HANDBOOK OF MATRICES

Helmut Lütkepohl

Handbook of Matrices

H. Lütkepohl
Humboldt-Universität zu Berlin, Germany

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto · Singapore

National 01243 779777
International (+44) 1243 779777

e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk.
Visit our Home Page on http://www.wiley.co.uk
or http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

Jacaranda Wiley Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 0512

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0 471 96688 6; 0 471 97015 8 (pbk.)

Produced from camera-ready copy supplied by the authors using LaTeX
Printed and bound in Great Britain by Biddles Ltd, Guildford and King's Lynn
This book is printed on acid-free paper responsibly manufactured from sustainable forestation, or which at least two trees are planted for each one used for paper production.

Contents

Preface

Nowadays matrices are used in many fields of science. Accordingly they have become standard tools in statistics, econometrics, mathematics, engineering and natural sciences textbooks. In fact, many textbooks from these fields have chapters, appendices or sections on matrices. Often collections of those results used in a book are included. Of course, there are also numerous books and even journals on matrices. For my own work I have found it useful, however, to have a collection of important matrix results handy for quick reference in one source. Therefore I have started collecting matrix results which are important for my own work. Over the years, this collection has grown to an extent that it may now be a valuable source of results for other researchers and students as well. To make it useful for less advanced students I have also included many elementary results.

The idea is to provide a collection where special matrix results are easy to locate. Therefore there is some repetition because some results fit under different headings and are consequently listed more than once. For example, for suitable matrices A, B and C, $A(B+C) = AB + AC$ is a result on matrix multiplication as well as on matrix sums. It is therefore listed under both headings. Although reducing search costs has been an important objective in putting together this collection of results, it is still possible that a specific result is listed in a place where I would look for it but where not everyone else would, because too much repetition turned out to be counterproductive.

Of course, this collection very much reflects my own personal preferences in this respect. Also, it is, of course, not a complete collection of results. In fact, in this respect it is again very subjective. Therefore, to make the volume more useful to others in the future, I would like to hear of any readers' pet results that I have left out. Moreover, I would be grateful to learn about errors that readers discover. It seems unavoidable that flaws sneak in somewhere.

In this book, definitions and results are given only and no proofs. At the end of most sections there is a note regarding proofs. Often a reference is made to one or more textbooks where proofs can be found. No attempt has been made to reference the origins of the results. Also, computational algorithms are not given. These may again be found in the references.

As mentioned earlier, it is hoped that this book is useful as a reference for researchers and students. I should perhaps give a warning, however, that it requires some basic knowledge and understanding of matrix algebra. It is not meant to be a tool for self teaching at an introductory level. Before searching for specific matrix results readers may want to go over Chapter 1 to familiarize themselves with my notation and terminology. Definitions and brief explanations of many terms related to matrices are given in an appendix.

Generally the results are stated for complex matrices (matrices of complex numbers). Of course, they also hold for real matrices because the latter may be regarded as special complex matrices. In some sections and chapters many results are formulated for real matrices only. In those instances a note to that effect appears at the beginning of the section or chapter. To make sure that a specific result holds indeed for complex matrices it is therefore recommended to check the beginning of the section and chapter where the result is found to determine possible qualifications.

Finally, I would like to acknowledge the help of many people who commented on this volume and helped me avoiding errors. In particular, I am grateful to Alexander Benkwitz, Jörg Breitung, Maike Burda, Helmut Herwartz, Kirstin Hubrich, Martin Moryson, and Rolf Tschernig for their careful scrutiny. Of course, none of them bears responsibility for any remaining errors. Financial support was provided by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 373.

Berlin, April 1996 Helmut Lütkepohl

List of Symbols

General Notation

$\sum$	summation
$\prod$	product
$\equiv$	equivalent to, by definition equal to
$\Rightarrow$	implies that, only if
$\Longleftrightarrow$	if and only if
π	3.14159...
i	$\sqrt{-1}$, imaginary unit
$\bar{c}$	complex conjugate of a complex number c
ln	natural logarithm
exp	exponential function
sin	sine function
cos	cosine function
max	maximum
min	minimum
sup	supremum, least upper bound
inf	infimum, greatest lower bound
$\dfrac{\mathrm{d}f}{\mathrm{d}x}$	derivative of the function $f(x)$ with respect to x
$\dfrac{\partial f}{\partial x}$	partial derivative of the function $f(\cdot)$ with respect to x
$n!$	$= 1 \cdot 2 \cdots n$
$\dbinom{n}{k}$	$= \dfrac{n!}{k!(n-k)!}$ binomial coefficient

Sets and Spaces

$\{. : \ldots\}, \{\ldots\}$	set
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers

$\mathbb{N}$	positive integers
$\mathbb{Z}$	integers
$\mathbb{R}^m$	m-dimensional Euclidean space, space of real $(m \times 1)$ vectors
$\mathbb{R}^{m \times n}$	real $(m \times n)$ matrices
$\mathbb{C}^m$	m-dimensional complex space, space of complex $(m \times 1)$ vectors
$\mathbb{C}^{m \times n}$	complex $(m \times n)$ matrices

General Matrices

$A = [a_{ij}]$	matrix with typical element a_{ij}
$A\ (m \times n)$	matrix with m rows and n columns
$A = [A_1 : A_2]$	matrix consisting of submatrices A_1 and A_2
$A = [A_1 : \ldots : A_n]$	matrix consisting of submatrices $A_1, \ldots, A_n$
$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$	matrix consisting of submatrices A_1 and A_2
$A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$	matrix consisting of submatrices $A_1, \ldots, A_n$
$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$	partitioned matrix consisting of submatrices $A_{11}, A_{12}, A_{21}, A_{22}$
$A = \begin{bmatrix} A_{11} & \ldots & A_{1q} \\ \vdots & \ddots & \vdots \\ A_{p1} & \ldots & A_{pq} \end{bmatrix}$	partitioned matrix consisting of submatrices A_{ij}

Special Matrices

D_m	$(m^2 \times \frac{1}{2}m(m+1))$ duplication matrix, 9
I_m	$(m \times m)$ identity matrix, 10
K_{mn} or $K_{m,n}$	$(mn \times mn)$ commutation matrix, 9
L_m	$(\frac{1}{2}m(m+1) \times m^2)$ elimination matrix, 9
$O_{m \times n}$	$(m \times n)$ zero or null matrix, 2
0	zero, null matrix or zero vector

Matrix Operations

$A + B$	sum of matrices A and B, 3
$A - B$	difference between matrices A and B, 3
AB	product of matrices A and B, 3

cA, Ac	product of a scalar c and a matrix A, 3
$A \otimes B$	Kronecker product of matrices A and B, 3
$A \odot B$	Hadamard or elementwise product of matrices A and B, 3
$A \oplus B$	direct sum of matrices A and B, 4

Matrix Transformations and Functions

$\det A$, $\det(A)$	determinant of a matrix A, 6
$\operatorname{tr} A$, $\operatorname{tr}(A)$	trace of a matrix A, 4
$\|A\|$	norm of a matrix A, 101
$\|A\|_2$	Euclidean norm of a matrix A, 103
$\operatorname{rk} A$, $\operatorname{rk}(A)$	rank of a matrix A, 12
$\|A\|_{\mathrm{abs}}$	absolute value or modulus of a matrix A, 5
A'	transpose of a matrix A, 5
$\bar{A}$	conjugate of a matrix A, 5
A^H	conjugate transpose or Hermitian adjoint of a matrix A, 5
A^{adj}	adjoint of a matrix A, 6
A^{-1}	inverse of a matrix A, 7
A^-	generalized inverse of a matrix A, 7
A^+	Moore–Penrose inverse of a matrix A, 7
A^i	ith power of a matrix A, 7
$A^{1/2}$	square root of a matrix A, 8
$\mathrm{dg}([a_{ij}])$	$\equiv \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix}$
$\mathrm{diag}(a_{11}, \ldots, a_{mm})$	$\equiv \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix}$
vec, col	column stacking operator, 8
rvec	row stacking operator, 8
row	operator that stacks the rows of a matrix in a column vector, 8
vech	half-vectorization operator, 8

Matrix Inequalities

$A > 0$	all elements of A are positive real numbers, 4

$A \geq 0$	all elements of A are nonnegative real numbers, 4
$A < 0$	all elements of A are negative real numbers, 4
$A \leq 0$	all elements of A are nonpositive real numbers, 4
$A > B$	each element of A is greater than the corresponding element of B, 4
$A \geq B$	each element of A is greater than or equal to the corresponding element of B, 4

Other Symbols Related to Matrices

$\mathrm{minor}(a_{ij})$	minor of the element a_{ij} of a matrix A, 6
$\mathrm{cof}(a_{ij})$	cofactor of the element a_{ij} of a matrix A, 6
$\lambda(A)$	eigenvalue of a matrix A, 63
λ_{max} or $\lambda_{max}(A)$	maximum eigenvalue of a matrix A, 63
λ_{min} or $\lambda_{min}(A)$	minimum eigenvalue of a matrix A, 63
$\sigma(A)$	singular value of a matrix A, 64
σ_{max} or $\sigma_{max}(A)$	maximum singular value of a matrix A, 64
σ_{min} or $\sigma_{min}(A)$	minimum singular value of a matrix A, 64
$\rho(A)$	spectral radius of the matrix A, 64
$p_A(\cdot)$	characteristic polynomial of a matrix A, 63
$q_A(\cdot)$	minimal polynomial of a matrix A, 215

1

Definitions, Notation, Terminology

1.1 Basic Notation and Terminology

An $(m \times n)$ **matrix** A is an array of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = [a_{ij}].$$

Alternative notations are:

$$A\ (m \times n), \qquad \underset{(m \times n)}{A}, \qquad A = [a_{ij}]\ (m \times n).$$

m and n are positive integers denoting the **row dimension** and the **column dimension**, respectively. $(m \times n)$ is the **dimension** or **order** of A. The a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, are real (elements of $\mathbb{R}$) or complex numbers (elements of $\mathbb{C}$). They are the **elements** or **entries** of the matrix A and a_{ij} is the ijth element or entry of A. The matrix A is sometimes called a **real matrix** if all its elements are real numbers. If some of its elements are complex numbers the matrix A is said to be a **complex matrix**. The set of all complex $(m \times n)$ matrices is sometimes denoted by $\mathbb{C}^{m \times n}$ and the set of all real $(m \times n)$ matrices is denoted by $\mathbb{R}^{m \times n}$.

Further notations:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix} = [A_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,q}} = [A_{ij}].$$

Here the A_{ij} are $(m_i \times n_j)$ **submatrices** of A with $\sum_{i=1}^{p} m_i = m$ and $\sum_{j=1}^{q} n_j = n$. A matrix written in terms of submatrices rather than individual elements is often called a **partitioned matrix** or a **block matrix**. Special cases are

$$A = [A_1, \ldots, A_n] = [A_1 : \ldots : A_n],$$

where the A_i are blocks of columns, and

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix},$$

where the A_i are blocks of rows of A.

A matrix

$$\underset{(m \times m)}{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix}$$

having the same number of rows and columns is a **square** or **quadratic matrix**. The elements $a_{11}, a_{22}, \ldots, a_{mm}$ (a_{ii}, $i = 1, \ldots, m$) constitute its **principal diagonal**.

$$\underset{(m \times m)}{I_m} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 1$ for $i = 1, \ldots, m$, and $a_{ij} = 0$ for $i \neq j$ is an $(m \times m)$ **identity** or **unit matrix** and

$$O_{m \times n} = \underset{(m \times n)}{\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}} = [a_{ij}]$$

with $a_{ij} = 0$ for all i, j is an $(m \times n)$ **zero** or **null matrix**. It is sometimes simply denoted by 0. A $(1 \times n)$ matrix

$$a = [a_1, \ldots, a_n]$$

is an n-**dimensional row vector** or $(1 \times n)$ **vector**. An $(m \times 1)$ matrix

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

is an m-**vector**, m-**dimensional vector** or m-**dimensional column vector**. The set of all real m-vectors is often denoted by $\mathbb{R}^m$ and the set of all complex m-vectors is denoted by $\mathbb{C}^m$. Many more special matrices are listed in Section 1.5 and the Appendix.

1.2 Operations Relating Matrices

Addition: $A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(m \times n)$

$$A + B \equiv [a_{ij} + b_{ij}] \ (m \times n)$$

(for the rules see Section 2.1).

Subtraction: $A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(m \times n)$

$$A - B \equiv [a_{ij} - b_{ij}] \ (m \times n)$$

(for the rules see Section 2.1).

Multiplication by a scalar: $A = [a_{ij}]$ $(m \times n)$, c a number (a scalar)

$$cA \equiv [ca_{ij}] \ (m \times n)$$

$$Ac \equiv [ca_{ij}] \ (m \times n)$$

(for the rules see Section 2.3).

Matrix multiplication or **matrix product:**
$A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(n \times p)$

$$AB \equiv \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right] \ (m \times p)$$

(for the rules see Section 2.2).

Kronecker product, tensor product or **direct product:**
$A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(p \times q)$

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \ (mp \times nq)$$

(for the rules see Section 2.4).

Hadamard product, Schur product or **elementwise product:**
$A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(m \times n)$

$$A \odot B \equiv [a_{ij}b_{ij}] \ (m \times n)$$

(for the rules see Section 2.5).

Direct sum: $A\,(m \times m)$, $B\,(n \times n)$

$$A \oplus B \equiv \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad ((m+n) \times (m+n))$$

(for the rules see Section 2.6).

Hierarchy of operations: If more than two matrices are related by the foregoing operations, they are performed in the following order:

(i) operations in parentheses or brackets,

(ii) matrix multiplication and multiplication by a scalar,

(iii) Kronecker product and Hadamard product,

(iv) addition and subtraction,

(v) direct sum.

Operations of the same hierarchical level are performed from left to right.

1.3 Inequality Relations Between Matrices

$A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times n)$ real:

$$\begin{array}{lcl} A > 0 & \Longleftrightarrow & a_{ij} > 0,\ i = 1,\dots,m,\ j = 1,\dots,n, \\ A \geq 0 & \Longleftrightarrow & a_{ij} \geq 0,\ i = 1,\dots,m,\ j = 1,\dots,n, \\ A > B & \Longleftrightarrow & a_{ij} > b_{ij},\ i = 1,\dots,m,\ j = 1,\dots,n, \\ A \geq B & \Longleftrightarrow & a_{ij} \geq b_{ij},\ i = 1,\dots,m,\ j = 1,\dots,n. \end{array}$$

Warning: In other literature inequality signs between matrices are sometimes used to denote different relations between positive definite and semidefinite matrices.

1.4 Operations Related to Individual Matrices

Trace: $A = [a_{ij}]$ $(m \times m)$

$$\operatorname{tr} A = \operatorname{tr}(A) \equiv a_{11} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}$$

(for its properties see Section 4.1).

Absolute value or **modulus:** $A = [a_{ij}]$ $(m \times n)$

$$|A|_{\text{abs}} \equiv [\,|a_{ij}|_{\text{abs}}\,] = \begin{bmatrix} |a_{11}|_{\text{abs}} & |a_{12}|_{\text{abs}} & \cdots & |a_{1n}|_{\text{abs}} \\ |a_{21}|_{\text{abs}} & |a_{22}|_{\text{abs}} & \cdots & |a_{2n}|_{\text{abs}} \\ \vdots & \vdots & & \vdots \\ |a_{m1}|_{\text{abs}} & |a_{m2}|_{\text{abs}} & \cdots & |a_{mn}|_{\text{abs}} \end{bmatrix} \quad (m \times n)$$

where $|c|_{\text{abs}}$ denotes the modulus of a complex number $c = c_1 + ic_2$ which is defined as $|c|_{\text{abs}} = \sqrt{c_1^2 + c_2^2} = \sqrt{c\bar{c}}$. Here $\bar{c}$ is the complex conjugate of c. Properties of the absolute value of a matrix are given in Section 3.8.

Warning: In other literature $|A|$ sometimes denotes the determinant of the matrix A.

Transpose: $A = [a_{ij}]$ $(m \times n)$

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}' \equiv \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$
$$(n \times m)$$

that is, the rows of A are the columns of A' (see Section 3.1 for its properties). In some other literature the transpose of a matrix A is denoted by A^T.

Conjugate: $A = [a_{ij}]$ $(m \times n)$

$$\bar{A} \equiv [\bar{a}_{ij}],$$

where $\bar{a}_{ij}$ is the complex conjugate of a_{ij} (see Section 3.2 for its properties).

Conjugate transpose or **Hermitian adjoint:** $A = [a_{ij}]$ $(m \times n)$

$$A^H \equiv \bar{A}' = [\bar{a}_{ij}]' \ (n \times m)$$

(see Section 3.3 for its properties).

Diagonal matrices: $A = [a_{ij}]$ $(m \times m)$

$$\text{dg}(A) \equiv \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix}$$

$$\operatorname{diag}(a_{11}, \ldots, a_{mm}) = \operatorname{diag}\left(\begin{bmatrix} a_{11} \\ \vdots \\ a_{mm} \end{bmatrix}\right) \equiv \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix}$$

(see Section 9.4 for further details).

Determinant: $A = [a_{ij}]$ $(m \times m)$

$$\det A = \det(A) \equiv \sum (-1)^p a_{1i_1} a_{2i_2} \times \cdots \times a_{mi_m},$$

where the sum is taken over all products consisting of precisely one element from each row and each column of A multiplied by -1 or 1, if the permutation $i_1, \ldots, i_m$ is odd or even, respectively (see Section 4.2 for the properties of determinants).

Minor of a matrix: The determinant of a submatrix of an $(m \times m)$ matrix A is a minor of A.

Minor of an element of a matrix: $A = [a_{ij}]$ $(m \times m)$. The minor of a_{ij} is the determinant of the matrix obtained by deleting the ith row and jth column from A,

$$\operatorname{minor}(a_{ij}) \equiv \det \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,m} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{mm} \end{bmatrix}.$$

Principal minor: $A = [a_{ij}]$ $(m \times m)$

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

is a principal minor of A for $k = 1, \ldots, m-1$.

Cofactor of an element of a matrix: $A = [a_{ij}]$ $(m \times m)$. The cofactor of a_{ij} is

$$\operatorname{cof}(a_{ij}) \equiv (-1)^{i+j} \operatorname{minor}(a_{ij}).$$

Adjoint: $A = [a_{ij}]$ $(m \times m), m > 1$,

$$A^{adj} \equiv \begin{bmatrix} \operatorname{cof}(a_{11}) & \cdots & \operatorname{cof}(a_{1m}) \\ \vdots & \ddots & \vdots \\ \operatorname{cof}(a_{m1}) & \cdots & \operatorname{cof}(a_{mm}) \end{bmatrix}' = [\operatorname{cof}(a_{ij})]'.$$

$A^{adj} \equiv 1$ if $m = 1$ (see Section 3.4 for the properties of the adjoint).

Inverse: $A = [a_{ij}]$ $(m \times m)$ with $\det(A) \neq 0$. The inverse of A is the unique $(m \times m)$ matrix A^{-1} satisfying $AA^{-1} = A^{-1}A = I_m$ (see Section 3.5 for its properties).

Generalized inverse: An $(n \times m)$ matrix A^- is a generalized inverse of the $(m \times n)$ matrix A if it satisfies $AA^-A = A$ (see Section 3.6 for its properties).

Moore–Penrose (generalized) inverse: The $(n \times m)$ matrix A^+ is the Moore–Penrose (generalized) inverse of the $(m \times n)$ matrix A if it satisfies

(i) $AA^+A = A$,

(ii) $A^+AA^+ = A^+$,

(iii) $(AA^+)^H = AA^+$,

(iv) $(A^+A)^H = A^+A$

(see Section 3.6.2 for its properties).

Power of a matrix: A $(m \times m)$

$$A^i \equiv \begin{cases} \prod_{j=1}^{i} A = \underbrace{A \times \cdots \times A}_{i \text{ times}} & \text{for positive integers } i \\ I_m & \text{for } i = 0 \\ \left(\prod_{j=1}^{-i} A\right)^{-1} & \text{for negative integers } i, \text{ if } \det(A) \neq 0 \end{cases}$$

If A can be written as

$$A = U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} U^H$$

for some unitary matrix U (see Section 1.5 and Chapter 6) then the power of A is defined for any $\alpha \in \mathbb{R}, \alpha > 0$, as follows:

$$A^\alpha \equiv U \begin{bmatrix} \lambda_1^\alpha & & 0 \\ & \ddots & \\ 0 & & \lambda_m^\alpha \end{bmatrix} U^H.$$

This definition applies for instance for Hermitian and real symmetric matrices (see Section 1.5 for the definitions of Hermitian and symmetric matrices and Section 3.7 for the properties of powers of matrices).

Square root of a matrix: The $(m \times m)$ matrix $A^{1/2}$ is a square root of the $(m \times m)$ matrix A if $A^{1/2}A^{1/2} = A$. Elsewhere in the literature a matrix B satisfying $B'B = A$ or $BB' = A$ is sometimes regarded as a square root of A.

Vectorization: $A = [a_{ij}]$ $(m \times n)$

$$\text{vec}\, A = \text{vec}(A) = \text{col}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ a_{13} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1)$$

that is, vec stacks the columns of A in a column vector.

$$\text{rvec}(A) = [\text{vec}(A')]'$$

that is, rvec stacks the rows of A in a row vector and

$$\text{row}(A) = \text{vec}(A') = \text{rvec}(A)'$$

that is, row stacks the rows of A in a column vector. (See Chapter 7 for the properties of vectorization operators).

Half-vectorization: $A = [a_{ij}]$ $(m \times m)$

$$\text{vech}\, A = \text{vech}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{33} \\ \vdots \\ a_{mm} \end{bmatrix} \quad (\tfrac{1}{2}m(m+1) \times 1)$$

that is, vech stacks the columns of A from the principal diagonal downwards in a column vector (see Chapter 7 for the properties of the half-vectorization operator and more details).

1.5 Some Special Matrices

Commutation matrix: The $(mn \times mn)$ matrix K_{mn} is a commutation matrix if $\text{vec}(A') = K_{mn}\text{vec}(A)$ for any $(m \times n)$ matrix A. It is sometimes denoted by $K_{m,n}$. For example,

$$K_{32} = K_{3,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is a commutation matrix. (For the properties of commutation matrices see Section 9.2.)

Diagonal matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $i \neq j$ is a diagonal matrix. (For the properties of diagonal matrices see Section 9.4.)

Duplication matrix: An $(m^2 \times \frac{1}{2}m(m+1))$ matrix D_m is a duplication matrix if $\text{vec}(A) = D_m\text{vech}(A)$ for any symmetric $(m \times m)$ matrix A. For example,

$$D_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is a duplication matrix. (For the properties of duplication matrices see Section 9.5.)

Elimination matrix: A $(\frac{1}{2}m(m+1) \times m^2)$ elimination matrix L_m is defined such that $\text{vech}(A) = L_m\text{vec}(A)$ for any $(m \times m)$ matrix A.

For example,

$$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is an elimination matrix. (For the properties of elimination matrices see Section 9.6.)

Hermitian matrix: An $(m \times m)$ matrix A is Hermitian if $\bar{A}' = A^H = A$. (For its properties see Section 9.7.)

Idempotent matrix: An $(m \times m)$ matrix A is idempotent if $A^2 = A$. (For the properties see Section 9.8.)

Identity matrix: An $(m \times m)$ matrix

$$I_m = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 1$ for $i = 1, \ldots, m$ and $a_{ij} = 0$ for $i \neq j$ is an identity or unit matrix.

Nonnegative matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is nonnegative if $a_{ij} \geq 0$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$. (For the properties see Section 9.9.)

Nonsingular matrix: An $(m \times m)$ matrix A is said to be nonsingular or invertible or regular if $\det(A) \neq 0$ and thus A^{-1} exists. (For the rules for matrix inversion see Section 3.5.)

