ 4 & 6 \end{pmatrix}$, we can express the 3×4 matrix $\mathbf{P}$ in terms of its 3×2 submatrices $\mathbf{R}$ and $\mathbf{S}$ as $\mathbf{P} = \begin{pmatrix} \mathbf{R} & \mathbf{S} \end{pmatrix}$. The 3×4 matrix $\begin{pmatrix} \mathbf{S} & \mathbf{R} \end{pmatrix}$, the 6×2 matrix $\begin{pmatrix} \mathbf{R} \\ \mathbf{S} \end{pmatrix}$ and the 6×4 matrix $\begin{pmatrix} \mathbf{R} & \mathbf{O}_{3\times 2} \\ \mathbf{O}_{3\times 2} & \mathbf{S} \end{pmatrix}$ are given by

$$\begin{pmatrix} 5 & 9 & 1 & 6 \\ 8 & 1 & 3 & 1 \\ 4 & 6 & 2 & 7 \end{pmatrix}, \begin{pmatrix} 1 & 6 \\ 3 & 1 \\ 2 & 7 \\ 5 & 9 \\ 8 & 1 \\ 4 & 6 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 6 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 2 & 7 & 0 & 0 \\ 0 & 0 & 5 & 9 \\ 0 & 0 & 8 & 1 \\ 0 & 0 & 4 & 6 \end{pmatrix}$$

respectively.

1.2 Square matrices

A matrix is square if its number of rows equals its number of columns. The elements a_{11}, a_{22}, $\cdots$, $a_{(N-1)(N-1)}$ and a_{NN} in an $N \times N$ square matrix $\mathbf{A}$ are called the diagonal elements of $\mathbf{A}$.

1.2.1 Upper triangular matrices

A square matrix $\mathbf{A} = (a_{ij})$ is an upper triangular matrix if $a_{ij} = 0$ for all $i > j$, that is, if all the elements below the diagonal elements are zero.

Examples:
The matrices listed below are upper triangular.

1. $\begin{pmatrix} 2 & 9 \\ 0 & 6 \end{pmatrix}$ (diagonal elements are 2 and 6)

2. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 5 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ (diagonal elements are 1, 5 and 3)

3. $\begin{pmatrix} 1 & 5 & 3 & 2 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ (diagonal elements are 1, 3, 0 and 4)

1.2.2 Lower triangular matrices

A square matrix $\mathbf{A} = (a_{ij})$ is a lower triangular matrix if $a_{ij} = 0$ for all $i < j$, that is, if all the elements above the diagonal elements are zero.

Examples:
The matrices listed below are lower triangular.

1. $\begin{pmatrix} 1 & 0 \\ 5 & 2 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 & 0 \\ 8 & 5 & 0 \\ 6 & 1 & 0 \end{pmatrix}$

3. $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 1 & 8 & 0 \\ 1 & 0 & 9 & 4 \end{pmatrix}$

1.2.3 Diagonal matrices

If all the nondiagonal elements of a square matrix $\mathbf{A}$ are zero, then $\mathbf{A}$ is a diagonal matrix. A diagonal matrix is both lower and upper triangular.

Examples:

1. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ (3×3 diagonal matrix)

2. $\begin{pmatrix} -10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$ (5×5 diagonal matrix)

1.2.4 Identity matrices

A diagonal matrix is said to be an identity matrix if all the diagonal elements of the matrix is 1. The $N \times N$ identity matrix is denoted by $\mathbf{I}_{N\times N}$ (which may be simply written as $\mathbf{I}$ if the order of the matrix is understood).

Example:

$$\mathbf{I}_{5\times 5} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

1.3 Arithmetic of matrices

1.3.1 Equality of matrices

Two matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$, both of order $M \times N$, are said to be equal, that is, $\mathbf{A} = \mathbf{B}$, if $a_{ij} = b_{ij}$, that is, if the element in the i-th row and j-th column of $\mathbf{A}$ is equal to the element in the i-th row and j-th column of $\mathbf{B}$.

Example:

If the matrices $\begin{pmatrix} a+b & c \\ c & a-b \end{pmatrix}$ and $\begin{pmatrix} c & 3 \\ d & b \end{pmatrix}$ are equal to each other, find a, b, c and d.

Solution:
Since

$$\begin{pmatrix} a+b & c \\ c & a-b \end{pmatrix} = \begin{pmatrix} c & 3 \\ d & b \end{pmatrix},$$

we can equate the corresponding elements in the two matrices to obtain $a+b=c$, $c=3$, $c=d$ and $a-b=b$. Thus, $a=2$, $b=1$, $c=3$ and $d=3$.

1.3.2 Addition of matrices

Two matrices $\mathbf{A}=(a_{ij})$ and $\mathbf{B}=(b_{ij})$, both of order $M\times N$, can be added up to form another $M\times N$ matrix denoted by $\mathbf{A}+\mathbf{B}$. If $\mathbf{S}=(s_{ij})=\mathbf{A}+\mathbf{B}$ then $s_{ij}=a_{ij}+b_{ij}$.

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} -1 & 2 \\ 8 & 3 \\ 0 & 6 \end{pmatrix}$$
$$= \begin{pmatrix} 1-1 & 2+2 \\ 3+8 & 4+3 \\ 5+0 & 6+6 \end{pmatrix} = \begin{pmatrix} 0 & 4 \\ 11 & 7 \\ 5 & 12 \end{pmatrix}$$

It is obvious that $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$, that is, addition of matrices is commutative, and $\mathbf{A}+\mathbf{O}=\mathbf{O}+\mathbf{A}=\mathbf{A}$, where $\mathbf{O}$ is the null matrix of the same order as $\mathbf{A}$.

If $\mathbf{C}=(c_{ij})$ is another matrix of order $M\times N$ then

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}),$$

since $(a_{ij}+b_{ij})+c_{ij}=a_{ij}+(b_{ij}+c_{ij})$. Note that the usual rule of evaluating first the mathematical expression enclosed within a pair of brackets is also applicable in matrix operations.

1.3.3 Multiplication of a number to a matrix

If $\mathbf{A} = (a_{ij})$ is an $M \times N$ matrix and c is a number then $c\mathbf{A}$ is an $M \times N$ matrix defined by $c\mathbf{A} = (b_{ij})$, where $b_{ij} = ca_{ij}$. We write $(-1)\mathbf{A}$ as simply $-\mathbf{A}$ and define $\mathbf{B} - \mathbf{A} = \mathbf{B} + (-\mathbf{A})$ (if $\mathbf{A}$ and $\mathbf{B}$ are both $M \times N$).

Examples:

1. $2\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$

2. $(-1)\begin{pmatrix} 3 \\ 2 \end{pmatrix} = -\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \end{pmatrix}$

3. $\begin{pmatrix} 8 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$

It is obvious that $a\mathbf{A} + b\mathbf{A} = (a+b)\mathbf{A}$ for any numbers a and b and any matrix $\mathbf{A}$. Also, $0\mathbf{A} = \mathbf{O}$ and $\mathbf{A} - \mathbf{A} = \mathbf{O}$.

1.3.4 Product of matrices

Let $\mathbf{A} = (a_{ij})$ be an $M \times N$ matrix and $\mathbf{B} = (b_{ij})$ an $R \times Q$ matrix. We can form the product $\mathbf{AB}$, which is an $M \times Q$ matrix, only if $N = R$. If $N = R$ and $\mathbf{AB} = \mathbf{M} = (m_{ij})$ then the element in the k-th row and p-th column of $\mathbf{M}$, that is, m_{kp}, is defined in terms of the elements in the k-th row of $\mathbf{A}$ and the p-th column of $\mathbf{B}$ by