Normal matrix: An $(m \times m)$ matrix A is normal if $A^H A = AA^H$.

Null matrix: An $(m \times n)$ matrix is a null matrix or zero matrix, denoted by $O_{m \times n}$ or simply by 0, if all its elements are zero.

Orthogonal matrix: An $(m \times m)$ matrix A is orthogonal if A is nonsingular and $A' = A^{-1}$. (For the properties see Section 9.10.)

Positive and **negative definite** and **semidefinite matrices:** A Hermitian or real symmetric $(m \times m)$ matrix A is **positive definite** if $x^H Ax > 0$ for all $(m \times 1)$ vectors $x \neq 0$; it is **positive semidefinite**

if $x^H Ax \geq 0$ for all $(m \times 1)$ vectors x; it is **negative definite** if $x^H Ax < 0$ for all $(m \times 1)$ vectors $x \neq 0$; it is **negative semidefinite** if $x^H Ax \leq 0$ for all $(m \times 1)$ vectors x; it is **indefinite** if $(m \times 1)$ vectors x and y exist such that $x^H Ax > 0$ and $y^H Ay < 0$ (see Section 9.12).

Positive matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is positive if $a_{ij} > 0$ for $i = 1, \ldots, m,\ j = 1, \ldots, n$ (see Section 9.9).

Symmetric matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ with $a_{ij} = a_{ji}$, $i, j = 1, \ldots, m$ is symmetric. In other words, A is symmetric if $A' = A$ (see Section 9.13).

Triangular matrices: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ a_{m1} & \cdots & \cdots & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $j > i$ is **lower triangular**. An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1m} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $i > j$ is **upper triangular** (see Section 9.14).

Unitary matrix: An $(m \times m)$ matrix A is unitary if it is nonsingular and $A^H = A^{-1}$ (see Section 9.15).

Note: Many more special matrices are listed in the Appendix.

1.6 Some Terms and Quantities Related to Matrices

Linear independence of vectors: The m-dimensional row or column vectors $x_1, \ldots, x_k$ are **linearly independent** if, for complex numbers $c_1, \ldots, c_k$, $c_1 x_1 + \cdots + c_k x_k = 0$ implies $c_1 = \cdots = c_k = 0$. They are **linearly dependent** if $c_1 x_1 + \cdots + c_k x_k = 0$ holds with at least one $c_i \neq 0$. In other words, $x_1, \ldots, x_k$ are linearly dependent if $\alpha_j \in \mathbb{C}$ exist such that for some $i \in \{1, \ldots, k\}$, $x_i = \alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \cdots + \alpha_k x_k$.

Rank: $A = [a_{ij}]$ $(m \times n)$

rk A = rk(A)
$\equiv$ maximum number of linearly independent rows or columns of A

row rk A = row rk(A)
$\equiv$ maximum number of linearly independent rows of A

col rk A = col rk(A)
$\equiv$ maximum number of linearly independent columns of A

(for rules related to the rank of a matrix see Section 4.3).

Elementary operations: The following changes to a matrix are called elementary operations:

(i) interchanging two rows or two columns,

(ii) multiplying any row or column by a nonzero number,

(iii) adding a multiple of one row to another row,

(iv) adding a multiple of one column to another column.

Quadratic form: Given a real symmetric $(m \times m)$ matrix A, the function $Q : \mathbb{R}^m \longrightarrow \mathbb{R}$ defined by $Q(x) = x'Ax$ is called a quadratic form. The quadratic form is called **positive (semi) definite** if A is positive (semi) definite. It is called **negative (semi) definite** if A is negative (semi) definite. It is **indefinite** if A is indefinite.

Hermitian form: Given a Hermitian $(m \times m)$ matrix A, the function $Q : \mathbb{C}^m \longrightarrow \mathbb{R}$ defined by $Q(x) = x^H Ax$ is called a Hermitian form. The Hermitian form is called **positive (semi) definite** if A is positive (semi) definite. It is called **negative (semi) definite** if A is negative (semi) definite. It is **indefinite** if A is indefinite.

Characteristic polynomial: The polynomial in λ given by $\det(\lambda I_m - A)$ is the characteristic polynomial of the $(m \times m)$ matrix A (see Section 5.1).

Characteristic determinant: The determinant $\det(\lambda I_m - A)$ is the characteristic determinant of the $(m \times m)$ matrix A.

Characteristic equation: The equation $\det(\lambda I_m - A) = 0$ is the characteristic equation of the $(m \times m)$ matrix A.

Eigenvalue, characteristic value, characteristic root or **latent root:** The roots of the characteristic polynomial of an $(m \times m)$ matrix A

are the eigenvalues, the characteristic values, the characteristic roots or the latent roots of A (see Chapter 5).

Eigenvector or **characteristic vector:** An $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda v$, where λ is an eigenvalue of the $(m \times m)$ matrix A, is an eigenvector or characteristic vector of A corresponding to or associated with the eigenvalue λ (see Chapter 5).

Singular value: The singular values of an $(m \times n)$ matrix A are the nonnegative square roots of the eigenvalues of AA^H if $m \leq n$, and of $A^H A$, if $m \geq n$ (see Chapter 5).

2

Rules for Matrix Operations

In the following all matrices are assumed to be complex matrices unless otherwise stated. All rules for complex matrices also hold for real matrices because the latter may be regarded as special complex matrices.

2.1 Rules Related to Matrix Sums and Differences

(1) $A, B\ (m \times n), c \in \mathbb{C}$:

(a) $A \pm B = B \pm A$.

(b) $c(A \pm B) = (A \pm B)c = cA \pm cB$.

(2) $A, B, C\ (m \times n)$: $(A \pm B) \pm C = A \pm (B \pm C) = A \pm B \pm C$.

(3) $A\ (m \times n), B, C\ (n \times r)$: $A(B \pm C) = AB \pm AC$.

(4) $A, B\ (m \times n), C\ (n \times r)$: $(A \pm B)C = AC \pm BC$.

(5) $A\ (m \times n), B, C\ (p \times q)$:

(a) $A \otimes (B \pm C) = A \otimes B \pm A \otimes C$.

(b) $(B \pm C) \otimes A = B \otimes A \pm C \otimes A$.

(6) $A, B, C\ (m \times n)$:

(a) $A \odot (B \pm C) = A \odot B \pm A \odot C$.

(b) $(A \pm B) \odot C = A \odot C \pm B \odot C$.

(7) $A, C\ (m \times m), B, D\ (n \times n)$: $(A \oplus B) \pm (C \oplus D) = (A \pm C) \oplus (B \pm D)$.

(8) $A, B\ (m \times n)$:

(a) $|A + B|_{\text{abs}} \leq |A|_{\text{abs}} + |B|_{\text{abs}}$.

(b) $(A \pm B)' = A' \pm B'$.

(c) $(A \pm B)^H = A^H \pm B^H$.

(d) $\overline{A \pm B} = \bar{A} \pm \bar{B}$.

(9) $A, B\ (m \times m)$:

(a) $\text{tr}(A \pm B) = \text{tr}(A) \pm \text{tr}(B)$.

(b) $\text{dg}(A \pm B) = \text{dg}(A) \pm \text{dg}(B)$.

(c) $\text{vech}(A \pm B) = \text{vech}(A) \pm \text{vech}(B)$.

(10) $A, B\ (m \times n)$:

(a) $\text{rk}(A \pm B) \leq \text{rk}(A) + \text{rk}(B)$.

(b) $\text{vec}(A \pm B) = \text{vec}(A) \pm \text{vec}(B)$.

(11) $A\ (m \times m)$:

(a) $A + A'$ is a symmetric $(m \times m)$ matrix.

(b) $A - A'$ is a skew-symmetric $(m \times m)$ matrix.

(c) $A + A^H$ is a Hermitian $(m \times m)$ matrix.

(d) $A - A^H$ is a skew-Hermitian $(m \times m)$ matrix.

Note: All the rules of this section follow easily from basic principles by writing down the typical elements of the matrices involved. For further details consult introductory matrix books such as Bronson (1989), Barnett (1990), Horn & Johnson (1985) and Searle (1982).

2.2 Rules Related to Matrix Multiplication

(1) $A\ (m \times n), B\ (n \times p), C\ (p \times q)$: $(AB)C = A(BC) = ABC$.

(2) $A, B\ (m \times m)$: $AB \neq BA$ in general.

(3) $A\ (m \times n), B, C\ (n \times p)$: $A(B \pm C) = AB \pm AC$.

(4) $A, B\ (m \times n), C\ (n \times p)$: $(A \pm B)C = AC \pm BC$.

(5) $A\ (m \times n), B\ (p \times q), C\ (n \times r), D\ (q \times s)$:

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

(6) $A, C\ (m \times m), B, D\ (n \times n)$: $(A \oplus B)(C \oplus D) = AC \oplus BD$.

(7) $A\ (m \times m), B\ (m \times n)$: $|AB|_{\text{abs}} \leq |A|_{\text{abs}} |B|_{\text{abs}}$.

(8) $A\ (m \times n), B\ (n \times p)$:

(a) $(AB)' = B'A'$.

(b) $(AB)^H = B^H A^H$.

(c) $\overline{AB} = \bar{A}\bar{B}$.

(9) $A, B\ (m \times m)$ nonsingular : $(AB)^{-1} = B^{-1} A^{-1}$.

(10) $A = [A_{ij}]$ $(m \times n)$ with A_{ij} $(m_i \times n_j)$, $B = [B_{ij}]$ $(n \times p)$ with B_{ij} $(n_i \times p_j)$:

$$AB = \left[\sum_k A_{ik} B_{kj} \right].$$

(11) $A\ (m \times n), B\ (n \times m):\quad \mathrm{tr}(AB) = \mathrm{tr}(BA).$

(12) $A, B\ (m \times m):\quad \det(AB) = \det(A)\det(B).$

(13) $A\ (m \times n), B\ (n \times p):\quad \mathrm{rk}(AB) \leq \min\{\mathrm{rk}(A), \mathrm{rk}(B)\}.$

(14) $A\ (m \times n):\quad \mathrm{rk}(AA') = \mathrm{rk}(A'A) = \mathrm{rk}(A).$

(15) $A\ (m \times n), B\ (n \times p):$

$$\mathrm{vec}(AB) = (I_p \otimes A)\mathrm{vec}(B) = (B' \otimes I_m)\mathrm{vec}(A) = (B' \otimes A)\mathrm{vec}(I_n).$$

(16) $A\ (m \times n), B\ (n \times p), C\ (p \times q):\quad \mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B).$

(17) $A\ (m \times n), B\ (n \times m):\quad \mathrm{tr}(AB) = \mathrm{vec}(A')'\mathrm{vec}(B) = \mathrm{vec}(B')'\mathrm{vec}(A).$

(18) $A\ (m \times n), B\ (n \times p), C\ (p \times m):\quad \mathrm{tr}(ABC) = \mathrm{vec}(A')'(C' \otimes I)\mathrm{vec}(B).$

(19) $A\ (m \times n), B\ (n \times p), C\ (p \times q), D\ (q \times m):$

$$\mathrm{tr}(ABCD) = \mathrm{vec}(D')'(C' \otimes A)\mathrm{vec}(B) = \mathrm{vec}(D)'(A \otimes C')\mathrm{vec}(B').$$

(20) $A\ (m \times m), B\ (m \times n):\quad A$ nonsingular $\Rightarrow \mathrm{rk}(AB) = \mathrm{rk}(B).$

(21) $A\ (m \times n), B\ (n \times n):\quad B$ nonsingular $\Rightarrow \mathrm{rk}(AB) = \mathrm{rk}(A).$

(22) $A\ (m \times m), B\ (m \times n), C\ (n \times n):$
A, C nonsingular $\Rightarrow \mathrm{rk}(ABC) = \mathrm{rk}(B).$

(23) $A\ (m \times m)$ nonsingular : $\quad AA^{-1} = A^{-1}A = I_m.$

(24) $A\ (m \times n):$

(a) $AI_n = I_m A = A.$

(b) $AO_{n \times p} = O_{m \times p},\quad O_{p \times m}A = O_{p \times n}.$

(25) $A\ (m \times n)$ real:

(a) AA' and $A'A$ are symmetric positive semidefinite.

(b) $\mathrm{rk}(A) = m \Rightarrow AA'$ is positive definite.

(c) $\mathrm{rk}(A) = n \Rightarrow A'A$ is positive definite.

(26) $A\ (m \times n):$

(a) AA^H and A^HA are Hermitian positive semidefinite.

(b) $\mathrm{rk}(A) = m \Rightarrow AA^H$ is positive definite.

(c) $\mathrm{rk}(A) = n \Rightarrow A^HA$ is positive definite.

Note: Most of the results of this section can be proven by considering typical elements of the matrices involved or follow directly from the definitions (see also introductory books such as Bronson (1989), Barnett (1990), Horn & Johnson (1985), Lancaster & Tismenetsky (1985) and Searle (1982)). The rules involving the vec operator can be found, e.g., in Magnus & Neudecker (1988).

2.3 Rules Related to Multiplication by a Scalar

(1) $A\ (m \times n),\ c_1, c_2 \in \mathbb{C}$:

(a) $c_1(c_2A) = c_1(Ac_2) = (c_1c_2)A = A(c_1c_2)$.

(b) $(c_1 \pm c_2)A = c_1A \pm c_2A$.

(2) $A, B\ (m \times n),\ c \in \mathbb{C}$: $\quad c(A \pm B) = cA \pm cB$.

(3) $A\ (m \times n), B\ (n \times p),\ c_1, c_2 \in \mathbb{C}$: $\quad (c_1A)(c_2B) = c_1c_2AB$.

(4) $A\ (m \times n), B\ (p \times q),\ c_1, c_2 \in \mathbb{C}$:

(a) $c_1(A \otimes B) = (c_1A) \otimes B = A \otimes (c_1B)$.

(b) $c_1A \otimes c_2B = (c_1c_2)(A \otimes B)$.

(5) $A, B\ (m \times n),\ c_1, c_2 \in \mathbb{C}$:

(a) $c_1(A \odot B) = (c_1A) \odot B = A \odot (c_1B)$.

(b) $c_1A \odot c_2B = (c_1c_2)(A \odot B)$.

(6) $A\ (m \times m), B\ (n \times n), c \in \mathbb{C}$: $\quad c(A \oplus B) = cA \oplus cB$.

(7) $A\ (m \times n)$: $\quad 0A = A0 = O_{m \times n}$.

(8) $A\ (m \times n), c \in \mathbb{C}$:

(a) $|cA|_{\text{abs}} = |c|_{\text{abs}}|A|_{\text{abs}}$.

(b) $(cA)' = cA'$.

(c) $(cA)^H = \bar{c}A^H$.

(9) $A\ (m \times m), c \in \mathbb{C}$:

(a) $\text{tr}(cA) = c\ \text{tr}(A)$.

(b) $\text{dg}(cA) = c\,\text{dg}(A)$.

(c) A is nonsingular, $c \neq 0 \Rightarrow (cA)^{-1} = \frac{1}{c}A^{-1}$.

(d) $\det(cA) = c^m \det(A)$.

(e) $\text{vech}(cA) = c\ \text{vech}(A)$.

(f) λ is eigenvalue of $A \Rightarrow c\lambda$ is eigenvalue of cA.

(10) $A\ (m \times n), c \in \mathbb{C}$:

(a) $c \neq 0 \ \Rightarrow \ \mathrm{rk}(cA) = \mathrm{rk}(A)$.

(b) $c \neq 0 \ \Rightarrow \ (cA)^+ = \frac{1}{c}A^+$.

(c) $\mathrm{vec}(cA) = c \ \mathrm{vec}(A)$.

Note: All these rules are elementary and can be found in introductory matrix books such as Bronson (1989), Barnett (1990), Horn & Johnson (1985), Lancaster & Tismenetsky (1985) or follow directly from definitions.

2.4 Rules for the Kronecker Product

(1) $A\ (m \times n), B\ (p \times q)$: $\quad A \otimes B \neq B \otimes A$ in general.

(2) $A\ (m \times n),\ B, C\ (p \times q)$: $\quad A \otimes (B \pm C) = A \otimes B \pm A \otimes C$.

(3) $A\ (m \times n), B\ (p \times q), C\ (r \times s)$:

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C = A \otimes B \otimes C.$$

(4) $A, B\ (m \times n),\ C, D\ (p \times q)$:

$$(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D.$$

(5) $A\ (m \times n), B\ (p \times q), C\ (n \times r), D\ (q \times s)$:

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

(6) $A\ (m \times n), B\ (p \times q), C\ (r \times s), D\ (n \times k), E\ (q \times l), F\ (s \times t)$:

$$(A \otimes B \otimes C)(D \otimes E \otimes F) = AD \otimes BE \otimes CF.$$

(7) $A\ (m \times n), c \in \mathbb{C}$: $\quad c \otimes A = cA = A \otimes c$.

(8) $A\ (m \times n), B\ (p \times q),\ c, d \in \mathbb{C}$:

$$cA \otimes dB = cd(A \otimes B) = (cdA) \otimes B = A \otimes (cdB).$$

(9) $A\ (m \times m), B\ (n \times n), C\ (p \times p)$: $\quad (A \oplus B) \otimes C = (A \otimes C) \oplus (B \otimes C)$.

(10) $A\ (m \times n), B\ (p \times q)$:

(a) $(A \otimes B)' = A' \otimes B'$.

(b) $(A \otimes B)^H = A^H \otimes B^H$.

(c) $\overline{A \otimes B} = \bar{A} \otimes \bar{B}$.

(d) $(A \otimes B)^+ = A^+ \otimes B^+$.

(e) $|A \otimes B|_{\text{abs}} = |A|_{\text{abs}} \otimes |B|_{\text{abs}}$.

(f) $\text{rk}(A \otimes B) = \text{rk}(A)\ \text{rk}(B)$.

(11) $A\ (m \times m), B\ (n \times n)$:

(a) A, B nonsingular $\Rightarrow\ (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

(b) $\text{tr}(A \otimes B) = \text{tr}(A)\ \text{tr}(B)$.

(c) $\text{dg}(A \otimes B) = \text{dg}(A) \otimes \text{dg}(B)$.

(d) $\det(A \otimes B) = (\det A)^n (\det B)^m$.

(e) $\lambda(A)$ and $\lambda(B)$ are eigenvalues of A and B, respectively, with associated eigenvectors $v(A)$ and $v(B)$ $\Rightarrow$ $\lambda(A) \cdot \lambda(B)$ is eigenvalue of $A \otimes B$ with eigenvector $v(A) \otimes v(B)$.

(12) $x, y\ (m \times 1)$: $\quad x' \otimes y = yx' = y \otimes x'$.

(13) $x\ (m \times 1), y\ (n \times 1)$: $\quad \text{vec}(xy') = y \otimes x$.

(14) $A\ (m \times n), B\ (n \times p), C\ (p \times q)$: $\quad (C' \otimes A)\text{vec}(B) = \text{vec}(ABC)$.

(15) $A\ (m \times n), B\ (n \times p), C\ (p \times q), D\ (q \times m)$:

$$\begin{aligned} \text{tr}(ABCD) &= \text{vec}(D')'(C' \otimes A)\text{vec}(B) \\ &= \text{vec}(D)'(A \otimes C')\text{vec}(B'), \end{aligned}$$

$$\begin{aligned} \text{vec}(D')'(C' \otimes A)\text{vec}(B) &= \text{vec}(A')'(D' \otimes B)\text{vec}(C) \\ &= \text{vec}(B')'(A' \otimes C)\text{vec}(D) \\ &= \text{vec}(C')'(B' \otimes D)\text{vec}(A). \end{aligned}$$

Note: A substantial collection of results on Kronecker products including many of those given here can be found in Magnus (1988). Some results are also given in Magnus & Neudecker (1988) and Lancaster & Tismenetsky (1985). There are many more rules for Kronecker products related to special matrices, in particular to commutation, duplication and elimination matrices (see Sections 9.2, 9.5 and 9.6).

2.5 Rules for the Hadamard Product

(1) $A, B\ (m \times n), c \in \mathbb{C}$:

(a) $A \odot B = B \odot A$.

(b) $c(A \odot B) = (cA) \odot B = A \odot (cB)$.

(2) $A, B, C\ (m \times n)$:

(a) $A \odot (B \odot C) = (A \odot B) \odot C = A \odot B \odot C$.

(b) $(A \pm B) \odot C = A \odot C \pm B \odot C$.

(3) $A, B, C, D\ (m \times n)$:

$$(A + B) \odot (C + D) = A \odot C + A \odot D + B \odot C + B \odot D.$$

(4) $A, C\ (m \times m), B, D\ (n \times n)$: $(A \oplus B) \odot (C \oplus D) = (A \odot C) \oplus (B \odot D)$.

(5) $A, B\ (m \times n)$:

(a) $(A \odot B)' = A' \odot B'$.

(b) $(A \odot B)^H = A^H \odot B^H$.

(c) $\overline{A \odot B} = \bar{A} \odot \bar{B}$.

(d) $|A \odot B|_{\mathrm{abs}} = |A|_{\mathrm{abs}} \odot |B|_{\mathrm{abs}}$.

(6) $A\ (m \times n)$: $A \odot O_{m \times n} = O_{m \times n} \odot A = O_{m \times n}$.

(7) $A\ (m \times m)$: $A \odot I_m = I_m \odot A = \mathrm{dg}(A)$.

(8) $A\ (m \times n), J = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} (m \times m)$: $A \odot J = A = J \odot A$.

(9) $A, B, C\ (m \times n)$: $\mathrm{tr}[A'(B \odot C)] = \mathrm{tr}[(A' \odot B')C]$.

(10) $A, B\ (m \times m)$, $j = (1, \dots, 1)'\ (m \times 1)$: $\mathrm{tr}(AB') = j'(A \odot B)j$.

(11) $A, B, D\ (m \times m)$, $j = (1, \dots, 1)'\ (m \times 1)$:

$$D \text{ diagonal} \Rightarrow \mathrm{tr}(ADB'D) = j'D(A \odot B)Dj.$$

(12) $A, B, D\ (m \times m)$: D diagonal $\Rightarrow (DA) \odot (BD) = D(A \odot B)D$.

(13) $A, B\ (m \times n)$:

(a) $\mathrm{vec}(A \odot B) = \mathrm{diag}(\mathrm{vec}\, A)\mathrm{vec}(B) = \mathrm{diag}(\mathrm{vec}\, B)\mathrm{vec}(A)$.

(b) $\mathrm{vec}(A \odot B) = (\mathrm{vec}\, A) \odot \mathrm{vec}(B)$.

(14) $A, B\ (m \times m)$:

(a) $\mathrm{vech}(A \odot B) = \mathrm{diag}(\mathrm{vech}\, A)\mathrm{vech}(B) = \mathrm{diag}(\mathrm{vech}\, B)\mathrm{vech}(A)$.

(b) $\mathrm{vech}(A \odot B) = (\mathrm{vech}\, A) \odot \mathrm{vech}(B)$.

(15) (Schur product theorem)
$A, B\ (m \times m)$: A, B positive (semi) definite $\Rightarrow$ $A \odot B$ positive (semi) definite.

Note: Most results of this section follow directly from definitions. The remaining ones can be found in Magnus & Neudecker (1988). The Schur product theorem is, for example, given in Horn & Johnson (1985).

2.6 Rules for Direct Sums

(1) $A\ (m \times m), B\ (n \times n),\ c \in \mathbb{C}$:

(a) $c(A \oplus B) = cA \oplus cB$.

(b) $A \neq B \iff A \oplus B \neq B \oplus A$.

(2) $A, B\ (m \times m),\ C, D\ (n \times n)$:

(a) $(A \pm B) \oplus (C \pm D) = (A \oplus C) \pm (B \oplus D)$.

(b) $(A \oplus C)(B \oplus D) = AB \oplus CD$.