$$\begin{aligned} m_{kp} &= a_{k1}b_{1p} + a_{k2}b_{2p} + \cdots + a_{kN}b_{Np} \\ &= \sum_{n=1}^{N} a_{kn}b_{np}, \end{aligned}$$

that is, m_{kp} is the sum of the products of all the N corresponding pairs of elements in the submatrices $\begin{pmatrix} a_{k1} & a_{k2} & \cdots & a_{kN} \end{pmatrix}$ (the k-th row of $\mathbf{A}$) and $\begin{pmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{Np} \end{pmatrix}$ (the p-th column of $\mathbf{B}$). Corresponding pairs of elements are formed by taking the element in the n-th column of $\begin{pmatrix} a_{k1} & a_{k2} & \cdots & a_{kN} \end{pmatrix}$ and that in the n-th row of $\begin{pmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{Np} \end{pmatrix}$. For example, a_{k3} and b_{3p} form a corresponding pair of elements.

For given integer values K and P, where $1 \leq K \leq M$ and $1 \leq P \leq Q$, if we regard (m_{KP}) as a 1×1 submatrix of $\mathbf{AB}$, we may write:

$$\begin{aligned} (m_{KP}) &= \begin{pmatrix} a_{K1} & a_{K2} & \cdots & a_{KN} \end{pmatrix} \begin{pmatrix} b_{1P} \\ b_{2P} \\ \vdots \\ b_{NP} \end{pmatrix} \\ &= (a_{K1}b_{1P} + a_{K2}b_{2P} + \cdots + a_{KN}b_{NP}). \end{aligned}$$

From the above definition of the product of matrices, it is clear that if g is a number then $g(\mathbf{AB}) = (g\mathbf{A})\mathbf{B} = \mathbf{A}(g\mathbf{B})$. Hence the products $g(\mathbf{AB})$, $(g\mathbf{A})\mathbf{B}$ and $\mathbf{A}(g\mathbf{B})$ are the same matrix and may be simply written as $g\mathbf{AB}$.

Examples:

1. Since $\mathbf{R} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} 5 & 1 & 2 & 2 \\ 3 & 3 & 1 & 2 \end{pmatrix}$ are of order 3×2 and 2×4 respectively, we can form the product $\mathbf{RS}$ which is a 3×4 matrix. The element in the k-th row and the p-th column of $\mathbf{RS}$ can be calculated by using the elements in the k-th row of $\mathbf{R}$ and the p-th column of $\mathbf{S}$. For example, the element in the second row and third column of $\mathbf{RS}$ is given by the element in the 1×1 submatrix $\begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = (6 + 4) = (10)$. If we calculate all the elements of of $\mathbf{RS}$, we obtain

$$\mathbf{RS} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 & 1 & 2 & 2 \\ 3 & 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 11 & 7 & 4 & 6 \\ 27 & 15 & 10 & 14 \\ 43 & 23 & 16 & 22 \end{pmatrix}.$$

In this example, we cannot form the product **SR** (that is, **SR** does not exist), since the number of columns of **S** is not equal to the number of rows of **R**.

2. If $\mathbf{P} = \begin{pmatrix} 1 & 3 \\ 6 & 4 \\ 5 & 6 \end{pmatrix}$ and $\mathbf{Q} = \begin{pmatrix} 7 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$, both **PQ** and **QP** exist and can be worked out as follows:

$$\begin{aligned} \mathbf{PQ} &= \begin{pmatrix} 1 & 3 \\ 6 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 7 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 16 & 7 & 5 \\ 54 & 14 & 16 \\ 53 & 17 & 16 \end{pmatrix}, \\ \mathbf{QP} &= \begin{pmatrix} 7 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 6 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 23 & 37 \\ 20 & 23 \end{pmatrix}. \end{aligned}$$

The products **PQ** and **QP** cannot be equal to each other, as the order of **PQ** is different from that of **QP**.