(3) $A\ (m \times m), B\ (n \times n), C\ (p \times p)$:

(a) $A \oplus (B \oplus C) = (A \oplus B) \oplus C = A \oplus B \oplus C$.

(b) $(A \oplus B) \otimes C = (A \otimes C) \oplus (B \otimes C)$.

(4) $A, D\ (m \times m),\ B, E\ (n \times n),\ C, F\ (p \times p)$:

$$(A \oplus B \oplus C)(D \oplus E \oplus F) = AD \oplus BE \oplus CF.$$

(5) $A\ (m \times m), B\ (n \times n)$:

(a) $|A \oplus B|_{\mathsf{abs}} = |A|_{\mathsf{abs}} \oplus |B|_{\mathsf{abs}}$.

(b) $(A \oplus B)' = A' \oplus B'$.

(c) $(A \oplus B)^H = A^H \oplus B^H$.

(d) $\overline{A \oplus B} = \bar{A} \oplus \bar{B}$.

(e) $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$, if A and B are nonsingular.

(f) $(A \oplus B)^{+} = A^{+} \oplus B^{+}$.

(6) $A\ (m \times m), B\ (n \times n)$:

(a) $\mathrm{tr}(A \oplus B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

(b) $\mathrm{dg}(A \oplus B) = \mathrm{dg}(A) \oplus \mathrm{dg}(B)$.

(c) $\mathrm{rk}(A \oplus B) = \mathrm{rk}(A) + \mathrm{rk}(B)$.

(d) $\det(A \oplus B) = \det(A)\det(B)$.

(e) $\lambda(A)$ and $\lambda(B)$ are eigenvalues of A and B, respectively, with associated eigenvectors $v(A)$ and $v(B)$ $\Rightarrow$ $\lambda(A)$ and $\lambda(B)$ are eigenvalues of $A \oplus B$ with eigenvectors

$$\begin{bmatrix} v(A) \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ v(B) \end{bmatrix},$$

respectively.

Note: All results of this section follow easily from the definitions (see also Horn & Johnson (1985)).

3

Matrix Valued Functions of a Matrix

In this chapter all matrices are assumed to be complex matrices unless otherwise stated. All rules for complex matrices also hold for real matrices as the latter may be regarded as special complex matrices.

3.1 The Transpose

Definition: The $(n \times m)$ matrix

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}' \equiv \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \quad (n \times m)$$

is the transpose of the $(m \times n)$ matrix $A = [a_{ij}]$.

(1) $A, B\ (m \times n): \quad (A \pm B)' = A' \pm B'$.

(2) $A\ (m \times n), c \in \mathbb{C}: \quad (cA)' = cA'$.

(3) $A\ (m \times n), B\ (n \times p): \quad (AB)' = B'A'$.

(4) $A\ (m \times n), B\ (p \times q): \quad (A \otimes B)' = A' \otimes B'$.

(5) $A, B\ (m \times n): \quad (A \odot B)' = A' \odot B'$.

(6) $A\ (m \times m), B\ (n \times n): \quad (A \oplus B)' = A' \oplus B'$.

(7) $A\ (m \times n):$

- (a) $|A'|_{\text{abs}} = |A|'_{\text{abs}}$.
- (b) $\text{rk}(A') = \text{rk}(A)$.
- (c) $(A')' = A$.

(d) $(A')^H = (A^H)' = \bar{A}$.

(e) $\overline{A'} = \bar{A}' = A^H$.

(8) $A\ (m \times m)$:

(a) $\mathrm{dg}(A') = \mathrm{dg}(A)$.

(b) $\mathrm{tr}(A') = \mathrm{tr}(A)$.

(c) $\det(A') = \det(A)$.

(d) $(A')^{-1} = (A^{-1})'$, if A is nonsingular.

(e) $(A')^{adj} = (A^{adj})'$.

(f) λ is eigenvalue of $A \Rightarrow \lambda$ is eigenvalues of A'.

(9) $A\ (m \times n)$ real: $(A')^+ = (A^+)'$.

(10) $A\ (m \times n)$, $K_{mn}\ (mn \times mn)$ commutation matrix:
$\mathrm{vec}(A') = K_{mn}\mathrm{vec}(A)$.

(11) $A, B\ (m \times n)$: $\mathrm{vec}(B')'\mathrm{vec}(A) = \mathrm{tr}(AB)$.

(12) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}' = \begin{bmatrix} A' & C' \\ B' & D' \end{bmatrix}.$$

(13) $A\ (m \times m)$:

(a) A is diagonal $\Rightarrow A' = A$.

(b) A is symmetric $\iff A' = A$.

(14) $A\ (m \times m)$ nonsingular: A orthogonal $\iff A' = A^{-1}$.

(15) $A\ (m \times n)$: $A'A$ and AA' are symmetric matrices.

Note: All these results are elementary. They follow either directly from definitions or can be obtained by considering the individual elements of the matrices involved (see, e.g., Lancaster & Tismenetsky (1985)).

3.2 The Conjugate

Definition: The $(m \times n)$ matrix $\bar{A} = [\bar{a}_{ij}]$ is the conjugate of the $(m \times n)$ matrix $A = [a_{ij}]$. Here $\bar{a}_{ij}$ denotes the complex conjugate of a_{ij}.

(1) $A, B\ (m \times n)$: $\overline{A \pm B} = \bar{A} \pm \bar{B}$.

(2) $A\ (m \times n), B\ (n \times p)$: $\overline{AB} = \bar{A}\bar{B}$.

(3) $A\ (m \times n), c \in \mathbb{C}$: $\overline{cA} = \bar{c}\bar{A}$.

(4) $A\ (m \times n), B\ (p \times q)$: $\overline{A \otimes B} = \bar{A} \otimes \bar{B}$.

(5) $A, B\ (m \times n)$: $\overline{A \odot B} = \bar{A} \odot \bar{B}$.

(6) $A\ (m \times m), B\ (n \times n)$: $\overline{A \oplus B} = \bar{A} \oplus \bar{B}$.

(7) $A\ (m \times n)$:

(a) $|\bar{A}|_{\text{abs}} = |A|_{\text{abs}}$.

(b) $\text{rk}(\bar{A}) = \text{rk}(A)$.

(c) $\bar{\bar{A}} = A$.

(d) $\bar{A}' = \overline{A'} = A^H$.

(e) $\bar{A}^H = \overline{A^H} = A'$.

(f) $\text{vec}(\bar{A}) = \overline{\text{vec}(A)}$.

(g) A is real $\Rightarrow$ $\bar{A} = A$.

(8) $A\ (m \times m)$:

(a) $\text{tr}(\bar{A}) = \overline{\text{tr}(A)}$.

(b) $\det(\bar{A}) = \overline{\det A}$.

(c) $\bar{A}^{adj} = \overline{A^{adj}}$.

(d) $\bar{A}^{-1} = \overline{A^{-1}}$, if A is nonsingular.

(e) $\text{vech}(\bar{A}) = \overline{\text{vech}(A)}$.

(f) $\text{dg}(\bar{A}) = \overline{\text{dg}(A)}$.

(g) λ is eigenvalue of A with eigenvector v $\Rightarrow$ $\bar{\lambda}$ is eigenvalue of $\bar{A}$ with eigenvector $\bar{v}$.

(9) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p)$:

$$\overline{\begin{bmatrix} A & B \\ C & D \end{bmatrix}} = \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix}.$$

(10) $a_1, \ldots, a_m \in \mathbb{C}$: $\overline{\text{diag}(a_1, \ldots, a_m)} = \text{diag}(\bar{a}_1, \ldots, \bar{a}_m)$.

Note: The results in this section are easily obtained from basic definitions or elementwise considerations and the fact that for $c_1, c_2 \in \mathbb{C}$, $\overline{c_1 c_2} = \bar{c}_1 \bar{c}_2$.

3.3 The Conjugate Transpose

Definition: The conjugate transpose of the $(m \times n)$ matrix $A = [a_{ij}]$ is the $(n \times m)$ matrix $A^H \equiv \bar{A}' = [\bar{a}_{ij}]'$.

(1) $A, B\ (m \times n)$: $(A \pm B)^H = A^H \pm B^H$.

(2) $A\ (m \times n), B\ (n \times p)$: $(AB)^H = B^H A^H$.

(3) $A\ (m \times n), c \in \mathbb{C}$: $(cA)^H = \bar{c}A^H$.

(4) $A\ (m \times n), B\ (p \times q)$: $(A \otimes B)^H = A^H \otimes B^H$.

(5) $A, B\ (m \times n)$: $(A \odot B)^H = A^H \odot B^H$.

(6) $A\ (m \times m), B\ (n \times n)$: $(A \oplus B)^H = A^H \oplus B^H$.

(7) $A\ (m \times n)$:

(a) $|A^H|_{\text{abs}} = |A|'_{\text{abs}}$.

(b) $\text{rk}(A^H) = \text{rk}(A)$.

(c) $(A^H)^H = A$.

(d) $(A')^H = \bar{A} = (A^H)'$.

(e) $\bar{A}^H = A' = \overline{(A^H)}$.

(f) $(A^H)^+ = (A^+)^H$.

(g) $\text{vec}(A^H) = \overline{\text{vec}(A')}$.

(h) A is real $\Rightarrow$ $A^H = A'$.

(8) $A\ (m \times m)$:

(a) $\text{tr}(A^H) = \text{tr}(\bar{A}) = \overline{\text{tr}(A)}$.

(b) $\det(A^H) = \det(\bar{A}) = \overline{\det A}$.

(c) $(A^H)^{adj} = (A^{adj})^H$.

(d) $(A^H)^{-1} = (A^{-1})^H$, if A is nonsingular.

(e) $\text{dg}(A^H) = \text{dg}(\bar{A}) = (\text{dg}\, A)^H$.

(9) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^H = \begin{bmatrix} A^H & C^H \\ B^H & D^H \end{bmatrix}.$$

(10) $a_1, \ldots, a_m \in \mathbb{C}$: $[\text{diag}(a_1, \ldots, a_m)]^H = \text{diag}(\bar{a}_1, \ldots, \bar{a}_m)$.

(11) $A\ (m \times m)$: A Hermitian $\iff$ $A^H = A$.

(12) $A\ (m \times m)$ nonsingular: A unitary $\iff$ $A^H = A^{-1}$.

(13) $A\ (m \times n)$:

(a) $A^H A$ and AA^H are Hermitian positive semidefinite.

(b) $\text{rk}(A) = m \Rightarrow AA^H$ is Hermitian positive definite.

(c) $\text{rk}(A) = n \Rightarrow A^H A$ is Hermitian positive definite.

Note: The results in this section are easily obtained from basic definitions or elementwise considerations (see, e.g., Barnett (1990)).

3.4 The Adjoint of a Square Matrix

Definition: For $m > 1$, the $(m \times m)$ matrix $A^{adj} = [\text{cof}(a_{ij})]'$ is the adjoint of the $(m \times m)$ matrix $A = [a_{ij}]$. Here $\text{cof}(a_{ij})$ is the cofactor of a_{ij}. For $m = 1$, $A^{adj} = 1$. For instance, for $m = 3$,

$$A^{adj} = \begin{bmatrix} \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ -\det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \\ \det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{bmatrix}'.$$

(1) $A, B\ (m \times m)$: $\quad (AB)^{adj} = B^{adj} A^{adj}$.

(2) $A\ (m \times m), c \in \mathbb{C}$: $\quad (cA)^{adj} = c^{m-1} A^{adj}$.

(3) $A\ (m \times m)$:

(a) $A^{adj} = \det(A) A^{-1}$, if A is nonsingular.

(b) $AA^{adj} = A^{adj} A = \det(A) I_m$.

(c) $(A')^{adj} = (A^{adj})'$.

(d) $(A^H)^{adj} = (A^{adj})^H$.

(e) $\det(A^{adj}) = (\det A)^{m-1}$.

(f) $\text{rk}(A) < m - 1 \Rightarrow A^{adj} = 0$.

(4) $A\ (m \times m), m \geq 2$:

$$\text{rk}(A^{adj}) = \begin{cases} m & \text{if} \quad \text{rk}(A) = m \\ 1 & \text{if} \quad \text{rk}(A) = m - 1 \\ 0 & \text{if} \quad \text{rk}(A) < m - 1 \end{cases}.$$

Note: These results follow easily from basic principles and definitions (see, e.g., Lancaster & Tismenetsky (1985) and Magnus & Neudecker (1988)).

3.5 The Inverse of a Square Matrix

Definition: An $(m \times m)$ matrix A^{-1} is the inverse of the $(m \times m)$ matrix A if $AA^{-1} = A^{-1}A = I_m$.

3.5.1 General Results

(1) $A, B\ (m \times m)$ nonsingular : $\quad (AB)^{-1} = B^{-1} A^{-1}$.

(2) A $(m \times m)$ nonsingular, $c \in \mathbb{C}, c \neq 0$: $(cA)^{-1} = c^{-1}A^{-1}$.

(3) A $(m \times m)$ nonsingular, B $(n \times n)$ nonsingular:

(a) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

(b) $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$.

(4) A $(m \times m)$ nonsingular:

(a) $A^{-1} = \dfrac{1}{\det(A)} A^{adj}$.

(b) $(A')^{-1} = (A^{-1})'$.

(c) $(\bar{A})^{-1} = \overline{A^{-1}}$.

(d) $(A^H)^{-1} = (A^{-1})^H$.

(e) $(A^{adj})^{-1} = \dfrac{1}{\det(A)} A$.

(f) $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

(g) λ is eigenvalue of A with eigenvector v $\Rightarrow$ $\frac{1}{\lambda}$ is eigenvalue of A^{-1} with eigenvector v.

(5) A $(m \times m)$:

(a) $\text{rk}(A) = m \iff A^{-1}$ exists.

(b) $\text{rk}(A) = m \Rightarrow A^{+} = A^{-1}$.

(c) A is diagonal dominant $\Rightarrow$ A is nonsingular.

(6) $a_1, \ldots, a_m \in \mathbb{C}, a_i \neq 0$ for $i = 1, \ldots m$:

$$A = \text{diag}(a_1, \ldots, a_m) \Rightarrow A^{-1} = \text{diag}(a_1^{-1}, \ldots, a_m^{-1}).$$

(7) I_m $(m \times m)$ identity matrix: $I_m^{-1} = I_m$.

(8) A $(m \times m)$ nonsingular:

(a) A orthogonal $\iff A^{-1} = A'$.

(b) A unitary $\iff A^{-1} = A^H$.

(9) A $(m \times m)$ positive definite: $(A^{1/2})^{-1}$ is a square root of A^{-1}.

Note: Many of these results are standard rules which can be found in introductory textbooks such as Barnett (1990) and Horn & Johnson (1985) or follow immediately from definitions.

3.5.2 Inverses Involving Sums and Differences

(1) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$, $|\lambda_i|_{\text{abs}} < 1$, $i = 1, \ldots, m$:

(a) $(I_m + A)^{-1} = \sum_{i=0}^{\infty}(-A)^i$.

(b) $(I_m - A)^{-1} = \sum_{i=0}^{\infty} A^i$.

(c) $(I_{m^2} + A \otimes A)^{-1} = \sum_{i=0}^{\infty}(-A)^i \otimes A^i$.

(d) $(I_{m^2} - A \otimes A)^{-1} = \sum_{i=0}^{\infty} A^i \otimes A^i$.

(2) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}$$

if all involved inverses exist.

(3) $A\ (m \times n)$:

$$\det(I_m + AA^H) \neq 0 \quad \Rightarrow \quad (I_m + AA^H)^{-1} = I_m - A(I_n + A^H A)^{-1} A^H.$$

(4) $A\ (m \times m)$ nonsingular, $B\ (m \times m)$: $(A + BB^H)$, $(I_m + B^H A^{-1} B)$ nonsingular $\Rightarrow$ $(A + BB^H)^{-1}B = A^{-1}B(I_m + B^H A^{-1}B)^{-1}$.

(5) $A, B\ (m \times m)$ nonsingular:

(a) $A^{-1} + B^{-1} = A^{-1}(A + B)B^{-1}$.

(b) $A^{-1} + B^{-1}$ nonsingular
$\Rightarrow$ $(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B = B(A + B)^{-1}A$.

(c) $A^{-1} + B^{-1} = (A + B)^{-1}$ $\Rightarrow$ $AB^{-1}A = BA^{-1}B$.

(6) $A, B\ (m \times m)$:

(a) $I_m + AB$ nonsingular $\Rightarrow$ $(I_m + AB)^{-1}A = A(I_m + BA)^{-1}$.

(b) $A + B$ nonsingular $\Rightarrow$ $A - A(A + B)^{-1}A = B - B(A + B)^{-1}B$.

Note: Most results of this subsection may be found in Searl (1982, Chapter 5) or follow from results given there. For (1) see Section 5.4.

3.5.3 Partitioned Inverses

(1) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \ A \text{ and } (D - CA^{-1}B) \text{ nonsingular}$$

$$\Rightarrow \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}.$$

(2) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \ D \text{ and } (A - BD^{-1}C) \text{ nonsingular}$$

$$\Rightarrow \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}.$$

(3) $A\ (m \times m)$ symmetric, $B\ (m \times n), C\ (m \times p), D\ (n \times n)$ symmetric, $E\ (p \times p)$ symmetric :

$$\begin{bmatrix} A & B & C \\ B' & D & 0 \\ C' & 0 & E \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} F & -FBD^{-1} & -FCE^{-1} \\ -D^{-1}B'F & D^{-1} + D^{-1}B'FBD^{-1} & D^{-1}B'FCE^{-1} \\ -E^{-1}C'F & E^{-1}C'FBD^{-1} & E^{-1} + E^{-1}C'FCE^{-1} \end{bmatrix}$$

if all inverses exist and $F = (A - BD^{-1}B' - CE^{-1}C')^{-1}$.

(4) $A_i\ (m_i \times m_i)$ nonsingular, $i = 1, \ldots, r$:

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} & & 0 \\ & \ddots & \\ 0 & & A_r^{-1} \end{bmatrix}.$$

(5) $m, n \in \mathbb{N}, m > n, A\ (m \times n), B\ (m \times (m - n))$:

$$\text{rk}(A) = n, \ \text{rk}(B) = m - n, \ A^H B = 0$$

$$\Rightarrow [A : B]^{-1} = \begin{bmatrix} (A^H A)^{-1}A^H \\ (B^H B)^{-1}B^H \end{bmatrix}.$$

(6) $m, n \in \mathbb{N}, m < n, A\ (m \times n), B\ ((n - m) \times n)$:

$$\text{rk}(A) = m, \ \text{rk}(B) = n - m, \ AB^H = 0$$

$$\Rightarrow \begin{bmatrix} A \\ B \end{bmatrix}^{-1} = [A^H(AA^H)^{-1} : B^H(BB^H)^{-1}].$$

Note: The inverses of the partitioned matrices in (1) – (3) are given in Magnus & Neudecker (1988, Chapter 1). The other results are straightforward consequences of the definition of the inverse.

3.5.4 Inverses Involving Commutation, Duplication and Elimination Matrices

Reminder:

- K_{mn} or $K_{m,n}$ denotes an $(mn \times mn)$ commutation matrix (for details, see Section 9.2).
- D_m denotes the $(m^2 \times \frac{1}{2}m(m+1))$ duplication matrix (see Section 9.5).
- L_m denotes the $(\frac{1}{2}m(m+1) \times m^2)$ elimination matrix (see Section 9.6).

(1) $K_{mn}^{-1} = K_{nm}$.

(2) $(D_m' D_m)^{-1} = D_m^+ D_m^{+'}$.

(3) $(D_{m+1}' D_{m+1})^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}I_m & 0 \\ 0 & 0 & (D_m' D_m)^{-1} \end{bmatrix}$.

(4) $(L_m L_m')^{-1} = L_m L_m'$.

(5) A $(m \times m)$ nonsingular : $\quad [D_m'(A \otimes A)D_m]^{-1} = D_m^+(A^{-1} \otimes A^{-1})D_m^{+'}$.

(6) A $(m \times m)$, $I_m \otimes A + A \otimes I_m$ nonsingular:

(a) $[D_m^+(A \otimes I_m)D_m]^{-1} = 2D_m^+(I_m \otimes A + A \otimes I_m)^{-1} D_m$.

(b) $[D_m^+(I_m \otimes A + A \otimes I_m)D_m]^{-1} = D_m^+(I_m \otimes A + A \otimes I_m)^{-1} D_m$.

(7) A, B $(m \times m)$, $A \otimes B + B \otimes A$ nonsingular:

$$[D_m^+(A \otimes B)D_m]^{-1} = 2D_m^+(A \otimes B + B \otimes A)^{-1} D_m.$$

(8) A $(m \times m)$ real symmetric nonsingular, $c \in \mathbb{R}$:

$$\begin{aligned} &\{D_m^+[A \otimes A + c\,\mathrm{vec}(A)\mathrm{vec}(A)']D_m\}^{-1} \\ &= D_m^+ \left[A^{-1} \otimes A^{-1} - \frac{c}{1+cm}\,\mathrm{vec}(A^{-1})\mathrm{vec}(A^{-1})' \right] D_m. \end{aligned}$$

(9) A, B $(m \times m)$ real, lower triangular, invertible :

$$[L_m(A' \otimes B)L_m']^{-1} = L_m((A')^{-1} \otimes B^{-1})L_m'.$$

(10) A $(m \times m)$ real, lower triangular, invertible, $c \in \mathbb{R}$:

$$\{L_m[A' \otimes A + c\,\mathrm{vec}(A)\mathrm{vec}(A')']L_m'\}^{-1}$$

$$= L_m \left[A'^{-1} \otimes A^{-1} - \frac{c}{1+cm} \operatorname{vec}(A^{-1})\operatorname{vec}(A'^{-1})' \right] L_m'.$$

Note: The results of this subsection may be found in Magnus (1988).

3.6 Generalized Inverses

Definition: An $(n \times m)$ matrix A^- is a generalized inverse of the $(m \times n)$ matrix A if it satisfies $AA^-A = A$.

3.6.1 General Results

(1) $A\ (m \times m)$: $\operatorname{rk}(A) = m \iff A^- = A^{-1}$.

(2) $A\ (m \times n)$:

(a) A^- is not unique in general.

(b) AA^- and A^-A are idempotent.

(c) $I_m - AA^-$ and $I_n - A^-A$ are idempotent.

(d) $A(A^HA)^-A^H$ is idempotent.

(3) $A\ (m \times n)$:

(a) $\operatorname{rk}(A) = \operatorname{rk}(A^-A) = \operatorname{rk}(AA^-)$.

(b) $\operatorname{rk}(A) = \operatorname{tr}(A^-A) = \operatorname{tr}(AA^-)$.

(c) $\operatorname{rk}(AA^-) = \operatorname{tr}(AA^-)$.

(d) $\operatorname{rk}(A^-) \geq \operatorname{rk}(A)$.

(e) $\operatorname{rk}(A^-) = \operatorname{rk}(A) \iff A^-AA^- = A^-$.

(f) $\operatorname{rk}(A) = n \iff A^-A = I_n$.

(g) $\operatorname{rk}(A) = m \iff AA^- = I_m$.

(4) $A\ (m \times n)$:

(a) $A(A^HA)^-A^HA = A$.

(b) $A^HA(A^HA)^-A^H = A^H$.

(c) $A(A^HA)^-A^H$ is Hermitian.

(d) $(A^-)^H$ is a generalized inverse of A^H.

(5) $O_{n\times m}$ is a generalized inverse of $O_{m\times n}$.

(6) $A\ (m\times m)$: A is idempotent $\Rightarrow$ A is a generalized inverse of itself.

(7) $A\ (m \times n), B\ (n \times m)$: A^- is a generalized inverse of A $\Rightarrow$ $A^- + B - A^-ABAA^-$ is a generalized inverse of A.