3. If $\mathbf{R} = \begin{pmatrix} 1 & 3 \\ 6 & 4 \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} 7 & 1 \\ 3 & 2 \end{pmatrix}$, we find that

$$\begin{aligned} \mathbf{RS} &= \begin{pmatrix} 1 & 3 \\ 6 & 4 \end{pmatrix} \begin{pmatrix} 7 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 16 & 7 \\ 54 & 14 \end{pmatrix}, \\ \mathbf{SR} &= \begin{pmatrix} 7 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 6 & 4 \end{pmatrix} = \begin{pmatrix} 13 & 25 \\ 15 & 17 \end{pmatrix} \neq \mathbf{RS}. \end{aligned}$$

For general matrices **A** and **B** such that **AB** and **BA** exist, **AB** needs not necessarily be equal to **BA**, even if **AB** and **BA** have the same order. For some matrices, we may possibly find that **AB** = **BA**. In general, multiplication of matrices is, however, not commutative.

4. If $\mathbf{U} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$ and $\mathbf{V} = \begin{pmatrix} 10 & 1 \end{pmatrix}$ then

$$\begin{aligned} \mathbf{I}_{2\times 2}\mathbf{U} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix} = \mathbf{U}, \\ \mathbf{V}\mathbf{I}_{2\times 2} &= \begin{pmatrix} 10 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 1 \end{pmatrix} = \mathbf{V}. \end{aligned}$$

If the elements in the identity matrix $\mathbf{I}_{N\times N}$ are denoted by δ_{ij} then

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

If $\mathbf{A} = (a_{ij})$ is an $M \times N$ matrix and $\mathbf{AI}_{N\times N} = \mathbf{C} = (c_{ij})$ then

$$c_{kp} = a_{k1}\delta_{1p} + a_{k2}\delta_{2p} + \cdots + a_{kN}\delta_{Np}.$$

Depending on the value of p, all the terms, except one, in the sum $a_{k1}\delta_{1p} + a_{k2}\delta_{2p} + \cdots + a_{kN}\delta_{Np}$ are zero. If $p = 1$, then $c_{k1} = a_{k1}$. Similarly, if $p = 2$, then $c_{k2} = a_{k2}$. In general, we find that $c_{kp} = a_{kp}$, that is, $\mathbf{C} = \mathbf{A}$. Thus, if the product $\mathbf{AI}$ exists, then $\mathbf{AI} = \mathbf{A}$.

Also, if $\mathbf{I}_{M\times M}\mathbf{A} = \mathbf{D} = (d_{ij})$ then

$$d_{kp} = \delta_{k1}a_{1p} + \delta_{k2}a_{2p} + \cdots + \delta_{kM}a_{Mp} = a_{kp}.$$

Thus, if $\mathbf{IA}$ exists then $\mathbf{IA} = \mathbf{A}$.

Let $\mathbf{A} = (a_{ij})$ be an $M \times N$ matrix and $\mathbf{B} = (b_{ij})$ and $\mathbf{C} = (c_{ij})$ be $N \times P$ matrices. If $\mathbf{W} = (w_{ij}) = \mathbf{A}(\mathbf{B} + \mathbf{C})$ then we can write

$$\begin{aligned} w_{kp} &= \sum_{n=1}^{N} a_{kn}(b_{np} + c_{np}) \\ &= \sum_{n=1}^{N} a_{kn}b_{np} + \sum_{n=1}^{N} a_{kn}c_{np}. \end{aligned}$$

Note that $\sum_{n=1}^{N} a_{kn}b_{np}$ and $\sum_{n=1}^{N} a_{kn}c_{np}$ are the elements in the k-th row and the p-th column of the product matrices $\mathbf{AB}$ and $\mathbf{AC}$ respectively. Thus, we find that $\mathbf{W} = \mathbf{AB} + \mathbf{AC}$, that is,

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

Similarly, if $\mathbf{A}$ is an $M \times P$ matrix and $\mathbf{B}$ and $\mathbf{C}$ are $N \times M$ matrices, we can show that

$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}.$$

Note the order of appearance of $\mathbf{A}$ in the matrix products on both sides of each of the formulae above. The order is important as multiplication of matrices is, in general, not commutative.