(8) A $(m \times n)$, B, C $(n \times m)$: A^- is a generalized inverse of $A \Rightarrow A^- + B(I_m - AA^-) + (I_n - A^-A)C$ is a generalized inverse of A.

(9) A $(m \times n)$, B $(m \times m)$, C $(n \times n)$:
B, C nonsingular $\Rightarrow$ $C^{-1}A^-B^{-1}$ is a generalized inverse of BAC.

(10) A $(m \times n)$, B $(p \times q)$: $A^- \otimes B^-$ is a generalized inverse of $A \otimes B$.

(11) A $(m \times n)$, B $(m \times r)$, C $(m \times r)$: Generalized inverse matrices A^- and C^- of A and C, respectively, exist such that $AA^-BC^-C = B \Rightarrow$ the system of equations $AXC = B$ can be solved for X and $X = A^-BC^- + Y - A^-AYCC^-$ is a solution for any $(n \times m)$ matrix Y.

Partitioned Matrices

(12) A $(m \times m)$ nonsingular, B $(m \times n)$, C $(r \times m)$, D $(r \times n)$:

$$\text{rk}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = m, \; D = CA^{-1}B$$

$$\Rightarrow \begin{bmatrix} A^{-1} & O_{m \times r} \\ O_{n \times m} & O_{n \times r} \end{bmatrix} \text{ is a generalized inverse of } \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

(13) A $(m \times n)$, $\text{rk}(A) = r$: B $(m \times m)$, C $(n \times n)$ are nonsingular and such that

$$BAC = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow C\begin{bmatrix} I_r & D \\ E & F \end{bmatrix} B \text{ is a generalized inverse of } A,$$

for any D $(r \times (n-r))$, E $((m-r) \times r)$, F $((m-r) \times (n-r))$.

(14) A $(m_i \times n_i)$, $i = 1, \ldots, r$:

$$\begin{bmatrix} A_1^- & & 0 \\ & \ddots & \\ 0 & & A_r^- \end{bmatrix} \text{ is a generalized inverse of } \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}.$$

Note: A number of books on generalized inverses exist which contain the foregoing results and more on generalized inverses (e.g., Rao & Mitra (1971), Pringle & Rayner (1971), Boullion & Odell (1971), Ben-Israel & Greville (1974)). Some of these books contain also extensive lists of references.

3.6.2 The Moore–Penrose Inverse

Definition: The $(n \times m)$ matrix A^+ is the Moore–Penrose (generalized) inverse of the $(m \times n)$ matrix A if it satisfies the following four conditions:

(i) $AA^+A = A$,

(ii) $A^+AA^+ = A^+$,

(iii) $(AA^+)^H = AA^+$,

(iv) $(A^+A)^H = A^+A$.

Properties

(1) $A\ (m \times n)$: A^+ exists and is unique.

(2) $A\ (m \times n)$:

$$A = U^H \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V \text{ is the singular value decomposition of } A$$

$$\Rightarrow \quad A^+ = V^H \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U.$$

(3) $A\ (m \times n), U\ (m \times m), V\ (n \times n)$:

$$U, V \text{ unitary} \Rightarrow (UAV)^+ = V^H A^+ U^H.$$

(4) $A\ (m \times n)$, $\text{rk}(A) = r$:

$$B\ (m \times r),\ C\ (r \times n) \text{ such that } A = BC \Rightarrow A^+ = C^+B^+.$$

(5) $c \in \mathbb{C}$: $\quad c^+ = \begin{cases} c^{-1} & \text{if } c \neq 0 \\ 0 & \text{if } c = 0 \end{cases}.$

(6) $A\ (m \times n)$, $c \in \mathbb{C}$: $\quad (cA)^+ = c^+A^+$.

(7) $A\ (m \times n), B\ (r \times s)$: $\quad (A \otimes B)^+ = A^+ \otimes B^+$.

(8) $A\ (m \times m)$:

(a) A is nonsingular $\Rightarrow A^+ = A^{-1}$.

(b) A is Hermitian $\Rightarrow A^+$ is Hermitian.

(c) A is Hermitian and idempotent $\Rightarrow A^+ = A$.

(9) $A\ (m \times n)$:

(a) $(A^+)^+ = A$.

(b) $(A^H)^+ = (A^+)^H$.

(c) $A^H A A^+ = A^H$.

(d) $A^+ A A^H = A^H$.

(e) $A^H (A^+)^H A^+ = A^+$.

(f) $A^+ (A^+)^H A^H = A^+$.

(g) $(A^H A)^+ = A^+ (A^+)^H$.

(h) $(A A^H)^+ = (A^+)^H A^+$.

(i) $A(A^H A)^+ A^H A = A$.

(j) $A A^H (A A^H)^+ A = A$.

(k) $A^+ = (A^H A)^+ A^H = A^H (A A^H)^+$.

(10) $A\ (m \times n)$:

(a) $\mathrm{rk}(A) = m \iff AA^+ = I_m$.

(b) $\mathrm{rk}(A) = n \iff A^+ A = I_n$.

(c) $\mathrm{rk}(A) = n \Rightarrow A^+ = (A^H A)^{-1} A^H$.

(d) $\mathrm{rk}(A) = m \Rightarrow A^+ = A^H (A A^H)^{-1}$.

(e) $\mathrm{rk}(A) = n \Rightarrow (A A^H)^+ = A(A^H A)^{-2} A^H$.

(f) $\mathrm{rk}(A) = 1 \Rightarrow A^+ = [\mathrm{tr}(A A^H)]^{-1} A^H$.

(g) $A = O_{m \times n} \iff A^+ = O_{n \times m}$.

(11) $A\ (m \times n)$:

(a) $\mathrm{rk}(A^+) = \mathrm{rk}(A)$.

(b) $\mathrm{rk}(AA^+) = \mathrm{rk}(A^+ A) = \mathrm{rk}(A)$.

(c) $\mathrm{tr}(AA^+) = \mathrm{rk}(A)$.

(12) $A\ (m \times n)$:

(a) AA^+ and $A^+ A$ are idempotent.

(b) $I_m - AA^+$ and $I_n - A^+ A$ are idempotent.

(13) $A\ (m \times n),\ B\ (n \times r)$: $\quad AB = O_{m \times r} \iff B^+ A^+ = O_{r \times m}$.

(14) $A\ (m \times n),\ B\ (m \times r)$: $\quad A^H B = O_{n \times r} \iff A^+ B = O_{n \times r}$.

(15) $A, B\ (m \times n),\ AB^H = 0$:

$$\begin{aligned}&(A+B)^+ \\ &= A^+ + (I_n - A^+ B)[C^+ + (I_n - C^+ C) M B^H (A^+)^H A^+ (I_m - BC^+)],\end{aligned}$$

where $C = (I_m - AA^+)B$ and

$$M = [I_n + (I_n - C^+ C) B^H (A^+)^H A^+ B (I_n - C^+ C)]^{-1}.$$

(16) $A\ (m \times n)$, $B\ (n \times r)$, $C\ (m \times r)$:

$$A^H AB = A^H C \iff AB = AA^+C.$$

(17) $A\ (m \times n)$, $B\ (n \times r)$: $\det(BB^H) \neq 0 \Rightarrow AB(AB)^+ = AA^+$.

(18) $A\ (m \times m)$ Hermitian idempotent, $B\ (m \times n)$:

(a) $AB = B \Rightarrow A - BB^+$ is Hermitian idempotent with $\mathrm{rk}(A - BB^+) = \mathrm{rk}(A) - \mathrm{rk}(B)$.

(b) $AB = 0$ and $\mathrm{rk}(A) + \mathrm{rk}(B) = m \Rightarrow A = I_m - BB^+$.

(19) $\Omega\ (m \times m)$ Hermitian positive definite, $A\ (m \times n)$:

$$A^H\Omega^{-1}A(A^H\Omega^{-1}A)^+A^H = A^H.$$

(20) $A\ (m \times m)$ Hermitian: $\lambda \neq 0$ is eigenvalue of A with associated eigenvector $v \Rightarrow \lambda^{-1}$ is eigenvalue of A^+ with associated eigenvector v.

Partitioned Matrices

(21) $A_i\ (m_i \times n_i)$, $i = 1, \ldots, r$:

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}^+ = \begin{bmatrix} A_1^+ & & 0 \\ & \ddots & \\ 0 & & A_r^+ \end{bmatrix}.$$

(22) $A\ (m \times n)$:

$$\begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix}^+ = \begin{bmatrix} 0 & 0 \\ 0 & A^+ \end{bmatrix}, \quad \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} A^+ & 0 \\ 0 & 0 \end{bmatrix}.$$

(23) $A\ (m \times n)$, $B\ (m \times p)$:

$$[A : B]^+ = \begin{bmatrix} A^+ - A^+B(C^+ + D) \\ C^+ + D \end{bmatrix},$$

where $C = (I_m - AA^+)B$ and

$$\begin{aligned} D &= (I_p - C^+C)[I_p + (I_p - C^+C)B^H(A^+)^H A^+B(I_p - C^+C)]^{-1} \\ &\quad \times B^H(A^+)^H A^+(I_m - BC^+). \end{aligned}$$

(24) $A\ (m \times n)$, $B\ (p \times n)$:

$$\begin{bmatrix} A \\ B \end{bmatrix}^+ = [A^+ - TBA^+ : T],$$

where $T = E^+ + (I_n - E^+B)A^+(A^+)^H B^H K(I_p - EE^+)$ with $E = B(I_n - A^+A)$ and $K = [I_p + (I_p - EE^+)BA^+(A^+)^H B^H (I_p - EE^+)]^{-1}$.

Note: These results on Moore–Penrose generalized inverses are also contained in books on generalized inverses such as Rao & Mitra (1971), Pringle & Rayner (1971), Boullion & Odell (1971), Ben-Israel & Greville (1974). A good collection of results is also contained in Magnus & Neudecker (1988), including many of those given here. Many of the results follow easily by verifying the defining properties of a Moore–Penrose inverse.

3.7 Matrix Powers

Definition: For $i \in \mathbb{Z}$, the ith power of the $(m \times m)$ matrix A, denoted by A^i, is defined as follows:

$$A^i = \begin{cases} \prod_{j=1}^{i} A & \text{for positive integers } i \\ I_m & \text{for } i = 0 \\ \left(\prod_{j=1}^{-i} A\right)^{-1} & \text{for negative integers } i, \text{ if } \det(A) \neq 0. \end{cases}$$

If A can be written as

$$A = U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} U^H$$

for some unitary matrix U (see Chapter 6) then the power of A is defined for any $\alpha \in \mathbb{R}, \alpha > 0$, as follows:

$$A^\alpha \equiv U \begin{bmatrix} \lambda_1^\alpha & & 0 \\ & \ddots & \\ 0 & & \lambda_m^\alpha \end{bmatrix} U^H.$$

This definition applies, for instance, for Hermitian matrices. The definitions are equivalent for integer values of α.

Properties

(1) $A\ (m \times m), c \in \mathbb{C}, i \in \mathbb{N}: \quad (cA)^i = c^i A^i.$

(2) (Binomial formula)
$A, B\ (m \times m), i \in \mathbb{N}, i \geq 1:$

$$(A + B)^i = \sum_{j=0}^{i} \sum A^{k_1} B A^{k_2} B \cdots A^{k_j} B A^{k_{j+1}},$$

where the second sum is taken over all $k_1, \ldots, k_{j+1} \in \{0, \ldots, i\}$ with $k_1 + \cdots + k_{j+1} = i - j$.

(3) (Binomial formula)
$A, B\ (m \times m): \quad (A+B)^2 = A^2 + AB + BA + B^2.$

(4) (Binomial formula for commuting matrices)
$A, B\ (m \times m), i \in \mathbb{N}, i \geq 1:$

$$AB = BA \Rightarrow (A+B)^i = \sum_{j=0}^{i} \binom{i}{j} A^j B^{i-j}.$$

(5) $A, B\ (m \times m), i \in \mathbb{N}, i \geq 1:$

$$A^i - B^i = \sum_{j=0}^{i-1} A^j (A-B) B^{i-1-j}.$$

(6) $A\ (m \times m), B\ (n \times n), i \in \mathbb{N}:$

(a) $(A \otimes B)^i = A^i \otimes B^i.$

(b) $(A \oplus B)^i = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^i = \begin{bmatrix} A^i & 0 \\ 0 & B^i \end{bmatrix} = A^i \oplus B^i.$

(7) $A\ (m \times m), i \in \mathbb{N}:$

(a) $|A^i|_{\text{abs}} \leq |A|^i_{\text{abs}}.$
(b) $(A^i)' = (A')^i.$
(c) $(A^i)^H = (A^H)^i.$
(d) $\overline{A^i} = \bar{A}^i.$
(e) $(A^i)^{-1} = (A^{-1})^i.$
(f) $\text{rk}(A^i) \leq \text{rk}(A).$
(g) $\det(A^i) = (\det A)^i.$

(8) $A\ (m \times m), i \in \mathbb{N}$ even : $\quad \text{tr}(A^i) = \text{vec}(A^{i/2'})'\text{vec}(A^{i/2}).$

(9) $I_m\ (m \times m)$ identity matrix, $i \in \mathbb{N}: \quad I_m^i = I_m.$

(10) $i \in \mathbb{N}, i \neq 0: \quad O_{m\times m}^i = O_{m\times m}.$

(11) $i \in \mathbb{N}$:

$$\underset{(m \times n)}{\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & & \lambda & 1 \\ 0 & 0 & \dots & & \lambda \end{bmatrix}^i} = \begin{bmatrix} \lambda^i & \binom{i}{1}\lambda^{i-1} & \dots & \binom{i}{m-1}\lambda^{i-m+1} \\ 0 & \lambda^i & \dots & \binom{i}{m-2}\lambda^{i-m+2} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \binom{i}{1}\lambda^{i-1} \\ 0 & 0 & \dots & \lambda^i \end{bmatrix}$$

(12) $A\ (m \times m)$:

(a) A is idempotent $\Rightarrow$ $A^i = A$ for $i = 1, 2, \dots$

(b) A is nilpotent $\iff$ $A^i = O_{m \times m}$ for some $i > 0$.

(c) A is symmetric $\Rightarrow$ A^i is symmetric for $i = 1, 2, \dots$

(d) A is Hermitian $\Rightarrow$ A^i is Hermitian for $i = 1, 2, \dots$

(13) $A\ (m \times m), i \in \mathbb{N}$:

$$\begin{aligned} \text{vec}(A^i) &= (I_m \otimes A^i)\text{vec}(I_m) \\ &= ((A^i)' \otimes I_m)\text{vec}(I_m) \\ &= ((A^{i/2}) \otimes A^{i/2})\text{vec}(I_m), \quad \text{if } A \text{ is positive definite} \\ &= ((A^{i/2})' \otimes A^{i/2})\text{vec}(I_m), \quad \text{if } i \text{ is even} \\ &= (A' \otimes A)\text{vec}(A^{i-2}), \quad \text{if } i \geq 2. \end{aligned}$$

(14) $A\ (m \times m)$:
$A^i \to_{i \to \infty} 0 \iff$ all eigenvalues of A have modulus less than 1.

Note: Most results of this section are basic and follow immediately from definitions. The binomial formulae may be found in Johansen (1995, Section A.2) and (13) is a consequence of an important relation between Kronecker products and the vec operator (see Section 2.4).

3.8 The Absolute Value

Definition: Given an $(m \times n)$ matrix $A = [a_{ij}]$ its absolute value or modulus is

$$|A|_{\text{abs}} \equiv [\ |a_{ij}|_{\text{abs}}\ ] = \begin{bmatrix} |a_{11}|_{\text{abs}} & |a_{12}|_{\text{abs}} & \cdots & |a_{1n}|_{\text{abs}} \\ |a_{21}|_{\text{abs}} & |a_{22}|_{\text{abs}} & \cdots & |a_{2n}|_{\text{abs}} \\ \vdots & \vdots & & \vdots \\ |a_{m1}|_{\text{abs}} & |a_{m2}|_{\text{abs}} & \cdots & |a_{mn}|_{\text{abs}} \end{bmatrix} \quad (m \times n)$$

where the modulus of a complex number $c = c_1 + ic_2$ is defined as $|c|_{\text{abs}} = \sqrt{c_1^2 + c_2^2} = \sqrt{c\bar{c}}$. Here $\bar{c}$ is the complex conjugate of c.

(1) $A\ (m \times n)$:

(a) $|A|_{\text{abs}} \geq O_{m\times n}$.

(b) $|A|_{\text{abs}} = O_{m\times n} \iff A = O_{m\times n}$.

(2) $A\ (m \times n), c \in \mathbb{C}$: $\quad |cA|_{\text{abs}} = |c|_{\text{abs}}\,|A|_{\text{abs}}$.

(3) $A, B\ (m \times n)$: $\quad |A + B|_{\text{abs}} \leq |A|_{\text{abs}} + |B|_{\text{abs}}$.

(4) $A\ (m \times n), B\ (n \times p)$: $\quad |AB|_{\text{abs}} \leq |A|_{\text{abs}}\,|B|_{\text{abs}}$.

(5) $A\ (m \times m), i \in \mathbb{N}$: $\quad |A^i|_{\text{abs}} \leq |A|_{\text{abs}}^i$.

(6) $A\ (m \times n), B\ (p \times q)$: $\quad |A \otimes B|_{\text{abs}} = |A|_{\text{abs}} \otimes |B|_{\text{abs}}$.

(7) $A\ (m \times m), B\ (n \times n)$: $\quad |A \oplus B|_{\text{abs}} = |A|_{\text{abs}} \oplus |B|_{\text{abs}}$.

(8) $A\ (m \times n)$:

(a) $|A'|_{\text{abs}} = |A|'_{\text{abs}}$.

(b) $|\bar{A}|_{\text{abs}} = |A|_{\text{abs}}$.

(c) $|A^H|_{\text{abs}} = |A|'_{\text{abs}}$.

(9) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p)$:

$$\left|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right|_{\text{abs}} = \begin{bmatrix} |A|_{\text{abs}} & |B|_{\text{abs}} \\ |C|_{\text{abs}} & |D|_{\text{abs}} \end{bmatrix}.$$

Note: These results may be found in Horn & Johnson (1985, Chapter 8) or follow easily from the definition of the absolute value.

4

Trace, Determinant and Rank of a Matrix

4.1 The Trace

Definition: The trace of an $(m \times m)$ matrix $A = [a_{ij}]$ is defined as

$$\mathrm{tr}A = \mathrm{tr}(A) \equiv a_{11} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}.$$

4.1.1 General Results

(1) $A, B\ (m \times m)$: $\quad \mathrm{tr}(A \pm B) = \mathrm{tr}(A) \pm \mathrm{tr}(B)$.

(2) $A\ (m \times m), c \in \mathbb{C}$: $\quad \mathrm{tr}(cA) = c\ \mathrm{tr}(A)$.

(3) $A, B\ (m \times m),\ c_1, c_2 \in \mathbb{C}$: $\quad \mathrm{tr}(c_1 A \pm c_2 B) = c_1 \mathrm{tr}(A) \pm c_2 \mathrm{tr}(B)$.

(4) $A\ (m \times m)$:

(a) $\mathrm{tr}(A') = \mathrm{tr}(A)$.

(b) $\mathrm{tr}(\bar{A}) = \overline{\mathrm{tr}(A)}$.

(c) $\mathrm{tr}(A^H) = \overline{\mathrm{tr}(A)}$.

(5) $A\ (m \times n)$:

(a) $\mathrm{tr}(AA^+) = \mathrm{rk}(A)$.

(b) $\mathrm{tr}(A^H A) = 0 \iff A = O_{m \times n}$.

(6) $A\ (m \times m)$: $\quad A$ idempotent $\Rightarrow \mathrm{tr}(A) = \mathrm{rk}(A)$.

(7) $A\ (m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$: $\quad \mathrm{tr}(A) = \lambda_1 + \cdots + \lambda_m$.

(8) $A\ (m \times n), B\ (n \times m)$:

(a) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

(b) $\mathrm{tr}(AB) = \mathrm{vec}(A')'\mathrm{vec}(B) = \mathrm{vec}(B')'\mathrm{vec}(A)$.

(9) $A\ (m \times n), B\ (n \times p), C\ (p \times q), D\ (q \times m)$:

$$\begin{aligned}\mathrm{tr}(ABCD) &= \mathrm{vec}(D')'(C' \otimes A)\mathrm{vec}(B) \\ &= \mathrm{vec}(A')'(D' \otimes B)\mathrm{vec}(C) \\ &= \mathrm{vec}(B')'(A' \otimes C)\mathrm{vec}(D) \\ &= \mathrm{vec}(C')'(B' \otimes D)\mathrm{vec}(A).\end{aligned}$$

(10) $A, B\ (m \times m)$: $\quad B$ nonsingular $\Rightarrow \mathrm{tr}(BAB^{-1}) = \mathrm{tr}(A)$.

(11) $A, B\ (m \times n), j_k = (1, \ldots, 1)'\ (k \times 1)$: $\quad \mathrm{tr}(AB') = j_m'(A \odot B)j_n$.

(12) $A, B, D\ (m \times m),\ j = (1, \ldots, 1)'\ (m \times 1)$:

$$D \text{ diagonal} \quad \Rightarrow \quad \mathrm{tr}(ADB'D) = j'D'(A \odot B)Dj.$$

(13) $A, B, C\ (m \times n)$: $\quad \mathrm{tr}[A'(B \odot C)] = \mathrm{tr}[(A' \odot B')C]$.

(14) $A\ (m \times m)$: $\quad \mathrm{tr}(A \odot I_m) = \mathrm{tr}(A)$.

(15) $A\ (m \times m), B\ (n \times n)$:

(a) $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B)$.

(b) $\mathrm{tr}(A \oplus B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

(16) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$\mathrm{tr}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \mathrm{tr}(A) + \mathrm{tr}(D).$$

(17) $K_{mm}\ (m^2 \times m^2)$ commutation matrix : $\quad \mathrm{tr}(K_{mm}) = m$.

(18) $D_m\ (m^2 \times \frac{1}{2}m(m \times 1))$ duplication matrix :

(a) $\mathrm{tr}(D_m'D_m) = \mathrm{tr}(D_mD_m') = m^2$.

(b) $\mathrm{tr}(D_m'D_m)^{-1} = m(m+3)/4$.

(19) $L_m\ (\frac{1}{2}m(m+1) \times m^2)$ elimination matrix :

$$\mathrm{tr}(L_mL_m') = \mathrm{tr}(L_m'L_m) = \tfrac{1}{2}m(m+1).$$

(20) $A = [a_{ij}], B\ (m \times m)$ real positive semidefinite, $\alpha \in \mathbb{R}, \alpha > 0, \alpha \neq 1$:

$$\mathrm{tr}(A^\alpha) = \sum_{i=1}^{m} a_{ii}^\alpha \iff A \text{ is diagonal}.$$

(21) $A, B \neq 0\ (m \times m)$ real positive semidefinite, $\alpha \in \mathbb{R}, 0 < \alpha < 1$:

$$\mathrm{tr}(A^\alpha B^{1-\alpha}) = (\mathrm{tr}\,A)^\alpha(\mathrm{tr}\,B)^{1-\alpha} \iff B = cA \text{ for some } c \in \mathbb{R}, c > 0.$$

(22) A, B $(m \times m)$ real positive semidefinite, $\alpha \in \mathbb{R}, \alpha > 1$:

$$[\mathrm{tr}(A+B)^\alpha]^{1/\alpha} = (\mathrm{tr}\, A^\alpha)^{1/\alpha} + (\mathrm{tr}\, B^\alpha)^{1/\alpha}$$
$$\iff B = cA \text{ for some } c \in \mathbb{R}, c > 0.$$

Note: The rules involving the vec operator, the commutation, duplication and elimination matrices are given in Magnus & Neudecker (1988) and Magnus (1988). (20) – (22) are given in Magnus & Neudecker (1988, Chapter 11). The other rules follow from basic principles.

4.1.2 Inequalities Involving the Trace

In this subsection all matrices are real unless otherwise stated.

(1) A $(m \times n)$ complex:

(a) $|\mathrm{tr}\, A|_{\mathrm{abs}} \leq \mathrm{tr}\, |A|_{\mathrm{abs}}$.

(b) $\mathrm{tr}(A^H A) = \mathrm{tr}(AA^H) \geq 0$.

(2) A, B $(m \times n)$:

(a) (Cauchy–Schwarz inequality)
$\mathrm{tr}(A'B)^2 \leq \mathrm{tr}(A'A)\mathrm{tr}(B'B)$.

(b) $\mathrm{tr}(A'B)^2 \leq \mathrm{tr}(A'AB'B)$.

(c) $\mathrm{tr}(A'B)^2 \leq \mathrm{tr}(AA'BB')$.

(3) (Schur's inequality)
A $(m \times m)$: $\quad \mathrm{tr}(A^2) \leq \mathrm{tr}(A'A)$.

(4) A $(m \times m)$ positive semidefinite: $\quad (\det A)^{1/m} \leq \frac{1}{m}\mathrm{tr}(A)$.

(5) A $(m \times m)$:
All eigenvalues of A are real $\Rightarrow$ $|\frac{1}{m}\mathrm{tr}(A)|_{\mathrm{abs}} \leq [\frac{1}{m}\mathrm{tr}(A^2)]^{1/2}$.

(6) A $(m \times m)$ symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, X $(m \times n)$ with $X'X = I_n$:

$$\sum_{i=1}^{n} \lambda_i \leq \mathrm{tr}(X'AX) \leq \sum_{i=1}^{n} \lambda_{m-n+i}.$$

(7) $A = [a_{ij}]$ $(m \times m)$ positive semidefinite:

$$\mathrm{tr}(A^\alpha) \begin{cases} \geq \sum_{i=1}^m a_{ii}^\alpha & \text{for} \quad \alpha > 1 \\ \leq \sum_{i=1}^m a_{ii}^\alpha & \text{for} \quad 0 < \alpha < 1. \end{cases}$$

(8) $A, B \neq 0$ $(m \times m)$ positive semidefinite, $\alpha \in \mathbb{R}$:

(a) (Hölder's inequality)
$0 < \alpha < 1 \;\Rightarrow\; \mathrm{tr}(A^\alpha B^{1-\alpha}) \leq (\mathrm{tr}\, A)^\alpha (\mathrm{tr}\, B)^{1-\alpha}$.

(b) $\alpha > 1, \beta = \alpha/(\alpha - 1) \;\Rightarrow\; \mathrm{tr}(AB) \leq (\mathrm{tr}\, A^\alpha)^{1/\alpha} (\mathrm{tr}\, B^\beta)^{1/\beta}$.

(c) (Minkowski's inequality)
$\alpha > 1 \;\Rightarrow\; [\mathrm{tr}(A + B)^\alpha]^{1/\alpha} \leq (\mathrm{tr}\, A^\alpha)^{1/\alpha} + (\mathrm{tr}\, B^\alpha)^{1/\alpha}$.

(9) A, B $(m \times m)$ positive semidefinite:

(a) $\alpha \in \mathbb{R}, 0 < \alpha < 1 \;\Rightarrow\; \mathrm{tr}(A^\alpha B^{1-\alpha}) \leq \mathrm{tr}(\alpha A + (1 - \alpha) B)$.

(b) $\frac{1}{m}\mathrm{tr}(AB) \geq (\det A)^{1/m} (\det B)^{1/m}$.

(10) A $(m \times m)$ positive definite: $\ln \det(A) \leq \mathrm{tr}(A) - m$.

(11) A $(m \times m)$ positive definite, B $(n \times m), C$ $(m \times n)$:

$$BC = I_n \;\Rightarrow\; \mathrm{tr}(C'AC) \geq \mathrm{tr}(BA^{-1}B')^{-1}.$$

(12) A $(m \times m)$ positive semidefinite with maximum eigenvalue $\lambda_{max}(A)$. B $(n \times m)$:

$$\mathrm{tr}(BAB') \leq \lambda_{max}(A)\mathrm{tr}(BB').$$

(13) A $(m \times m)$ positive semidefinite with maximum eigenvalue $\lambda_{max}(A)$:

$$\mathrm{tr}(A^2) \leq \lambda_{max}(A)\mathrm{tr}(A).$$

(14) A, B $(m \times m)$: $\mathrm{tr}[(A + B)(A + B)'] \leq 2[\mathrm{tr}(AA') + \mathrm{tr}(BB')]$.

(15) A $(m \times m)$ positive definite, B $(m \times m)$ positive semidefinite:

$$\exp[\mathrm{tr}(A^{-1}B)] \geq \frac{\det(A + B)}{\det(A)}.$$

(16) A, B $(m \times m)$ positive semidefinite:

(a) $\mathrm{tr}(AB) \leq \mathrm{tr}(A \otimes B)$.

(b) $\mathrm{tr}(AB) \leq \frac{1}{4}(\mathrm{tr}\, A + \mathrm{tr}\, B)^2$.

(c) $\mathrm{tr}(A \otimes B) \leq \frac{1}{4}(\mathrm{tr}\, A + \mathrm{tr}\, B)^2$.

(d) $\mathrm{tr}(A \odot B) \geq 0$.

(e) $\mathrm{tr}(A \odot B) \leq \mathrm{tr}(A \otimes B)$.

(17) A $(m \times m)$, B $(n \times n)$ positive semidefinite: $\mathrm{tr}(A \otimes B) \geq 0$.

(18) A, B, C, D $(m \times m)$ positive semidefinite: $C - A$, $D - B$ positive semidefinite $\Rightarrow \mathrm{tr}(AB) \leq \mathrm{tr}(CD)$.

(19) A, B $(m \times m)$ symmetric: $\mathrm{tr}(AB) \leq \frac{1}{2}\mathrm{tr}(A^2 + B^2)$.

(20) A, B $(m \times m)$:

$$\mathrm{tr}(A \otimes B) \leq \tfrac{1}{2}\mathrm{tr}(A \otimes A + B \otimes B),$$

$$\operatorname{tr}(A \odot B) \leq \tfrac{1}{2}\operatorname{tr}(A \odot A + B \odot B).$$

Note: Inequalities (16) – (20) are from Neudecker & Shuangzhe (1993) (see also Neudecker & Shuangzhe (1995)). The other inequalities can be found in Chapter 11 of Magnus & Neudecker (1988).

4.1.3 Optimization of Functions Involving the Trace

In this subsection again all matrices are real if not otherwise stated.

(1) Ω $(m \times m)$ real symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\operatorname{tr}(B'\Omega B) \;:\; B\ (m \times n)\text{ real}, B'B = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\max\{\operatorname{tr}(B'\Omega B) \;:\; B\ (m \times n)\text{ real}, B'B = I_n\} = \lambda_m + \cdots + \lambda_{m-n+1}.$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(2) Ω $(m \times m)$ complex Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\operatorname{tr}(B^H\Omega B) \;:\; B\ (m \times n)\text{ complex}, B^H B = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\begin{aligned}\max\{\operatorname{tr}(B^H\Omega B) \;:\; B\ (m \times n)\text{ complex}, B^H B = I_n\}\\ = \lambda_m + \cdots + \lambda_{m-n+1}.\end{aligned}$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(3) Ω $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$, $0 \leq r < m$:

$$\begin{aligned}\min\{\operatorname{tr}(\Omega - A)^2 \;:\; A\ (m \times m)\text{ positive semidefinite}, \operatorname{rk}(A) = r\}\\ = \lambda_1^2 + \cdots + \lambda_{m-r}^2.\end{aligned}$$

The minimum is attained for

$$A = \sum_{i=m-r+1}^{m} \lambda_i v_i v_i'.$$

(4) X $(m \times n)$, $\lambda_1 \leq \cdots \leq \lambda_n$ eigenvalues of $X'X$ with associated orthonormal eigenvectors $v_1, \ldots, v_n$, $0 \leq r < n$:

$$\min\{\mathrm{tr}[(X - AB)(X - AB)'] : A\ (m \times r), B\ (r \times n) \text{ with } BB' = I_r\} \\ = \lambda_1 + \lambda_2 + \cdots + \lambda_{m-r}.$$

The minimum is attained for

$$B = [v_n, \ldots, v_{n-r+1}]' \quad \text{and} \quad A = XB'.$$

(5) Y $(m \times n), X$ $(p \times n)$ with $\mathrm{rk}(X) = p, \Omega$ $(m \times m)$ positive definite, $\lambda_1 \leq \cdots \leq \lambda_m$ eigenvalues of $\Omega^{1/2} YX'(XX')^{-1}XY'\Omega^{1/2'}$ with associated orthonormal eigenvectors $v_1, \ldots, v_m$, $1 \leq r \leq m \leq p$:

$$\min\{\mathrm{tr}[(Y - ABX)'\Omega(Y - ABX)] : \\ A\ (m \times r), B\ (r \times p), \mathrm{rk}(A) = \mathrm{rk}(B) = r\} \\ = \mathrm{tr}[(Y - \hat{A}\hat{B}X)'\Omega(Y - \hat{A}\hat{B}X)],$$

where

$$\hat{A} = \Omega^{-1/2}[v_m, \ldots, v_{m-r+1}]$$

and

$$\hat{B} = [v_m, \ldots, v_{m-r+1}]'\Omega^{1/2} YX'(XX')^{-1}.$$

(6) Y $(m \times n)$, X $(p \times n), \mathrm{rk}(X) = p$, Ω $(m \times m)$ positive definite:

$$\min\{\mathrm{tr}[(Y - AX)'\Omega(Y - AX)] \ : \ A\ (m \times p)\} \\ = \mathrm{tr}(\Omega YY' - \Omega YX'(XX')^{-1}XY').$$

The minimum is attained for $A = YX'(XX')^{-1}$.

(7) Y $(m \times n)$, X $(p \times n)$ with $\mathrm{rk}(X) = p$, Ω $(m \times m)$ positive definite, R $(q \times m)$ with $\mathrm{rk}(R) = q$, C $(q \times p)$:

$$\min\{\mathrm{tr}[(Y - AX)'\Omega(Y - AX)] \ : \ A\ (m \times p),\ RA = C\} \\ = \mathrm{tr}[(Y - \hat{A}X)'\Omega(Y - \hat{A}X)],$$

where

$$\hat{A} = YX'(XX')^{-1} + \Omega^{-1}R'(R\Omega^{-1}R')^{-1}(C - RYX'(XX')^{-1}).$$

(8) B, C $(m \times m)$ positive definite, B diagonal, $\lambda_1 \leq \cdots \leq \lambda_m$ eigenvalues of $B^{-1/2}CB^{-1/2}$ with associated orthonormal eigenvectors $v_1, \ldots, v_m$, $1 \leq r \leq m$:

$$\min\{\ln\det(AA' + B) + \mathrm{tr}[(AA' + B)^{-1}C] \ : \ A\ (m \times r)\} \\ = m + \ln\det(C) + \textstyle\sum_{i=1}^{m-r}(\lambda_i - \ln\lambda_i - 1).$$

The minimum is attained for

$$A = B^{1/2}[v_m, \ldots, v_{m-r+1}]\left(\begin{bmatrix} \lambda_m & & 0 \\ & \ddots & \\ 0 & & \lambda_{m-r+1} \end{bmatrix} - I_r\right)^{1/2}.$$

(9) B $(m \times m)$ positive definite:

$$\min\{\operatorname{tr}(B^{-1}A) - \ln|\det(B^{-1}A)|_{\mathrm{abs}} : \\ A\ (m \times m) \text{ positive semidefinite}\} = m.$$

The minimum is attained for $A = B$.

Note: The results of this subsection are partly given in Magnus & Neudecker (1988, Chapter 17), Horn & Johnson (1985, Chapter 4) and Lütkepohl (1991, Section A.14). Result (9) is from Johansen (1995, Section A.1).

4.2 The Determinant

Definition: The determinant of the $(m \times m)$ matrix $A = [a_{ij}]$ is defined as

$$\det A = \det(A) \equiv \sum(-1)^p a_{1i_1} a_{2i_2} \times \cdots \times a_{mi_m},$$

where the sum is taken over all products consisting of precisely one element from each row and each column of A multiplied by -1 or 1, if the permutation $i_1, \ldots, i_m$ is odd or even, respectively.

4.2.1 General Results

(1) $\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$

(2) $A = [a_{ij}]$ $(m \times m)$, $m > 2$:

$$\begin{aligned} \det(A) &= a_{i1} \operatorname{cof}(a_{i1}) + \cdots + a_{im} \operatorname{cof}(a_{im}) \\ &= a_{1j} \operatorname{cof}(a_{1j}) + \cdots + a_{mj} \operatorname{cof}(a_{mj}) \end{aligned}$$

for any $i, j \in \{1, \ldots, m\}$. Here $\operatorname{cof}(a_{ij})$ denotes the cofacor of a_{ij} (see Section 1.4).

(3) A $(m \times m), c \in \mathbb{C}$: $\quad \det(cA) = c^m \det(A)$.

(4) A, B $(m \times m)$:

(a) $\det(AB) = \det(A)\det(B)$.

(b) B is nonsingular $\Rightarrow$ $\det(BAB^{-1}) = \det(A)$.

(5) I_m $(m \times m)$ identity matrix: $\det(I_m) = 1$.

(6) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:

$$\det(A) = \lambda_1 \cdots \lambda_m = \prod_{i=1}^{m} \lambda_i .$$

(7) A $(m \times m)$:

(a) $\det(A') = \det(A)$.

(b) $\det(A^H) = \overline{\det(A)}$.

(c) $\det(\bar{A}) = \overline{\det(A)}$.

(d) $\det(A^{-1}) = (\det A)^{-1}$, if A is nonsingular.

(e) $\det(A^{adj}) = (\det A)^{m-1}$.

(f) $\det(A) I_m = A^{adj} A = A A^{adj}$.

(8) A $(m \times m)$, B $(n \times n)$:

(a) $\det(A \otimes B) = (\det A)^n (\det B)^m$.

(b) $\det(A \oplus B) = \det(A)\det(B)$.

(9) $A = [a_{ij}]$ $(m \times m)$:

(a) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm})$ $\Rightarrow$ $\det(A) = a_{11} \cdots a_{mm} = \prod_{i=1}^{m} a_{ii}$.

(b) A is triangular $\Rightarrow$ $\det(A) = a_{11} \cdots a_{mm} = \prod_{i=1}^{m} a_{ii}$.

(10) A $(m \times n)$ real : $\det(I_m + AA') = \det(I_n + A'A)$.

(11) A $(m \times m)$:

(a) $\mathrm{rk}(A) < m \iff \det(A) = 0$.

(b) $\mathrm{rk}(A) = m \iff \det(A) \neq 0$.

(c) A is singular $\iff \det(A) = 0$.

(d) The rows of A are linearly independent $\iff \det(A) \neq 0$.

(e) The columns of A are linearly independent $\iff \det(A) \neq 0$.

(f) A has a row or column of zeros $\Rightarrow$ $\det(A) = 0$.

(g) A has two identical rows or columns $\Rightarrow$ $\det(A) = 0$.

(12) A, B $(m \times m)$:

(a) B is obtained from A by adding to one row (column) a scalar multiple of another row (column) $\Rightarrow$ $\det(A) = \det(B)$.

(b) B is obtained from A by interchanging two rows or columns $\Rightarrow$ $\det(B) = -\det(A)$.

(13) (Binet-Cauchy formula)
$A, B\ (m \times n), m \leq n$: $\det(AB') = \sum \det(A_m)\det(B'_m)$, where the sum is taken over all $(m \times m)$ submatrices A_m of A and B_m are the corresponding submatrices of B.

(14) A $(m \times m)$ positive definite, B $(m \times m)$ positive semidefinite:

$$\det(A + B) = \det(A) \iff B = 0.$$

(15) $A = [a_{ij}]$ $(m \times m)$ positive definite:

$$\det(A) = \prod_{i=1}^{m} a_{ii} \iff A = \operatorname{diag}(a_{11}, \ldots, a_{mm}).$$

(16) (Vandermonde determinant)
$\lambda_1, \ldots, \lambda_m \in \mathbb{C}$:

$$\det \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_m \\ \vdots & \vdots & & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \ldots & \lambda_m^{m-1} \end{bmatrix} = \prod_{j<i} (\lambda_i - \lambda_j).$$

(17) $p_0, \ldots, p_{m-1} \in \mathbb{C}$:

$$\det \begin{bmatrix} 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & \cdots & -p_{m-2} & -p_{m-1} \end{bmatrix} = (-1)^m p_0.$$

Note: Most results of this subsection are straightforward implications of the definitions and elementary matrix rules. They can be found in textbooks such as Lancaster & Tismenetsky (1985) and Barnett (1990).

4.2.2 Determinants of Partitioned Matrices

(1) A $(m \times n)$, B $(m \times p)$, $C = [A : B]$ $(m \times (n + p))$:

(a) $\det(C^H C) = \det(A^H A)\det[B^H(I_m - AA^+)B]$
$= \det(B^H B)\det[A^H(I_m - BB^+)A]$.

(b) $\mathrm{rk}(A) = n$
$\Rightarrow \det(C^H C) = \det(A^H A)\det[B^H B - B^H A(A^H A)^{-1} A^H B]$.

(c) $\mathrm{rk}(B) = p$
$\Rightarrow \det(C^H C) = \det(B^H B)\det[A^H A - A^H B(B^H B)^{-1} B^H A]$.

(2) $\det \begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix} = (-1)^m$.

(3) $A\ (m \times m), B\ (m \times n), C\ (n \times n)$:

$$\det \begin{bmatrix} A & B \\ O_{n \times m} & C \end{bmatrix} = \det(A)\det(C).$$

(4) $A\ (m \times m), B\ (n \times m), C\ (n \times n)$:

$$\det \begin{bmatrix} A & O_{m \times n} \\ B & C \end{bmatrix} = \det(A)\det(C).$$

(5) $A, B, C\ (m \times m)$:

$$\det \begin{bmatrix} A & B \\ C & O_{m \times m} \end{bmatrix} = (-1)^m \det(B)\det(C).$$

(6) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$A \text{ nonsingular} \quad \Rightarrow \quad \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B).$$

$$D \text{ nonsingular} \quad \Rightarrow \quad \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(D)\det(A - BD^{-1}C).$$

(7) $A\ (m \times m),\ a, b\ (m \times 1)$:

$$A \text{ nonsingular} \quad \Rightarrow \quad \det \begin{bmatrix} A & a \\ b^H & 0 \end{bmatrix} = -b^H A^{adj} a.$$

(8) $A\ (m \times m),\ a, b\ (m \times 1),\ c \in \mathbb{C}$: A nonsingular

$$\Rightarrow \quad \det \begin{bmatrix} A & a \\ b^H & c \end{bmatrix} = c\det(A) - b^H A^{adj} a = \det(A)(c - b^H A^{-1} a).$$

Note: Some of these results on partitioned matrices are given in Magnus (1988). Most results are straightforward implications of the definitions and elementary matrix rules, however.

4.2.3 Determinants Involving Duplication Matrices

Reminder: D_m denotes the $(m^2 \times \frac{1}{2}m(m+1))$ duplication matrix (see Section 9.5 for more details).

(1) $\det(D_m' D_m) = 2^{m(m-1)/2}$.

(2) $\det(D_m D_m') = 0$.

(3) A $(m \times m)$:

(a) $\det(D_m'(A \otimes A)D_m) = 2^{m(m-1)/2}\det(A)^{m+1}$.

(b) $\det(D_m^+(A \otimes A)D_m^{+\prime}) = 2^{-m(m-1)/2}\det(A)^{m+1}$.

(c) $\det(D_m^+(A \otimes A)D_m) = (\det A)^{m+1}$.

(4) A, B $(m \times m)$:

(a) $\det(B) = 0 \;\Rightarrow\; \det(D_m^+(A \otimes B)D_m) = 0$.

(b) $\det(B) \neq 0$ and $\lambda_1, \dots, \lambda_m$ are the eigenvalues of AB^{-1} $\Rightarrow$

$$\det[D_m^+(A \otimes B)D_m] = 2^{-m(m-1)/2}\det(A)(\det B)^m \prod_{i>j}(\lambda_i + \lambda_j).$$

(c) $\det(A) \neq 0$ and $\lambda_1, \dots, \lambda_m$ are the eigenvalues of BA^{-1}

$$\Rightarrow \det[D_m^+(A \otimes A \pm B \otimes B)D_m] = (\det A)^{m+1} \prod_{i \geq j}(1 \pm \lambda_i \lambda_j).$$

(5) A $(m \times m)$ with eigenvalues $\lambda_1, \dots, \lambda_m$:

$$\det[D_m^+(A \otimes I_m)D_m] = 2^{-m(m-1)/2}\det(A) \prod_{i>j}(\lambda_i + \lambda_j),$$

$$\det[D_m^+(I_m \otimes A + A \otimes I_m)D_m] = 2^m \det(A) \prod_{i>j}(\lambda_i + \lambda_j).$$

(6) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ lower triangular:

$$\det[D_m^+(A \otimes A \pm B \otimes B)D_m] = \prod_{i \geq j}(a_{ii}a_{jj} \pm b_{ii}b_{jj}).$$

(7) A $(m \times m)$ real, symmetric, nonsingular, $c \in \mathbb{R}$:

$$\det(D_m^+[A \otimes A + c\,\mathrm{vec}(A)\mathrm{vec}(A)']D_m) = (1 + cm)\det(A)^{m+1}.$$

(8) A $(m \times m)$ with eigenvalues $\lambda_1, \dots, \lambda_m$, $i \in \mathbb{N}, i > 1$:

$$\det\left(D_m^+ \sum_{j=0}^{i-1}(A^{i-1-j} \otimes A^j)D_m\right) = i^m(\det A)^{i-1} \prod_{k>l} \mu_{kl}$$

where

$$\mu_{kl} = \begin{cases} (\lambda_k^i - \lambda_l^i)/(\lambda_k - \lambda_l) & \text{if} \quad \lambda_k \neq \lambda_l \\ i\lambda_k^{i-1} & \text{if} \quad \lambda_k = \lambda_l \end{cases} .$$

Note: The rules presented in this subsection may be found in Magnus (1988).

4.2.4 Determinants Involving Elimination Matrices

Reminder: L_m denotes the $(\frac{1}{2}m(m+1)\times m^2)$ elimination matrix (see Section 9.6 for more details).

(1) $\det(L_m L_m') = 1$.

(2) $\det(L_m' L_m) = 0$.

(3) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ upper (lower) triangular:

$$\det(L_m(A \odot B)L_m') = \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}.$$

(4) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ lower triangular:

$$\det(L_m(A' \odot B)L_m') = \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}.$$

(5) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$:

$$\det(L_m(A \odot B)L_m') = \prod_{i=1}^{m} \det(A_{(i)})\det(B^{(i)}),$$

where

$$A_{(i)} = \begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix}, \quad B^{(i)} = \begin{bmatrix} b_{ii} & \dots & b_{im} \\ \vdots & \ddots & \vdots \\ b_{mi} & \dots & b_{mm} \end{bmatrix},$$

$i = 1, \dots, m$.