Let $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$ and $\mathbf{C} = (c_{ij})$ be matrices of order $M \times N$, $N \times R$ and $R \times S$ respectively. Define the $M \times R$ matrix $\mathbf{F} = (f_{ij}) = \mathbf{AB}$, where

$$f_{kp} = \sum_{n=1}^{N} a_{kn} b_{np}.$$

If the product $\mathbf{FC} = (\mathbf{AB})\mathbf{C}$ is given by the $M \times S$ matrix $\mathbf{G} = (g_{ij})$ then

$$\begin{aligned} g_{kp} &= \sum_{m=1}^{R} f_{km} c_{mp} \\ &= \sum_{m=1}^{R} (\sum_{n=1}^{N} a_{kn} b_{nm}) c_{mp} = \sum_{n=1}^{N} a_{kn} (\sum_{m=1}^{R} b_{nm} c_{mp}). \end{aligned}$$

Note that $\sum_{m=1}^{R} b_{nm} c_{mp}$ gives the element in the n-th row and p-th column of the product matrix $\mathbf{BC}$. Hence we can write $\mathbf{G} = \mathbf{A}(\mathbf{BC})$, that is,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

Multiplication of matrices is associative. Since $(\mathbf{AB})\mathbf{C}$ and $\mathbf{A}(\mathbf{BC})$ are equal to each other, they can just be written as $\mathbf{ABC}$. Note that $\mathbf{ABC} \neq \mathbf{ACB}$ as multiplication of matrices is not commutative.

1.3.5 Powers of square matrices

Let $\mathbf{A}$ be an $N \times N$ matrix.

We define:

$$\begin{aligned} \mathbf{A}^1 &= \mathbf{A}, \\ \mathbf{A}^2 &= \mathbf{AA}^1 = \mathbf{AA} \\ \mathbf{A}^3 &= \mathbf{AA}^2 = \mathbf{AAA} \\ \mathbf{A}^4 &= \mathbf{AA}^3 = \mathbf{AAAA} \\ &\vdots \\ \mathbf{A}^m &= \mathbf{AA}^{m-1} = \underbrace{\mathbf{AAA} \cdots \mathbf{A}}_{\text{product of } m \text{ matrices}} \\ &\vdots \end{aligned}$$

Note that $\mathbf{A}^m$ is of order $N \times N$ and the matrices $\mathbf{A}^2$, $\mathbf{A}^3$, $\mathbf{A}^4$ and so on do not exist if $\mathbf{A}$ is not a square matrix.

It is obvious that if c is a number then $(c\mathbf{A})^m = c^m \mathbf{A}^m$.

If $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same order then

$$\begin{aligned}(\mathbf{A}+\mathbf{B})^2 &= (\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B}) \\ &= (\mathbf{A}+\mathbf{B})\mathbf{A}+ (\mathbf{A}+\mathbf{B})\mathbf{B} \\ &= \mathbf{A}^2+\mathbf{BA}+\mathbf{AB}+\mathbf{B}^2.\end{aligned}$$

Note that we cannot write $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$ or $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{BA} + \mathbf{B}^2$ because $\mathbf{BA} \neq \mathbf{AB}$ in general.

If the diagonal element on the k-th row of an $N \times N$ diagonal matrix $\mathbf{D}$ is p_k then the matrix $\mathbf{D}^m$ (m is a positive integer) is also an $N \times N$ diagonal matrix with diagonal element p_k^m on the k-th row.

Example:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}^6 = \begin{pmatrix} (-1)^6 & 0 & 0 \\ 0 & 2^6 & 0 \\ 0 & 0 & 3^6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 64 & 0 \\ 0 & 0 & 729 \end{pmatrix}$$

We can prove the proposition that $\mathbf{D}^m$ is diagonal with diagonal element p_k^m on the k-th row by using the principle of mathematical induction.

The proposition is true for $m = 1$, since $\mathbf{D}^1 = \mathbf{D}$.

We shall now show that if the proposition is true for $m = M$ (M is a given positive integer) then it is also true for $m = M + 1$.

Let $\mathbf{D} = (d_{ij})$ where

$$d_{ij} = \begin{cases} p_k & \text{if} \quad i = j = k, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Define $\mathbf{D}^M = (f_{ij})$ and $\mathbf{D}^{M+1} = (g_{ij})$.

If the above mentioned proposition is true for $m = M$ then

$$f_{ij} = \begin{cases} p_k^M & \text{if} \quad i = j = k, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Since $\mathbf{D}^{M+1} = \mathbf{D}\mathbf{D}^M$, we can write

$$g_{ij} = \sum_{s=1}^{N} d_{is} f_{sj} = d_{i1} f_{1j} + d_{i2} f_{2j} + d_{i3} f_{3j} + \cdots + d_{iN} f_{Nj}.$$

For any fixed value of s, if $i \neq j$, either d_{is} or f_{sj} (or both) is zero. Hence $g_{ij} = 0$ if $i \neq j$, that is, $\mathbf{D}^{M+1} = (g_{ij})$ is a diagonal matrix.

If $i = j = k$, the term $d_{is} f_{sj}$ is zero for all values of s except $s = k$. It follows that

$$g_{kk} = d_{kk} f_{kk} = p_k p_k^M = p_k^{M+1},$$

that is, the diagonal element in the k-th row of $\mathbf{D}^{M+1}$ is p_k^{M+1}.

Thus, if the proposition that $\mathbf{D}^m$ is diagonal with diagonal element p_k^m on the k-th row is true for $m = M$, it is also true for $m = M + 1$. As pointed out above, the proposition is true for $m = 1$. By induction, we deduce that it is true for $m = 2, 3, 4, \cdots$.

1.4 Transpose of a matrix

Let $\mathbf{A} = (a_{ij})$ be an $M \times N$ matrix. We define the transpose of $\mathbf{A}$ to be the $N \times M$ matrix $\mathbf{B} = (b_{ij})$, where $b_{ij} = a_{ji}$ for $i = 1, 2, \cdots, N$ and $j = 1, 2, \cdots, M$. The elements in the i-th column of $\mathbf{B}$ are the elements in the i-th row of $\mathbf{A}$ in the same order of appearance. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}^T$. It should be obvious that $(\mathbf{A}^T)^T = \mathbf{A}$ and $(c\mathbf{A})^T = c\mathbf{A}^T$ where c is a number.

Examples:

1. $\begin{pmatrix} 1 & 10 & 9 & 3 \end{pmatrix}^T = \begin{pmatrix} 1 \\ 10 \\ 9 \\ 3 \end{pmatrix}$

2. $\begin{pmatrix} 2 \\ 3 \end{pmatrix}^T = \begin{pmatrix} 2 & 3 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & 5 \end{pmatrix}$

4. $\left(\begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & 5 \end{pmatrix}^T\right)^T = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & 5 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 6 & 5 & 0 \\ 2 & 7 & 5 & 7 \\ 4 & 1 & 8 & 2 \\ 5 & 4 & 9 & 8 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 & 4 & 5 \\ 6 & 7 & 1 & 4 \\ 5 & 5 & 8 & 9 \\ 0 & 7 & 2 & 8 \end{pmatrix}$

6. $\left(2\begin{pmatrix} 1 & 6 \\ 2 & 7 \end{pmatrix}\right)^T = \begin{pmatrix} 2 & 4 \\ 12 & 14 \end{pmatrix} = 2\begin{pmatrix} 1 & 6 \\ 2 & 7 \end{pmatrix}^T$