(6) A $(m \times m)$ real, lower triangular, nonsingular, $c \in \mathbb{R}$:

$$\det(L_m[A' \odot A + c\,\mathrm{vec}(A)\mathrm{vec}(A')']L_m') = (1 + cm)\det(A)^{m+1}.$$

(7) A, B $(m \times m)$ lower triangular:

$$\begin{aligned} \det(L_m(AB' \odot B'A)L_m') &= \det(L_m(AB' \odot A'B)L_m') \\ &= \det(L_m(AB \odot B'A)D_m) \\ &= \det(L_m(AB' \odot A'B')D_m) \\ &= (\det A)^{m+1}(\det B)^{m+1}. \end{aligned}$$

(8) $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$, $D = [d_{ij}]$ $(m \times m)$ lower triangular:

$$\det(L_m(AB' \otimes C'D)L_m') = \prod_{i=1}^{m} (c_{ii}d_{ii})^i (a_{ii}b_{ii})^{m-i+1},$$

$$\det[L_m(A' \otimes B + C' \otimes D)L_m'] = \prod_{i \geq j} (b_{ii}a_{jj} + d_{ii}c_{jj}).$$

(9) $A = [a_{ij}]$ $(m \times m)$ lower triangular, $n \in \mathbb{N}$:

$$\det\left(L_m\left(\sum_{j=0}^{n-1} (A')^{n-1-j} \otimes A^j\right) L_m'\right) = n^m (\det A)^{n-1} \prod_{k>l} \mu_{kl},$$

where

$$\mu_{kl} = \begin{cases} (a_{kk}^n - a_{ll}^n)/(a_{kk} - a_{ll}), & \text{if } a_{kk} \neq a_{ll} \\ n a_{kk}^{n-1}, & \text{if } a_{kk} = a_{ll} \end{cases}.$$

Note: The rules presented in this subsection may be found in Magnus (1988).

4.2.5 Determinants Involving Both Duplication and Elimination Matrices

(1) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ lower triangular:

$$\begin{aligned} \det(L_m(A \otimes B)D_m) &= \det(L_m(A' \otimes B')D_m) \\ &= \det(L_m(A \otimes B')D_m) \\ &= \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}. \end{aligned}$$

(2) A, B $(m \times m)$:

$$\det(L_m(A \otimes B)D_m) = \begin{cases} 0 & \text{if } \det(B) = 0 \\ \det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)}) & \text{if } \det(B) \neq 0 \end{cases}$$

where

$$C_{(n)} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}$$

denotes the nth principal submatrix of $AB^{-1} = [c_{ij}]$.

(3) $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]$ $(m \times m)$ lower triangular :

$$\det[L_m(A \otimes B'C)D_m] = \prod_{i=1}^{m} (b_{ii}c_{ii})^i a_{ii}^{m-i+1},$$

$$\det[L_m(AB' \otimes C')D_m] = \prod_{i=1}^{m} c_{ii}^i (a_{ii}b_{ii})^{m-i+1}.$$

(4) A, B $(m \times m)$:

$$\det(L_m(A \otimes B)D_m^{+\prime})$$
$$= \begin{cases} 0 & \text{if } \det(B) = 0 \\ 2^{-m(m-1)/2}\det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)}) & \text{if } \det(B) \neq 0 \end{cases}$$

where

$$C_{(n)} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{bmatrix}$$

are the principal submatrices of $C = [c_{ij}] = AB^{-1}$.

Note: The rules presented in this subsection are given in Magnus (1988).

4.2.6 Inequalities Related to Determinants

(1) (Cauchy-Schwarz inequality)
A, B $(m \times n)$: $(\det A^H B)^2 \leq \det(A^H A)\det(B^H B)$.

(2) (Hadamard's inequality)

$$A = [a_{ij}] \ (m \times m) : \quad (\det A)^2 \leq \prod_{i=1}^{m} \left(\sum_{j=1}^{m} |a_{ij}|_{\text{abs}}^2 \right).$$

(3) (Hadamard's inequality)

$$A = [a_{ij}] \ (m \times m) \text{ positive semidefinite:} \quad \det(A) \leq \prod_{i=1}^{m} a_{ii}.$$

(4) $A = [a_{ij}]$ $(m \times m)$: $|\det(A)|_{\text{abs}} \leq \{\max_{i,j} |a_{ij}|_{\text{abs}}\}^m m^{m/2}$.

(5) (Fischer's inequality)
A $(m \times m)$, B $(m \times n)$, C $(n \times n)$:

$$D = \begin{bmatrix} A & B \\ B^H & C \end{bmatrix} \text{positive definite} \quad \Rightarrow \quad \det(D) \leq \det(A)\det(C).$$

(6) $A = [a_{ij}]$ $(m \times m)$ positive definite:

$$\det \begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix} > 0, \quad i = 1, \dots, m.$$

(7) $A = [a_{ij}]$ $(m \times m)$ positive semidefinite:

$$\det \begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix} \geq 0, \quad i = 1, \dots, m.$$

(8) $A = [a_{ij}]$ $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \dots \leq \lambda_m$:

$$\prod_{i=1}^{k} \lambda_i \leq \det \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix} \leq \prod_{i=1}^{k} \lambda_{m-k+i}.$$

(9) A $(m \times m)$ positive definite: $\ln \det(A) \leq \mathrm{tr}(A) - m$

(10) A $(m \times m)$ positive semidefinite: $(\det A)^{1/m} \leq \frac{1}{m}\mathrm{tr}(A)$.

(11) A $(m \times m)$ positive definite, B $(m \times m)$ positive semidefinite:

$$\det(A + B) \geq \det(A).$$

(12) A, B $(m \times m)$ positive semidefinite:

$$\det(A + B) \geq \det(A) + \det(B).$$

(13) A $(m \times m)$ positive definite, B $(m \times m)$ negative semidefinite:

$$\det(A + B) \leq \det(A).$$

(14) (Minkowski's inequality)
$A \neq 0, B \neq 0$ $(m \times m)$ positive semidefinite:

$$[\det(A + B)]^{1/m} \geq (\det A)^{1/m} + (\det B)^{1/m}.$$

(15) A, B $(m \times m)$ positive semidefinite, $\alpha \in \mathbb{R}$, $0 < \alpha < 1$:

$$(\det A)^{\alpha}(\det B)^{1-\alpha} \leq \det(\alpha A + (1 - \alpha)B).$$

(16) A, B $(m \times m)$ positive semidefinite:

$$(\det A)^{1/m}(\det B)^{1/m} \leq \frac{1}{m}\mathrm{tr}(AB).$$

(17) A $(m \times m)$ positive definite, B $(m \times m)$ positive semidefinite:

$$\frac{\det(A+B)}{\det(A)} \leq \exp[\mathrm{tr}(A^{-1}B)].$$

(18) (Oppenheim's inequality)
$A, B = [b_{ij}]$ $(m \times m)$ positive semidefinite:

$$\det(A \odot B) \geq \det(A) \prod_{i=1}^{m} b_{ii}.$$

(19) A, B $(m \times m)$ positive definite: $\det(A \odot B) \geq \det(A)\det(B)$.

(20) (Ostrowski Taussky inequality)
A $(m \times m)$:
$\frac{1}{2}(A + A^H)$ positive definite $\Rightarrow$ $\det \frac{1}{2}(A + A^H) \leq |\det(A)|_{\mathrm{abs}}$.

(21) A, B $(m \times m)$ real: $\det(AB) \leq \frac{1}{2}[(\det A)^2 + (\det B)^2]$.

(22) A, B $(m \times m)$ real: $\det(A \odot B) \leq \frac{1}{2}[\det(A \odot A) + \det(B \odot B)]$.

Note: Inequalities (21) and (22) are from Neudecker & Shuangzhe (1993, 1995). The other inequalities may be found in Horn & Johnson (1985) and Magnus & Neudecker (1988).

4.2.7 Optimization of Functions Involving a Determinant

In this subsection all matrices are real if not stated otherwise.

(1) Ω $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B'\Omega B) \ : \ B \ (m \times n), B'B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\max\{\det(B'\Omega B) \ : \ B \ (m \times n), B'B = I_n\} = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(2) Ω $(m \times m)$ complex Hermitian positive definite with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B^H\Omega B) \ : \ B \ (m \times n) \text{ complex}, B^H B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\max\{\det(B^H\Omega B) : B \ (m \times n) \text{ complex}, B^H B = I_n\} \\ = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(3) $Y\ (m \times n), X\ (p \times n), \mathrm{rk}(X) = p$:

$$\begin{aligned}\min\{\det[(Y - AX)(Y - AX)'] \ :\ A\ (m \times p)\} \\ = \det[Y(I_n - X'(XX')^{-1}X)Y'].\end{aligned}$$

The minimum is attained for $A = YX'(XX')^{-1}$.

(4) $Y\ (m \times n),\ X\ (p \times n)$ with $\mathrm{rk}(X) = p,\ R\ (q \times m)$ with $\mathrm{rk}(R) = q$, $C\ (q \times p)$:

$$\begin{aligned}\min\{\det[(Y - AX)(Y - AX)'] \ :\ A\ (m \times p), RA = C\} \\ = \det[(Y - \dot{A}X)(Y - \dot{A}X)'],\end{aligned}$$

where $\dot{A} = YX'(XX')^{-1} + R'(RR')^{-1}(C - RYX'(XX')^{-1})$.

(5) $Y, X\ (m \times n),\ \mathrm{rk}(Y) = \mathrm{rk}(X) = m,\ \lambda_1 \leq \cdots \leq \lambda_m$ the eigenvalues of

$$(XX')^{-1/2}XY'(YY')^{-1}YX'(XX')^{-1/2'}$$

with associated orthonormal eigenvectors $v_1, \ldots, v_m$:

$$\begin{aligned}&\min\{\det[(Y - BCX)(Y - BCX)'] :\\ &\quad B\ (m \times r), C\ (r \times m), \mathrm{rk}(B) = \mathrm{rk}(C) = r\}\\ &\quad = \det(YY')(1 - \lambda_m)\cdots(1 - \lambda_{m-r+1}).\end{aligned}$$

The minimum is attained for

$$C = [v_m, \ldots, v_{m-r+1}]'(XX')^{-1/2}$$

and

$$B = YX'C'(CXX'C')^{-1}.$$

(6) $B, C\ \ (m \times m)$ positive definite, B diagonal, $\lambda_1 \leq \cdots \leq \lambda_m$ the eigenvalues of $B^{-1/2}CB^{-1/2}$ with associated orthonormal eigenvectors $v_1, \ldots, v_m,\ 1 \leq r \leq m$:

$$\begin{aligned}\min\{\ln\det(AA' + B) + \mathrm{tr}[(AA' + B)^{-1}C] \ :\ A\ (m \times r)\} \\ = m + \ln\det(C) + \sum_{i=1}^{m-r}(\lambda_i - \ln\lambda_i - 1).\end{aligned}$$

The minimum is attained for

$$A = B^{1/2}[v_m, \ldots, v_{m-r+1}]'\left(\begin{bmatrix} \lambda_m & & 0 \\ & \ddots & \\ 0 & & \lambda_{m-r+1} \end{bmatrix} - I_r\right)^{1/2}.$$

Note: The results of this subsection may be obtained from Magnus & Neudecker (1988) and Lütkepohl (1991, Section A.14).

4.3 The Rank of a Matrix

Definitions: The rank of a matrix A, denoted by rk A or rk(A), is the maximum number of linearly independent rows or columns of A. The column rank of a matrix A, denoted by col rk A or col rk(A), is the number of linearly independent columns of A. The row rank of a matrix A, denoted by row rk A or row rk(A), is the number of linearly independent rows of A.

4.3.1 General Results

(1) $A\ (m \times n)$: row rk(A) = col rk(A) = rk(A).

(2) $A\ (m \times n), c \in \mathbb{C}$: $c \neq 0 \Rightarrow \text{rk}(cA) = \text{rk}(A)$.

(3) $A\ (m \times n)$:

(a) $\text{rk}(A') = \text{rk}(A)$.

(b) $\text{rk}(\bar{A}) = \text{rk}(A)$.

(c) $\text{rk}(A^H) = \text{rk}(A)$.

(d) $\text{rk}(A^+) = \text{rk}(A)$.

(e) $\text{rk}(A^-) = \text{rk}(A) \iff A^-AA^- = A^-$.

(4) $A\ (m \times m)$:

$$\text{rk}(A^{adj}) = \begin{cases} m & \text{if} \quad \text{rk}(A) = m \\ 1 & \text{if} \quad \text{rk}(A) = m-1 \\ 0 & \text{if} \quad \text{rk}(A) < m-1 \end{cases}.$$

(5) $A\ (m \times n)$:

(a) $\text{rk}(AA') = \text{rk}(A'A) = \text{rk}(A)$.

(b) $\text{rk}(AA^H) = \text{rk}(A^HA) = \text{rk}(A)$.

(6) $a, b\ (m \times 1)$: $a, b \neq 0 \Rightarrow \text{rk}(ab') = 1$.

(7) $A\ (m \times n)$: $\text{rk}(A) = r \Rightarrow$ there exist matrices $B\ (m \times r)$ and $C\ (r \times n)$ such that $A = BC$.

(8) $A\ (m \times n), B\ (n \times n)$: B nonsingular $\Rightarrow \text{rk}(AB) = \text{rk}(A)$.

(9) $A\ (m \times n), B\ (m \times m)$: B nonsingular $\Rightarrow \text{rk}(BA) = \text{rk}(A)$.

(10) $A\ (m \times n), B\ (m \times m), C\ (n \times n)$:

$$B, C \text{ nonsingular} \Rightarrow \text{rk}(BAC) = \text{rk}(A).$$

(11) $A\ (m \times n), B\ (p \times q)$: $\text{rk}(A \otimes B) = \text{rk}(A)\text{rk}(B)$.

(12) $A\ (m \times m), B\ (n \times n)$: $\text{rk}(A \oplus B) = \text{rk}(A) + \text{rk}(B)$.

(13) $A\ (m \times n), B\ (r \times m)$: $\operatorname{rk}(B) = r,\ \operatorname{rk}(A) = m \Rightarrow \operatorname{rk}(BA) = r.$

(14) $A\ (m \times n)$:

(a) $\operatorname{rk}(A) = m \Rightarrow A^+ = A^H(AA^H)^{-1}.$

(b) $\operatorname{rk}(A) = n \Rightarrow A^+ = (A^H A)^{-1} A^H.$

(15) $A\ (m \times m)$:

(a) $\operatorname{rk}(A) < m - 1 \Rightarrow A^{adj} = 0.$

(b) $\operatorname{rk}(A) = m \iff \det(A) \neq 0.$

(c) $\operatorname{rk}(A) = m \iff A$ is nonsingular.

(16) $A\ (m \times n)$: $\operatorname{rk}(A) = 0 \iff A = 0.$

(17) $A\ (m \times m)$ idempotent:

(a) $\operatorname{rk}(A) = m \Rightarrow A = I_m.$

(b) $\operatorname{rk}(A) = \operatorname{tr}(A).$

(c) $\operatorname{rk}(I_m - A) = m - \operatorname{rk}(A).$

(18) $\operatorname{rk}(I_m) = m.$

(19) $K_{mn}\ (mn \times mn)$ commutation matrix: $\operatorname{rk}(K_{mn}) = mn.$

(20) $D_m\ (m^2 \times \frac{1}{2}m(m+1))$ duplication matrix: $\operatorname{rk}(D_m) = \frac{1}{2}m(m+1).$

(21) $L_m\ (\frac{1}{2}m(m+1) \times m^2)$ elimination matrix: $\operatorname{rk}(L_m) = \frac{1}{2}m(m+1).$

(22) $A\ (m \times m)$:

(a) $\operatorname{rk}(\lambda I_m - A) < m \iff \lambda$ is an eigenvalue of A.

(b) 0 is a simple eigenvalue of $A \Rightarrow \operatorname{rk}(A) = m - 1.$

(23) $A\ (m \times n)$: $\operatorname{rk}(A) = r \Rightarrow \operatorname{rk}(K_{mn}(A' \otimes A)) = r^2.$

(24) $A\ (m \times m),\ A_{ij}\ (n \times n)$ submatrix of A :
$\operatorname{rk}(A) = r,\ n > r \Rightarrow \det(A_{ij}) = 0.$

(25) $A\ (m \times m)$:

(a) $\operatorname{rk}(A) = r \Rightarrow$ there exist $(m \times m)$ matrices $B_i,\ i = 1, \dots, r$, with $\operatorname{rk}(B_i) = 1$ and $A = B_1 + \cdots + B_r$.

(b) $\operatorname{rk}(A^i) = \operatorname{rk}(A^{i+1})$ for some $i \in \mathbb{N} \Rightarrow \operatorname{rk}(A^i) = \operatorname{rk}(A^j)$ for all $j \geq i$.

(c) There exists an $i \in \mathbb{N}$ such that $\operatorname{rk}(A^i) = \operatorname{rk}(A^{i+1})$.

(26) $A\ (m \times m)$ nonsingular, $B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$\operatorname{rk}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = m \iff D = CA^{-1}B.$$

Note: Many of these results are elementary and follow immediately from definitions (see, e.g., textbooks such as Horn & Johnson (1985) and Lancaster & Tismenetsky (1985)). See also Magnus (1988) for some of the results.

4.3.2 Matrix Decompositions Related to the Rank

(1) (Singular value decomposition)
A $(m \times n)$, $\text{rk}(A) = r$, $\sigma_1, \ldots, \sigma_r \neq 0$ singular values of A: There exist unitary matrices U $(m \times m)$, V $(n \times n)$ such that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^H,$$

where $D = \text{diag}(\sigma_1, \ldots, \sigma_r)$ and some or all of the zero submatrices disappear when $r = m$ and/or $r = n$.

(2) A $(m \times m)$ symmetric: $\text{rk}(A) = r \iff$ there exists a nonsingular $(m \times m)$ matrix T such that

$$A = T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T',$$

where the zero submatrices disappear when $r = m$.

(3) (Diagonal reduction)
A $(m \times n)$: $\text{rk}(A) = r \iff$ there exists a nonsingular $(m \times m)$ matrix S and a nonsingular $(n \times n)$ matrix T such that

$$A = S \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T,$$

where some or all of the zero submatrices disappear when $r = m$ and/or $r = n$.

(4) A $(m \times m)$: $\text{rk}(A) = \text{rk}(A^2) \iff$ there exist nonsingular matrices P and D such that

$$A = P \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} P^{-1}.$$

(5) (Rank factorization)
A $(m \times n)$: $\text{rk}(A) = r \iff$ there exists an $(m \times r)$ matrix B and an $(r \times n)$ matrix C with $\text{rk}(B) = \text{rk}(C) = r$ such that $A = BC$.

(6) (Triangular factorization)
$A = [a_{ij}]$ $(m \times m)$, $\text{rk}(A) = r$:

$$\det \begin{bmatrix} a_{11} & \ldots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \ldots & a_{ii} \end{bmatrix} \neq 0, \quad i = 1, \ldots, r$$

$\Rightarrow$ there exists a lower triangular $(m \times m)$ matrix L and an upper triangular $(m \times m)$ matrix U one of which is nonsingular such that $A = LU$.

(7) (Polar decomposition)
$A\ (m \times m)$: $\text{rk}(A) = r \iff$ there exists a positive semidefinite $(m \times m)$ matrix P with $\text{rk}(P) = r$ and a unitary $(m \times m)$ matrix U such that $A = PU$.

Note: For these and further decomposition theorems see Chapter 6.

4.3.3 Inequalities Related to the Rank

(1) $A\ (m \times n)$:

(a) $\text{rk}(A) \leq \min(m, n)$.

(b) $\text{rk}(A^-) \geq \text{rk}(A)$.

(2) $A\ (m \times n), B\ (n \times r)$:

(a) $\text{rk}(AB) \leq \min\{\text{rk}(A),\ \text{rk}(B)\}$.

(b) $\text{rk}(A) + \text{rk}(B) \leq \text{rk}(AB) + n$.

(c) $AB = 0 \ \Rightarrow\ \text{rk}(A) + \text{rk}(B) \leq n$.

(3) $A, B\ (m \times n)$: $\text{rk}(A + B) \leq \text{rk}(A) + \text{rk}(B)$.

(4) $A\ (m \times n), B\ (m \times r)$: $\text{rk}([A : B]) \leq \text{rk}(A) + \text{rk}(B)$.

(5) $A\ (m \times n)$: $Ax = 0$ for some $(n \times 1)$ vector $x \neq 0 \ \Rightarrow\ \text{rk}(A) \leq n - 1$.

(6) (Sylvester's law)
$A, B\ (m \times m)$: $\text{rk}(A) = r,\ \text{rk}(B) = s \ \Rightarrow\ \text{rk}(AB) \geq r + s - m$.

(7) $A, B\ (m \times m)$ Hermitian : $\text{rk}(A \odot B) \leq \text{rk}(A)\text{rk}(B)$.

(8) $A\ (m \times m)$ Hermitian : $\text{rk}(A) \geq [\text{tr}(A)]^2/\text{tr}(A^2)$.

(9) (Frobenius inequality)
$A\ (m \times n), B\ (n \times r), C\ (r \times s)$:

$$\text{rk}(AB) + \text{rk}(BC) \leq \text{rk}(B) + \text{rk}(ABC).$$

Note: These results can be found in many matrix textbooks (see, e.g., Horn & Johnson (1985) or Rao (1973)) or follow easily from rules given there.

5

Eigenvalues and Singular Values

In this chapter all matrices are complex unless otherwise stated.

5.1 Definitions

For an $(m \times m)$ matrix A the polynomial in λ, $p_A(\lambda) = \det(\lambda I_m - A)$, is called the **characteristic polynomial** of A. The roots of $p_A(\lambda)$ are said to be the **eigenvalues**, **characteristic values**, **characteristic roots** or **latent roots** of A.

Let $\lambda_1, \dots, \lambda_n$ be the distinct eigenvalues of the $(m \times m)$ matrix A. Then the characteristic polynomial can be represented as

$$p_A(\lambda) = p_0 + p_1\lambda + \cdots + p_{m-1}\lambda^{m-1} + \lambda^m = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_n)^{m_n},$$

where the m_i are positive integers with $\sum_{i=1}^n m_i = m$. The number m_i is the **multiplicity** or **algebraic multiplicity** of the eigenvalue λ_i, $i = 1, \dots, n$. An eigenvalue is called **simple**, if its multiplicity is 1. The **geometric multiplicity** of an eigenvalue λ of an $(m \times m)$ matrix A is the number of blocks

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & \ddots & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \dots & & \lambda \end{bmatrix}$$

with λ on the principal diagonal, in the *Jordan decomposition* of A which is given in the next section (see also Section 6.1). For an $(m \times m)$ matrix A,

$$\rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda \text{ is an eigenvalue of } A\}$$

is the **spectral radius** of A. The set $\{\lambda : \lambda$ is eigenvalue of $A\}$ is the **spectrum** of A.

For an $(m \times m)$ matrix A with eigenvalue λ any $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda v$ is said to be an **eigenvector** or **characteristic vector** of A corresponding to or associated with the eigenvalue λ.

For two $(m \times m)$ matrices A and B a root of the polynomial $p_{A,B}(\lambda) = \det(\lambda B - A)$ is sometimes called **eigenvalue** or **characteristic value of A in the metric** of B. Accordingly, for an eigenvalue λ of A in the metric of B an $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda Bv$, is called an **eigenvector** or **characteristic vector** of A **in the metric** of B corresponding to or associated with λ.

For an $(m \times n)$ matrix A the nonnegative square roots of the eigenvalues of $A^H A$, if $m \geq n$, and of AA^H, if $m \leq n$, are said to be the **singular values** of A.

5.2 Properties of Eigenvalues and Eigenvectors

Notation: $\lambda(A)$ denotes an eigenvalue of the matrix A.

$\lambda_{min}(A)$ is the smallest eigenvalue of the matrix A if all eigenvalues are real.

$\lambda_{max}(A)$ is the largest eigenvalue of the matrix A if all eigenvalues are real.

5.2.1 General Results

(1) A $(m \times m)$: A has exactly m eigenvalues if the multiplicities of the roots are taken into account.

(2) A, B $(m \times m)$: A, B have the same eigenvalues $\iff \text{tr}(A^k) = \text{tr}(B^k)$, $k = 1, \ldots, m$.

(3) A $(m \times m)$ with eigenvalue λ and associated eigenvector v :

(a) $\bar{\lambda}$ is eigenvalue of $\bar{A}$ with eigenvector $\bar{v}$.

(b) λ^i is eigenvalue of A^i with eigenvector v for $i \in \mathbb{N}$.

(c) λ^{-1} is eigenvalue of A^{-1} with eigenvector v, if A is nonsingular.

(4) A $(m \times m)$:

(a) λ is eigenvalue of $A \iff \lambda$ is eigenvalue of A'.

(b) λ is eigenvalue of $A \iff \bar{\lambda}$ is eigenvalue of A^H.

(c) λ is eigenvalue of $A \Rightarrow \lambda + \tau$ is eigenvalue of $A + \tau I_m$.

(5) $A = [a_{ij}]$ $(m \times m)$:

(a) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm}) \;\Rightarrow\; a_{11}, \ldots, a_{mm}$ are the eigenvalues of A.

(b) A is triangular $\Rightarrow$ $a_{11}, \ldots, a_{mm}$ are the eigenvalues of A.

(c) A is Hermitian $\Rightarrow$ all eigenvalues of A are real numbers.

(d) A is real symmetric $\Rightarrow$ all eigenvalues of A are real numbers.

(e) A is idempotent $\Rightarrow$ all eigenvalues of A are 0 or 1.

(f) A is nilpotent $\Rightarrow$ all eigenvalues of A are 0.

(g) A is real orthogonal $\Rightarrow$ all eigenvalues of A are 1 or -1.

(h) A is singular $\Longleftrightarrow$ 0 is an eigenvalue of A.

(i) A is nonsingular $\Longleftrightarrow$ all eigenvalues of A are nonzero.

(6) A $(m \times m)$ Hermitian:

(a) A is positive definite $\Longleftrightarrow$ all eigenvalues of A are real and greater than 0.

(b) A is positive semidefinite $\Longleftrightarrow$ all eigenvalues of A are real and greater than or equal to 0.

(7) A $(m \times m)$ real symmetric:

(a) A is positive definite $\Longleftrightarrow$ all eigenvalues of A are real and greater than 0.

(b) A is positive semidefinite $\Longleftrightarrow$ all eigenvalues of A are real and greater than or equal to 0.

(8) A $(m \times m)$, B $(m \times m)$ nonsingular:
λ is eigenvalue of A $\Longleftrightarrow$ λ is eigenvalue of $B^{-1}AB$.

(9) A $(m \times m)$, B $(m \times m)$ unitary:
λ is eigenvalue of A $\Longleftrightarrow$ λ is eigenvalue of $B^H AB$.

(10) A $(m \times m)$, B $(m \times m)$ orthogonal:
λ is eigenvalue of A $\Longleftrightarrow$ λ is eigenvalue of $B'AB$.

(11) A $(m \times m)$, B $(m \times m)$ positive definite:
λ is eigenvalue of BA $\Longleftrightarrow$ λ is eigenvalue of $B^{1/2}AB^{1/2}$.

(12) A $(m \times n)$, B $(n \times m)$, $n \geq m$:

(a) λ is eigenvalue of AB $\Rightarrow$ λ is eigenvalue of BA.

(b) $\lambda \neq 0$ is eigenvalue of BA $\Rightarrow$ λ is eigenvalue of AB.

(c) $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of AB $\Rightarrow$ $\lambda_1, \ldots, \lambda_m, 0, \ldots, 0$ are the eigenvalues of BA.

(13) A $(m \times m)$, $c \in \mathbb{C}$, $c \neq 0$:

(a) λ is eigenvalue of A with eigenvector v $\Rightarrow$ $c\lambda$ is eigenvalue of cA with eigenvector v.

(b) v is eigenvector of A corresponding to eigenvalue λ $\Rightarrow$ cv is eigenvector of A corresponding to eigenvalue λ.

(14) A $(m \times m), c_1, c_2 \in \mathbb{C}, c_1$ or $c_2 \neq 0$: v_1, v_2 eigenvectors of A corresponding to an eigenvalue λ $\Rightarrow$ $c_1 v_1 + c_2 v_2$ is eigenvector of A corresponding to eigenvalue λ.

(15) A $(m \times m)$ with eigenvalues λ_i, λ_j and associated eigenvectors $v_i, v_j \neq 0$: $\lambda_i \neq \lambda_j$ $\Rightarrow$ v_i and v_j are linearly independent.

(16) A $(m \times m)$, $\gamma_0, \gamma_1, \ldots, \gamma_p \in \mathbb{C}$: λ is eigenvalue of A with eigenvector v $\Rightarrow$ $\gamma_0 + \gamma_1\lambda + \cdots + \gamma_p\lambda^p$ is eigenvalue of $\gamma_0 I_m + \gamma_1 A + \cdots + \gamma_p A^p$ with eigenvector v.

(17) A $(m \times m)$, B $(n \times n)$, $\lambda(A), \lambda(B)$ eigenvalues of A and B, respectively, with associated eigenvectors $v(A)$ and $v(B)$, respectively:

(a) $\lambda(A) \cdot \lambda(B)$ is eigenvalue of $A \odot B$ with eigenvector $v(A) \odot v(B)$.

(b) $\lambda(A)$ and $\lambda(B)$ are eigenvalues of $A \oplus B$ with eigenvectors

$$\begin{bmatrix} v(A) \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ v(B) \end{bmatrix}, \text{ respectively.}$$

(18) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:

(a) $\det(A) = \lambda_1 \cdots \lambda_m = \prod_{i=1}^{m} \lambda_i$.

(b) $\operatorname{tr}(A) = \lambda_1 + \cdots + \lambda_m = \sum_{i=1}^{m} \lambda_i$.

(19) A $(m \times m)$:

(a) A has r nonzero eigenvalues $\Rightarrow$ $\operatorname{rk}(A) \geq r$.

(b) $\operatorname{rk}(\lambda I_m - A) \leq m - 1$ $\iff$ λ is eigenvalue of A.

(c) 0 is a simple eigenvalue of A $\Rightarrow$ $\operatorname{rk}(A) = m - 1$.

(20) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:

(a) $|\lambda_i|_{\text{abs}} < 1$, $i = 1, \ldots, m$
$\iff$ $\det(I_m - Az) \neq 0$ for $|z|_{\text{abs}} \leq 1$, $z \in \mathbb{C}$.

(b) $|\lambda_i|_{\text{abs}} < 1$, $i = 1, \ldots, m$ $\Rightarrow$ $I_m \pm A$ is nonsingular.

(c) $|\lambda_i|_{\text{abs}} < 1$, $i = 1, \ldots, m$ $\Rightarrow$ $I_{m^2} \pm A \odot A$ is nonsingular.

(21) (Cayley–Hamilton theorem)

$$A\ (m\times m):\quad \lambda_1,\ldots,\lambda_m \text{ are eigenvalues of } A \;\Rightarrow\; \prod_{i=1}^{m}(\lambda_i I_m - A) = 0.$$

(22) $A\ (m \times m)$: The geometric multiplicity of an eigenvalue λ of A, that is, the number of Jordan blocks with λ on the principal diagonal, is not greater than the algebraic multiplicity of λ. (See Section 5.2.3 for the Jordan decomposition.)

Note: Most of the results of this section can be found in Horn & Johnson (1985). The others are easy to derive from results given there. Chatelin (1993) is a reference for further results on eigenvalues and vectors including computational algorithms.

5.2.2 Optimization Properties of Eigenvalues

(1) (Rayleigh–Ritz theorem)
$A\ (m \times m)$ Hermitian:

$$\lambda_{min}(A) = \min\left\{\frac{x^H A x}{x^H x} : x\ (m \times 1),\ x \neq 0\right\},$$

$$\lambda_{max}(A) = \max\left\{\frac{x^H A x}{x^H x} : x\ (m \times 1),\ x \neq 0\right\}.$$

(2) (Rayleigh–Ritz theorem for real matrices)
$A\ (m \times m)$ real symmetric:

$$\lambda_{min}(A) = \min\left\{\frac{x' A x}{x' x} : x\ (m \times 1) \text{ real},\ x \neq 0\right\},$$

$$\lambda_{max}(A) = \max\left\{\frac{x' A x}{x' x} : x\ (m \times 1) \text{ real},\ x \neq 0\right\}.$$

(3) (Courant–Fischer theorem)
$A\ (m \times m)$ Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m, 1 < i < m$:

$$\lambda_i = \min_{y_1,\ldots,y_{m-i}\ (m\times 1)} \max\left\{\frac{x^H A x}{x^H x} :\right.$$
$$\left. x\ (m \times 1),\ x \neq 0,\ x^H y_j = 0, j = 1,\ldots,m-i\right\}$$

and

$$\lambda_i = \max_{y_1,\ldots,y_{i-1}\ (m\times 1)} \min\left\{\frac{x^H A x}{x^H x} :\right.$$
$$\left. x\ (m \times 1),\ x \neq 0,\ x^H y_j = 0, j = 1,\ldots,i-1\right\}.$$

(4) (Courant–Fischer theorem for real matrices)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, $1 < i < m$:

$$\lambda_i = \min_{\substack{y_1,\ldots,y_{m-i} \\ (m \times 1) \text{ real}}} \max\left\{ \frac{x'Ax}{x'x} : \right.$$
$$\left. x \ (m \times 1) \text{ real}, x \neq 0, \ x'y_j = 0, j = 1, \ldots, m-i \right\}$$

and

$$\lambda_i = \max_{\substack{y_1,\ldots,y_{i-1} \\ (m \times 1) \text{ real}}} \min\left\{ \frac{x'Ax}{x'x} : \right.$$
$$\left. x \ (m \times 1) \text{ real}, x \neq 0, \ x'y_j = 0, j = 1, \ldots, i-1 \right\}.$$

(5) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\mathrm{tr}(X^H AX) \ : \ X \ (m \times n), \ X^H X = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $X = [v_1, \ldots, v_n]$.

$$\max\{\mathrm{tr}(X^H AX) \ : \ X \ (m \times n), \ X^H X = I_n\} = \lambda_m + \cdots + \lambda_{m-n+1}.$$

The maximizing matrix is $X = [v_m, \ldots, v_{m-n+1}]$.

(6) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\mathrm{tr}(X'AX) \ : \ X \ (m \times n) \text{ real}, X'X = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $X = [v_1, \ldots, v_n]$.

$$\max\{\mathrm{tr}(X'AX) \ : \ X \ (m \times n) \text{ real}, X'X = I_n\} = \lambda_m + \cdots + \lambda_{m-n+1}.$$

The maximizing matrix is $X = [v_m, \ldots, v_{m-n+1}]$.

(7) A $(m \times m)$ Hermitian with eigenvalues $0 < \lambda_1 \leq \cdots \leq \lambda_m$, $n \in \{1, \ldots, m\}$:

$$\lambda_1 \cdot \lambda_2 \cdots \lambda_n =$$
$$\min\{(x_1^H Ax_1) \cdots (x_n^H Ax_n) :$$
$$x_i \ (m \times 1), \ X = [x_1, \ldots, x_n] \ (m \times n), \ X^H X = I_n\}.$$

(8) A $(m \times m)$ Hermitian positive definite with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B^H AB) \ : \ B \ (m \times n), B^H B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\max\{\det(B^H AB) \ : \ B \ (m \times n), B^H B = I_n\} = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(9) A $(m \times m)$ real symmetric positive definite with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B'AB) \ : \ B \ (m \times n) \text{ real}, \ B'B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$$\max\{\det(B'AB) : B \ (m \times n) \text{ real}, \ B'B = I_n\} = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

Note: The results of this section may be found in Horn & Johnson (1985) and partly in Magnus & Neudecker (1988).

5.2.3 Matrix Decompositions Involving Eigenvalues

(1) (Schur decomposition)
A $(m \times m)$: There exists a unitary $(m \times m)$ matrix U and an upper triangular matrix Λ with the eigenvalues of A on the principal diagonal such that $A = U \Lambda U^H$.

(2) (Schur decomposition of a real matrix)
A $(m \times m)$ real with real eigenvalues: There exists a real orthogonal $(m \times m)$ matrix U and a real upper triangular matrix Λ with the eigenvalues of A on the principal diagonal such that $A = U \Lambda U'$.

(3) (Spectral decomposition of a normal matrix)
A $(m \times m)$ normal with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = U \Lambda U^H$, where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the unitary $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i^H .$$

(4) (Spectral decomposition of a Hermitian matrix)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = U\Lambda U^H$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the unitary $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i^H .$$

(5) (Spectral decomposition of a real symmetric matrix)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = U\Lambda U'$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the real orthogonal $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i' .$$

Jordan Decompositions

(6) A $(m \times m)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_k \end{bmatrix}$$

is block diagonal with blocks

$$\Lambda_i = \begin{bmatrix} \lambda_{n_i} & 1 & 0 & \ldots & 0 \\ 0 & \lambda_{n_i} & 1 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & & \lambda_{n_i} \end{bmatrix}, \quad i = 1, \ldots, k \geq n,$$

on the diagonal and $\{n_1, \ldots, n_k\} = \{1, \ldots, n\}$, that is, the same eigenvalue may appear on the diagonal of more than one Λ_i. (For more details see Chapter 6.)

(7) A $(m \times m)$ with m distinct eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = T\Lambda T^{-1}$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and T is a nonsingular $(m \times m)$ matrix whose columns are eigenvectors of A associated with $\lambda_1, \ldots, \lambda_m$.

(8) (Near-diagonal Jordan decomposition)
A $(m \times m)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, $\epsilon > 0$ a given real number: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1(\epsilon) & & 0 \\ & \ddots & \\ 0 & & \Lambda_k(\epsilon) \end{bmatrix}$$

is block diagonal with

$$\Lambda_i(\epsilon) = \begin{bmatrix} \lambda_{n_i} & \epsilon & 0 & \ldots & 0 \\ 0 & \lambda_{n_i} & \epsilon & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \epsilon \\ 0 & 0 & \ldots & & \lambda_{n_i} \end{bmatrix}, \quad i = 1, \ldots, k \geq n,$$

and $\{n_1, \ldots, n_k\} = \{1, \ldots, n\}$.

(9) (Real Jordan decomposition)
A $(m \times m)$ real with distinct real eigenvalues $\lambda_1, \ldots, \lambda_p$ and distinct complex eigenvalues $\alpha_1 \pm \mathrm{i}\beta_1, \ldots, \alpha_q \pm \mathrm{i}\beta_q$: There exists a real nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \Lambda_k & & & \\ & & & \Gamma_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Gamma_l \end{bmatrix}$$

is block diagonal with

$$\Lambda_i = \begin{bmatrix} \lambda_{p_i} & 1 & 0 & \ldots & 0 \\ 0 & \lambda_{p_i} & 1 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & & \lambda_{p_i} \end{bmatrix}, \quad i = 1, \ldots, k \geq p,$$

$\{p_1, \ldots, p_k\} = \{1, \ldots, p\}$ and

$$\Gamma_i = \begin{bmatrix} \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} & I_2 & 0 & \ldots & 0 \\ 0 & \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} & I_2 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & I_2 \\ 0 & 0 & & \ldots & \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} \end{bmatrix},$$

$i = 1, \ldots, l \geq q$, $\{q_1, \ldots, q_l\} = \{1, \ldots, q\}$.

Note: The results of this section can be found in Horn & Johnson (1985). See Chapter 6 for further matrix decomposition results.

5.3 Eigenvalue Inequalities

5.3.1 Inequalities for the Eigenvalues of a Single Matrix

(1) A $(m \times m)$ Hermitian, $x \neq 0$ $(m \times 1)$:

$$\lambda_{min}(A) \leq \frac{x^H A x}{x^H x} \leq \lambda_{max}(A).$$

(2) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A), \ldots, \lambda_m(A)$:

$$\min\left\{ \frac{x^H A x}{x^H x} \ : \ x\ (m \times 1),\ x \neq 0 \right\}$$
$$\leq \lambda_i(A) \leq \max\left\{ \frac{x^H A x}{x^H x} \ : \ x\ (m \times 1),\ x \neq 0 \right\},$$

$i = 1, \ldots, m$.

(3) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$:

$$X^H X = I_n \quad \Rightarrow \quad \sum_{i=1}^{n} \lambda_i(A) \leq \mathrm{tr}(X^H A X) \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A).$$

(4) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$ real:

$$X'X = I_n \quad \Rightarrow \quad \sum_{i=1}^{n} \lambda_i(A) \leq \mathrm{tr}(X'AX) \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A).$$

(5) $A = [a_{ij}]$ $(m \times m)$ Hermitian:

$$\lambda_{min}(A) \leq a_{ii} \leq \lambda_{max}(A), \quad i = 1, \ldots, m.$$

(6) $A = [a_{ij}]$ $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\sum_{i=1}^{n} \lambda_i(A) \leq \sum_{i=1}^{n} a_{ii} \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A), \quad n = 1, \ldots, m.$$

(7) $A = [a_{ij}]$ $(m \times m)$ Hermitian with eigenvalues $0 < \lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\prod_{i=1}^{n} \lambda_i(A) \leq \prod_{i=1}^{n} a_{ii}, \quad n = 1, \ldots, m.$$

(8) A $(m \times m)$ Hermitian positive definite with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$:

$$X^H X = I_n \quad \Rightarrow \quad \prod_{i=1}^{n} \lambda_i(A) \leq \det(X^H A X) \leq \prod_{i=1}^{n} \lambda_{m-n+i}(A).$$

(9) A $(m \times m)$ real positive definite with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$ real:

$$X'X = I_n \quad \Rightarrow \quad \prod_{i=1}^{n} \lambda_i(A) \leq \det(X'AX) \leq \prod_{i=1}^{n} \lambda_{m-n+i}(A).$$

(10) $A = [a_{ij}]$ $(m \times m)$ Hermitian positive definite with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$A_{(n)} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}$$

$$\Rightarrow \quad \prod_{i=1}^{n} \lambda_i(A) \leq \det(A_{(n)}) \leq \prod_{i=1}^{n} \lambda_{m-n+i}(A).$$

(11) (Inclusion principle)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, $A_{(n)}$ $(n \times n)$ a principal submatrix of A with eigenvalues $\lambda_1(A_{(n)}) \leq \cdots \leq \lambda_n(A_{(n)})$:

$$\lambda_i(A) \leq \lambda_i(A_{(n)}) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n,$$

$$\lambda_{min}(A) \leq \lambda_{min}(A_{(n)}) \leq \lambda_{max}(A_{(n)}) \leq \lambda_{max}(A).$$

5.3.2 Relations Between Eigenvalues of More Than One Matrix

(1) A ($m \times m$) Hermitian, B ($m \times m$) positive semidefinite:

(a) $\lambda_{min}(A + B) \geq \lambda_{min}(A)$.

(b) $\lambda_{max}(A + B) \geq \lambda_{max}(A)$.

(2) A, B ($m \times m$) Hermitian:

(a) $\lambda_{min}(A + B) \geq \lambda_{min}(A) + \lambda_{min}(B)$.

(b) $\lambda_{max}(A + B) \leq \lambda_{max}(A) + \lambda_{max}(B)$.

(3) A, B ($m \times m$) Hermitian, $0 < a, b \in \mathbb{R}$:

(a) $\lambda_{min}(aA + bB) \geq a\lambda_{min}(A) + b\lambda_{min}(B)$.

(b) $\lambda_{max}(aA + bB) \leq a\lambda_{max}(A) + b\lambda_{max}(B)$.

(4) A ($m \times m$) Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, B ($m \times m$) Hermitian positive semidefinite, $\lambda_1(A + B) \leq \cdots \leq \lambda_m(A + B)$ eigenvalues of $A + B$:

$$\lambda_i(A + B) \geq \lambda_i(A), \quad i = 1, \ldots, m.$$

(5) A ($m \times m$) Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, x ($m \times 1$), $\lambda_1(A \pm xx^H) \leq \cdots \leq \lambda_m(A \pm xx^H)$ eigenvalues of $A \pm xx^H$:

$$\lambda_i(A \pm xx^H) \leq \lambda_{i+1}(A) \leq \lambda_{i+2}(A \pm xx^H), \quad i = 1, 2, \ldots, m - 2;$$

$$\lambda_i(A) \leq \lambda_{i+1}(A \pm xx^H) \leq \lambda_{i+2}(A), \quad i = 1, 2, \ldots, m - 2.$$

(6) A, B ($m \times m$) Hermitian, $\mathrm{rk}(B) \leq r$, $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ eigenvalues of A, $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

$$\lambda_i(A + B) \leq \lambda_{i+r}(A) \leq \lambda_{i+2r}(A + B), \quad i = 1, 2, \ldots, m - 2r;$$

$$\lambda_i(A) \leq \lambda_{i+r}(A + B) \leq \lambda_{i+2r}(A), \quad i = 1, 2, \ldots, m - 2r.$$

(7) (Poincare's separation theorem)
A ($m \times m$) Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X ($m \times n$) such that $X^H X = I_n$, $n \leq m$, $\lambda_1(X^H AX) \leq \cdots \leq \lambda_n(X^H AX)$ eigenvalues of $X^H AX$:

$$\lambda_i(A) \leq \lambda_i(X^H AX) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n.$$

(8) A, B ($m \times m$) Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A + B) \leq \cdots \leq \lambda_m(A + B)$ eigenvalues of $A + B$:

(a) $\displaystyle\sum_{i=1}^{n}[\lambda_i(A) + \lambda_i(B)] \leq \sum_{i=1}^{n}\lambda_i(A + B), \quad n = 1, \ldots, m.$

(b) $\lambda_1(B) \leq \lambda_n(A+B) - \lambda_n(A) \leq \lambda_m(B), \quad n = 1, \ldots, m.$

(c) $|\lambda_n(A+B) - \lambda_n(A)|_{\text{abs}}$
$\leq \max\{|\lambda_j(B)|_{\text{abs}} \; : \; j = 1, \ldots, m\} = \rho(B), \quad n = 1, \ldots, m.$

(d) $\lambda_n(A+B) \leq \min\{\lambda_i(A) + \lambda_j(B) \; : \; i + j = n + m\}$,
$n = 1, \ldots, m$.

(9) (Weyl's theorem)
A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

$$\lambda_i(A+B) \geq \begin{cases} \lambda_i(A) + \lambda_1(B) \\ \lambda_{i-1}(A) + \lambda_2(B) \\ \vdots \\ \lambda_1(A) + \lambda_i(B) \end{cases}$$

and

$$\lambda_i(A+B) \leq \begin{cases} \lambda_i(A) + \lambda_m(B) \\ \lambda_{i+1}(A) + \lambda_{m-1}(B) \\ \vdots \\ \lambda_m(A) + \lambda_i(B) \end{cases}$$

$i = 1, \ldots, m$.

(10) A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A-B) \leq \cdots \leq \lambda_m(A-B)$ eigenvalues of $A-B$:

$$\lambda_1(A-B) > 0 \quad \Rightarrow \quad \lambda_i(A) \geq \lambda_i(B), \; i = 1, \ldots, m.$$

(11) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, x $(m \times 1)$, $c \in \mathbb{R}$,

$$B = \begin{bmatrix} A & x \\ x^H & c \end{bmatrix}$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_{m+1}(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{i+1}(B), \quad i = 1, \ldots, m.$$

(12) A $(n \times n)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_n(A)$, D $(p \times p)$ Hermitian, C $(n \times p)$ and

$$B = \begin{bmatrix} A & C \\ C^H & D \end{bmatrix} (m \times m)$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{m-n+i}(B) \quad \text{for } i = 1, \ldots, n.$$

Note: The results of this section are given in Horn & Johnson (1985, Chapters 4 and 7) or follow easily from results given there. Some results, in particular those on real symmetric matrices, may also be found in Magnus & Neudecker (1988). Many of these and further results are also contained in Chatelin (1993).