Let $\mathbf{V} = (v_{ij})$ and $\mathbf{W} = (w_{ij})$ be matrices of order $M \times N$. If $\mathbf{P} = (p_{ij}) = \mathbf{V} + \mathbf{W}$, $\mathbf{P}^T = (q_{ij})$, $\mathbf{V}^T = (r_{ij})$ and $\mathbf{W}^T = (s_{ij})$ then

$$p_{ij} = v_{ij} + w_{ij} \text{ for } i = 1, 2, \cdots, M \text{ and } j = 1, 2, \cdots, N,$$

and

$$\begin{aligned} q_{ij} &= v_{ji} + w_{ji} = r_{ij} + s_{ij} \\ &\quad \text{for } i = 1, 2, \cdots, N \text{ and } j = 1, 2, \cdots, M. \end{aligned}$$

From $q_{ij} = r_{ij} + s_{ij}$, we know that $\mathbf{P}^T = \mathbf{V}^T + \mathbf{W}^T$, that is,

$$(\mathbf{V} + \mathbf{W})^T = \mathbf{V}^T + \mathbf{W}^T.$$

Let $\mathbf{V} = (v_{ij})$ and $\mathbf{W} = (w_{ij})$ be two matrices of order $M \times N$ and $N \times R$ respectively. If $\mathbf{P} = (p_{ij}) = \mathbf{VW}$, $\mathbf{P}^T = (q_{ij})$, $\mathbf{V}^T = (r_{ij})$, $\mathbf{W}^T = (s_{ij})$ and $\mathbf{Z} = (z_{ij}) = \mathbf{W}^T\mathbf{V}^T$ then

$$p_{k\ell} = \sum_{n=1}^{N} v_{kn} w_{n\ell} \text{ for } k = 1, 2, \cdots, M \text{ and } \ell = 1, 2, \cdots, R,$$

and

$$z_{k\ell} = \sum_{n=1}^{N} s_{kn} r_{n\ell} = \sum_{n=1}^{N} w_{nk} v_{\ell n} = \sum_{n=1}^{N} v_{\ell n} w_{nk} = p_{\ell k} = q_{k\ell}$$
$$\text{for } k = 1, 2, \cdots, R \text{ and } \ell = 1, 2, \cdots, M.$$

From $q_{k\ell} = z_{k\ell}$, we know that $\mathbf{P}^T = \mathbf{W}^T \mathbf{V}^T$, that is,

$$(\mathbf{VW})^T = \mathbf{W}^T \mathbf{V}^T.$$

1.5 Vectors

1.5.1 Ordered sets of numbers

In linear algebra, an N-dimensional vector is an ordered set of N numbers. A vector is defined not only by the numbers in the set but also by their order of appearance in the set.

An N-dimensional vector may be represented by either a $1 \times N$ matrix or an $N \times 1$ matrix. We use underlined bold letters such as $\underline{\mathbf{u}}$, $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ to denote vectors. Thus, if $\underline{\mathbf{u}}$ is a four-dimensional vector, we may represent it by using a 1×4 matrix or a 4×1 matrix, that is, we may write

$$\underline{\mathbf{u}} = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 \end{pmatrix} \text{ or } \underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}.$$

We refer to N-dimensional vectors represented by $1 \times N$ matrices and $N \times 1$ matrices as row and column vectors respectively.

The numbers in the ordered sets defining vectors are referred to as the vector components. The vector components of real vectors are restricted to real numbers. In general, the vector components may be complex.

In elementary physics, vectors are physical quantities in two or three-dimensional space, having magnitudes and directions. Examples of such vector quantities are displacement and velocity. The displacement of a body from a point is characterized by the distance and the direction of the body from the point. The distance which is given by a number is the magnitude of the displacement. The velocity of the body is characterized by the speed at