5.4 Results for the Spectral Radius

Reminder: $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times n)$ real:

$$\begin{array}{lcl} A > (\geq)\ 0 & \iff & a_{ij} > (\geq)\ 0,\ i = 1, \ldots, m,\ j = 1, \ldots, n. \\ A > (\geq)\ B & \iff & a_{ij} > (\geq)\ b_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n. \end{array}$$

(1) A $(m \times m)$, $i \in \mathbb{N}$: $\rho(A)^i = \rho(A^i)$.

(2) A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A + B) \leq \cdots \leq \lambda_m(A + B)$ eigenvalues of $A + B$:

$$|\lambda_i(A + B) - \lambda_i(A)|_{\text{abs}} \leq \rho(B), \quad i = 1, \ldots, m.$$

(3) A $(m \times m)$:

(a) $\rho(A) < 1 \iff A^n \to 0$ as $n \to \infty$, i.e., A is convergent.

(b) $\rho(A) < 1 \Rightarrow \displaystyle\sum_{i=0}^{\infty} A^i = (I_m - A)^{-1}$.

(c) $\rho(A) < 1 \Rightarrow \displaystyle\sum_{i=0}^{\infty} (-A)^i = (I_m + A)^{-1}$.

(d) $\rho(A) < 1 \Rightarrow \displaystyle\sum_{i=0}^{\infty} A^i \odot A^i = (I_{m^2} - A \odot A)^{-1}$.

(e) $\rho(A) < 1 \Rightarrow \displaystyle\sum_{i=0}^{\infty} (-A)^i \odot A^i = (I_{m^2} + A \odot A)^{-1}$.

(4) $A = [a_{ij}]$ $(m \times m)$ real:

(a) $A > 0 \Rightarrow \rho(A) > 0$.

(b) $A \geq 0$, $\displaystyle\sum_{j=1}^{m} a_{ij} > 0$, $i = 1, \ldots, m \Rightarrow \rho(A) > 0$.

(c) A stochastic $\Rightarrow \rho(A) > 0$.

(d) A doubly stochastic $\Rightarrow \rho(A) > 0$.

(e) $A \geq 0, A^i > 0$ for some $i \geq 1 \;\Rightarrow\; \rho(A) > 0$.

(5) A, B $(m \times m)$ real:

(a) $0 \leq A \leq B \;\Rightarrow\; \rho(A) \leq \rho(B)$.

(b) $0 < A < B \;\Rightarrow\; \rho(A) < \rho(B)$.

(6) $A = [a_{ij}]$ $(m \times m)$ real,

$$A_{(n)} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

principal submatrix of A :

$$A \geq 0 \;\Rightarrow\; \rho(A_{(n)}) \leq \rho(A), \; n = 1, \dots, m.$$

(7) $A = [a_{ij}] \geq 0$ $(m \times m)$ real:

(a) $\max\limits_{i=1,\dots,m} a_{ii} \leq \rho(A)$.

(b) $\min\limits_{i=1,\dots,m} \sum\limits_{j=1}^{m} a_{ij} \leq \rho(A) \leq \max\limits_{i=1,\dots,m} \sum\limits_{j=1}^{m} a_{ij}$.

(c) $\min\limits_{j=1,\dots,m} \sum\limits_{i=1}^{m} a_{ij} \leq \rho(A) \leq \max\limits_{j=1,\dots,m} \sum\limits_{i=1}^{m} a_{ij}$.

(d) $\rho(A) = \max\limits_{\substack{x=(x_1,\dots,x_m)' \geq 0 \\ x \neq 0}} \; \min\limits_{\substack{1 \leq i \leq m \\ x_i \neq 0}} \dfrac{1}{x_i} \sum\limits_{j=1}^{m} a_{ij} x_j$.

(8) $A = [a_{ij}] \geq 0$ $(m \times m)$ real, $x = (x_1, \dots, x_m)' > 0$ $(m \times 1)$ real:

(a) $\min\limits_{i=1,\dots,m} \dfrac{1}{x_i} \sum\limits_{j=1}^{m} a_{ij} x_j \leq \rho(A) \leq \max\limits_{i=1,\dots,m} \dfrac{1}{x_i} \sum\limits_{j=1}^{m} a_{ij} x_j$.

(b) $\min\limits_{j=1,\dots,m} x_j \sum\limits_{i=1}^{m} \dfrac{a_{ij}}{x_i} \leq \rho(A) \leq \max\limits_{j=1,\dots,m} x_j \sum\limits_{i=1}^{m} \dfrac{a_{ij}}{x_i}$.

(9) A $(m \times m)$ real, $A \geq 0, x$ $(m \times 1)$ real, $x > 0, a, b \in \mathbb{R}, a, b \geq 0$:

(a) $ax < Ax < bx \;\Rightarrow\; a < \rho(A) < b$.

(b) $ax \leq Ax \leq bx \;\Rightarrow\; a \leq \rho(A) \leq b$.

(c) $Ax = ax \;\Rightarrow\; a = \rho(A)$.

(10) A $(m \times m)$ real, λ eigenvalue of A with real eigenvector $x \neq 0$:
$A > 0, |\lambda|_{\text{abs}} = \rho(A)$
$\Rightarrow$ $|\lambda|_{\text{abs}}$ is eigenvalue of A with eigenvector $|x|_{\text{abs}}$.

(11) A $(m \times m)$ real, $A > 0$:

(a) $\rho(A)$ is a simple eigenvalue of A.

(b) $\rho(A)$ is eigenvalue of A with eigenvector $x > 0$.

(12) A $(m \times m)$ real, $A \geq 0$:
$\rho(A)$ is eigenvalue of A with eigenvector $x \geq 0, x \neq 0$.

(13) A $(m \times m)$ real, $A > 0, \lambda$ eigenvalue of A :
$\lambda \neq \rho(A) \;\Rightarrow\; |\lambda|_{\text{abs}} < \rho(A)$.

(14) (Hopf's theorem)
$A = [a_{ij}]$ $(m \times m)$ real, $A > 0, \lambda_{m-1}$ eigenvalue of A with second largest modulus, $M = \max\{a_{ij} : i, j = 1, \ldots, m\}, \mu = \min\{a_{ij} : i, j = 1, \ldots, m\}$:

$$\frac{|\lambda_{m-1}|_{\text{abs}}}{\rho(A)} \leq \frac{M - \mu}{M + \mu} < 1.$$

(15) A $(m \times m)$ real, $A > 0, x > 0$ $(m \times 1)$ real eigenvector of A corresponding to eigenvalue $\rho(A), y > 0$ $(m \times 1)$ real eigenvector of A' corresponding to eigenvalue $\rho(A'), x'y = 1$: $\lim_{i\to\infty} [\rho(A)^{-1}A]^i = xy'$.

(16) A $(m \times m)$ real, $A > 0$:

$$\lim_{i\to\infty} [\rho(A)^{-1}A]^i > 0,$$

$$\text{rk}\{\lim_{i\to\infty} [\rho(A)^{-1}A]^i\} = 1.$$

Note: Most results of this subsection are, for instance, given in Horn & Johnson (1985).

5.5 Singular Values

Notation: $\sigma(A)$ denotes a singular value of the matrix A.
$\sigma_{min}(A)$ denotes the smallest singular value of the matrix A.
$\sigma_{max}(A)$ denotes the largest singular value of the matrix A.

5.5.1 General Results

(1) A $(m \times n), m \geq n$:

$$\begin{aligned}\sigma_{min}(A) &= \min\left\{ \left(\frac{x^H A^H A x}{x^H x}\right)^{1/2} : x \ (n \times 1), \ x \neq 0 \right\} \\ &= \min\{(x^H A^H A x)^{1/2} : x \ (n \times 1), x^H x = 1\}.\end{aligned}$$

(2) $A\ (m \times n)$:

$$\begin{aligned}\sigma_{max}(A) &= \max\left\{\left(\frac{x^H A^H A x}{x^H x}\right)^{1/2} : x\ (n \times 1),\ x \neq 0\right\}\\ &= \max\{(x^H A^H A x)^{1/2} : x\ (n \times 1), x^H x = 1\}.\end{aligned}$$

(3) $A\ (m \times n), r = \min\{m, n\}$, $\sigma_1 \geq \cdots \geq \sigma_r$ singular values of A, $1 < i < r$:

$$\sigma_i = \min_{y_1,\dots,y_{i-1}\ (m\times 1)} \max\left\{\left(\frac{x^H A^H A x}{x^H x}\right)^{1/2} : x\ (m \times 1),\ x \neq 0,\ x^H y_j = 0, j = 1, \dots, i-1\right\}$$

and

$$\sigma_i = \max_{y_1,\dots,y_{m-i}\ (m\times 1)} \min\left\{\left(\frac{x^H A^H A x}{x^H x}\right)^{1/2} : x\ (m \times 1),\ x \neq 0,\ x^H y_j = 0, j = 1, \dots, m-i\right\}.$$

(4) $A\ (m \times m)$ nonsingular:

$$\frac{1}{\sigma_{min}(A)} = \max\left\{\left(\frac{x^H (A^{-1})^H A^{-1} x}{x^H x}\right)^{1/2} : x\ (m \times 1),\ x \neq 0\right\}.$$

(5) $A\ (m \times n)$: $\sigma_1, \dots, \sigma_r$ singular values of A $\Longleftrightarrow$ $\sigma_1, \dots, \sigma_r$ singular values of A^H.

(6) $A\ (m \times n), r = \min\{m, n\}$, $\sigma_1, \dots, \sigma_r$ singular values of A :

$$\mathrm{tr}(A^H A) = \sum_{i=1}^{r} \sigma_i^2.$$

(7) (Singular value decomposition)
$A\ (m \times n)$, $\mathrm{rk}(A) = r$, $\sigma_1, \dots, \sigma_r \neq 0$ singular values of A : There exist unitary matrices $U\ (m \times m)$, $V\ (n \times n)$ such that

$$U^H A V = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

where $D = \mathrm{diag}(\sigma_1, \dots, \sigma_r)$ and some or all of the zero submatrices disappear if $r = m$ and/or $r = n$.

(8) (Singular value decomposition of a real matrix)
A $(m \times n)$ real, $\text{rk}(A) = r$, $\sigma_1, \dots, \sigma_r \neq 0$ singular values of A : There exist real orthogonal matrices U $(m \times m)$, V $(n \times n)$ such that

$$U'AV = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

where $D = \text{diag}(\sigma_1, \dots, \sigma_r)$ and some or all of the zero submatrices disappear if $r = m$ and/or $r = n$.

(9) A $(m \times n)$:

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^H \text{ is the singular value decomposition of } A$$

$$\Rightarrow \quad A^+ = V \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^H.$$

(10) A $(m \times n), r = \min\{m, n\}$, $0 \leq \sigma_1, \dots, \sigma_r \in \mathbb{R}$: $\sigma_1, \dots, \sigma_r$ are the singular values of A $\iff$ $\sigma_1, \dots, \sigma_r, -\sigma_1, \dots, -\sigma_r$ and $|m - n|_{abs}$ zeros are the eigenvalues of

$$\begin{bmatrix} 0 & A \\ A^H & 0 \end{bmatrix}.$$

5.5.2 Inequalities

(1) A, B $(m \times n)$, $r = \min\{m, n\}$, $\sigma_1(A) \geq \cdots \geq \sigma_r(A) \geq 0$, $\sigma_1(B) \geq \cdots \geq \sigma_r(B) \geq 0$, $\sigma_1(A + B) \geq \cdots \geq \sigma_r(A + B) \geq 0$ singular values of A, B and $A + B$, respectively:

$$\sigma_{i+j-1}(A + B) \leq \sigma_i(A) + \sigma_j(B), \quad 1 \leq i, j \leq r \text{ with } i + j \leq r + 1.$$

(2) A, B $(m \times n)$: $\quad \sigma_{max}(A + B) \leq \sigma_{max}(A) + \sigma_{max}(B)$.

(3) A, B $(m \times n)$, $r = \min\{m, n\}$, $\sigma_1(A) \geq \cdots \geq \sigma_r(A) \geq 0$, $\sigma_1(B) \geq \cdots \geq \sigma_r(B) \geq 0$, $\sigma_1(AB^H) \geq \cdots \geq \sigma_r(AB^H) \geq 0$ singular values of A, B and AB^H, respectively:

$$\sigma_{i+j-1}(AB^H) \leq \sigma_i(A)\sigma_j(B), \quad 1 \leq i, j \leq r \text{ with } i + j \leq r + 1.$$

(4) $A = [a_1, \dots, a_m]$ $(m \times m)$, a_j $(m \times 1), \sigma_1(A) \geq \cdots \geq \sigma_m(A)$ singular values of A:

$$\sum_{j=1}^{n} \sigma_{m-n+j}(A)^2 \leq \sum_{j=1}^{n} a_j^H a_j \leq \sum_{j=1}^{n} \sigma_j(A)^2, \quad n = 1, 2, \dots, m.$$

(5) A $(m \times n)$, B $(m \times (n-1))$ is obtained from A by deleting any one column of A, $r = \min\{m, n\}$, $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$ and $\sigma_1(B) \geq \cdots \geq \sigma_{\min(m,n-1)}(B)$ singular values of A and B, respectively:

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \cdots \geq \sigma_{n-1}(B) \geq \sigma_n(A) \geq 0$$

if $m \geq n$ and

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \cdots \geq \sigma_m(A) \geq \sigma_m(B) \geq 0$$

if $m < n$.

(6) A $(m \times n)$, B $((m-1) \times n)$ is obtained from A by deleting any one row of A, $r = \min\{m, n\}$, $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$ and $\sigma_1(B) \geq \cdots \geq \sigma_{\min(m-1,n)}(B)$ singular values of A and B, respectively:

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \cdots \geq \sigma_{m-1}(B) \geq \sigma_m(A) \geq 0$$

if $m \leq n$ and

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \cdots \geq \sigma_n(A) \geq \sigma_n(B) \geq 0$$

if $m > n$.

(7) A $(m \times n)$: $\quad \sigma_{max}(A) \geq \left(\frac{1}{n}\mathrm{tr}(A^H A)\right)^{1/2}$.

Note: The results of this section can be found in Horn & Johnson (1985). Other books discussing singular values include Barnett (1990), Lancaster & Tismenetsky (1985) and, in particular, Chatelin (1993) where further results may be found.

6

Matrix Decompositions and Canonical Forms

All matrices in this chapter are complex unless otherwise specified.

6.1 Complex Matrix Decompositions

6.1.1 Jordan Type Decompositions

(1) (Jordan decomposition)
A $(m \times m)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^{-1}$, where Λ is a *Jordan form*, that is,

$$\Lambda = \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_k \end{bmatrix}$$

is block diagonal with blocks

$$\Lambda_i = \begin{bmatrix} \lambda_{n_i} & 1 & 0 & \ldots & 0 \\ 0 & \lambda_{n_i} & 1 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & & \lambda_{n_i} \end{bmatrix}, \qquad i = 1, \ldots, k \geq n,$$

on the diagonal and $\{n_1, \ldots, n_k\} = \{1, \ldots, n\}$, that is, the same eigenvalue may appear on the principal diagonal of more than one

Λ_i. For example,

$$\Lambda = \begin{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} & & 0 \\ & [2] & \\ 0 & & \begin{bmatrix} 3.2 & 1 \\ 0 & 3.2 \end{bmatrix} \end{bmatrix}$$

is a Jordan form of a matrix with the two distinct eigenvalues $\lambda_1 = 2, \lambda_2 = 3.2$ and

$$\Lambda_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad \Lambda_2 = [2], \quad \Lambda_3 = \begin{bmatrix} 3.2 & 1 \\ 0 & 3.2 \end{bmatrix}.$$

(2) (Jordan decomposition of a matrix whose eigenvalues are all distinct) A $(m \times m)$ with m distinct eigenvalues $\lambda_1, \dots, \lambda_m$: $\quad A = T\Lambda T^{-1}$, where $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$ and T is a nonsingular $(m \times m)$ matrix whose columns are eigenvectors of A corresponding to $\lambda_1, \dots, \lambda_m$.

(3) (Near-diagonal Jordan decomposition) A $(m \times m)$ with distinct eigenvalues $\lambda_1, \dots, \lambda_n, \epsilon > 0$ a given real number: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1(\epsilon) & & 0 \\ & \ddots & \\ 0 & & \Lambda_k(\epsilon) \end{bmatrix}$$

is block diagonal with

$$\Lambda_i(\epsilon) = \begin{bmatrix} \lambda_{n_i} & \epsilon & 0 & \dots & 0 \\ 0 & \lambda_{n_i} & \epsilon & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \epsilon \\ 0 & 0 & \dots & & \lambda_{n_i} \end{bmatrix}, \quad i = 1, \dots, k \geq n,$$

and $\{n_1, \dots, n_k\} = \{1, \dots, n\}$, that is, the same eigenvalue may appear on the principal diagonal of more than one $\Lambda_i(\epsilon)$.

Note: For proofs see, e.g., Horn & Johnson (1985, Chapter 3) or Barnett (1990, Chapter 8).

6.1.2 Diagonal Decompositions

(1) (Spectral decomposition of a normal matrix)
A $(m \times m)$ normal with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = U\Lambda U^H$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the unitary $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i^H .$$

(2) (Spectral decomposition of a Hermitian matrix)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = U\Lambda U^H$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the unitary $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i^H .$$

(3) (Singular value decomposition)
A $(m \times n)$, $\mathrm{rk}(A) = r$, $\sigma_1, \ldots, \sigma_r \neq 0$ singular values of A: There exist unitary matrices U $(m \times m)$, V $(n \times n)$ such that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^H ,$$

where $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ and some or all of the zero submatrices disappear when $r = m$ and/or $r = n$.

(4) (Diagonal reduction)
A $(m \times n)$, $\mathrm{rk}(A) = r$: There exists a nonsingular $(m \times m)$ matrix S and a nonsingular $(n \times n)$ matrix T such that

$$A = S \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T,$$

where some or all of the zero submatrices disappear when $r = m$ and/or $r = n$.

(5) (Congruent canonical form)
A $(m \times m)$ Hermitian: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^H$, where $\Lambda = \mathrm{diag}(d_1, \ldots, d_m)$ and the d_i are $+1, -1$ or 0 corresponding to the positive, negative and zero eigenvalues of A, respectively.

(6) A $(m \times m)$ symmetric, $\mathrm{rk}(A) = r$: There exists a nonsingular $(m \times m)$ matrix T such that

$$A = T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T',$$

where the zero submatrices disappear when $r = m$.

(7) A $(m \times m)$: There exists a nonsingular $(m \times m)$ matrix T, a unitary $(m \times m)$ matrix U and an $(m \times m)$ diagonal matrix Λ with nonnegative elements on the principal diagonal such that $A = TU\Lambda U'T^{-1}$.

Simultaneous Diagonalization of Two Matrices

(8) (Simultaneous diagonalization of two Hermitian matrices)
A, B $(m \times m)$ Hermitian, $AB = BA$: There exists a unitary $(m \times m)$ matrix U such that $A = UDU^H$ and $B = U\Lambda U^H$, where D and Λ are diagonal matrices.

(9) (Simultaneous diagonalization of a positive definite and a Hermitian matrix)
A $(m \times m)$ Hermitian positive definite, B $(m \times m)$ Hermitian: There exists a nonsingular $(m \times m)$ matrix T such that $A = TT^H$ and $B = T\Lambda T^H$, where Λ is a diagonal matrix.

(10) (Simultaneous diagonalization of a positive definite and a symmetric matrix)
A $(m \times m)$ positive definite, B $(m \times m)$ symmetric: There exists a nonsingular $(m \times m)$ matrix T such that $A = TT^H$ and $B = T\Lambda T'$, where Λ is a real diagonal matrix with nonnegative diagonal elements.

Note: The results of this subsection and further diagonal decompositions may be found in Horn & Johnson (1985, Chapters 3, 4 and 7) and Rao (1973, Chapter 1).

6.1.3 Other Triangular Decompositions and Factorizations

(1) (Schur decomposition)
A $(m \times m)$: There exists a unitary $(m \times m)$ matrix U and an upper triangular matrix Λ with the eigenvalues of A on the principal diagonal such that $A = U\Lambda U^H$.

(2) (Choleski decomposition)
A $(m \times m)$ positive definite: There exists a unique lower (upper) triangular $(m \times m)$ matrix B with real positive principal diagonal elements such that $A = BB^H$.

(3) (Triangular factorization)
$A = [a_{ij}]$ $(m \times m)$, $\text{rk}(A) = r$,

$$\det \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \neq 0, \quad n = 1, \dots, r:$$

There exists a lower triangular $(m \times m)$ matrix L and an upper triangular $(m \times m)$ matrix U, one of which is nonsingular, such that $A = LU$.

(4) (Triangular factorization of a nonsingular matrix)
$A = [a_{ij}]$ $(m \times m)$ nonsingular,

$$\det \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \neq 0, \quad n = 1, \dots, m:$$

There exists a unique lower triangular $(m \times m)$ matrix L and a unique upper triangular $(m \times m)$ matrix U both having unit principal diagonal such that $A = LDU$, where $D = \text{diag}(d_1, \dots, d_m)$ with

$$\prod_{j=1}^{n} d_j = \det \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}, \quad n = 1, \dots, m.$$

(5) (Gram–Schmidt triangular reduction)
A $(m \times n), m \geq n$: There exists an $(m \times n)$ matrix U with orthogonal columns and an upper triangular $(n \times n)$ matrix R such that $A = UR$.

(6) (Gram–Schmidt triangular reduction of a square matrix)
A $(m \times m)$: There exists a unitary $(m \times m)$ matrix U and an upper triangular $(m \times m)$ matrix R such that $A = UR$.

(7) (Gram–Schmidt triangular reduction of a nonsingular matrix)
A $(m \times m)$ nonsingular : There exists a unique unitary $(m \times m)$ matrix U and a unique upper triangular $(m \times m)$ matrix R with positive principal diagonal such that $A = UR$.

(8) A $(m \times m)$: There exist $(m \times m)$ permutation matrices P and Q, a lower triangular $(m \times m)$ matrix L and an upper triangular $(m \times m)$ matrix U so that $A = PLUQ$.

(9) A $(m \times m)$ nonsingular : There exists an $(m \times m)$ permutation matrix P, a lower triangular $(m \times m)$ matrix L and an upper triangular $(m \times m)$ matrix U such that $A = PLU$.

Note: For proofs and further details see, e.g., Horn & Johnson (1985, Chapters 2 and 3).

6.1.4 Miscellaneous Decompositions

(1) (Rank factorization)
A $(m \times n)$, $\text{rk}(A) = r$: There exists an $(m \times r)$ matrix B and an $(r \times n)$ matrix C with $\text{rk}(B) = \text{rk}(C) = r$ such that $A = BC$.

(2) (Polar decomposition)
A $(m \times m)$, $\text{rk}(A) = r$: There exists a positive semidefinite $(m \times m)$ matrix P with $\text{rk}(P) = r$ and a unitary $(m \times m)$ matrix U such that $A = PU$.

(3) A $(m \times m)$ nonsingular : There exists a symmetric $(m \times m)$ matrix P and a (complex) orthogonal $(m \times m)$ matrix Q such that $A = PQ$.

(4) (Square root decomposition)
A $(m \times m)$ positive (semi) definite: There exists a positive (semi) definite $(m \times m)$ matrix B such that $A = BB$, that is, B is a square root of A.

(5) A $(m \times m)$ unitary : There exists a real orthogonal $(m \times m)$ matrix Q and a real symmetric $(m \times m)$ matrix S such that $A = Q\exp(\mathrm{i}S)$.

(6) A $(m \times m)$ orthogonal : There exists a real orthogonal $(m \times m)$ matrix Q and a real skew-symmetric $(m \times m)$ matrix S such that $A = Q\exp(\mathrm{i}S)$.

(7) (Echelon form)
A $(m \times m)$: There exists a nonsingular $(m \times m)$ matrix B such that $BA = [c_{ij}]$ is in echelon form, that is, for each row i, either $c_{ij} = 0, j = 1, \ldots, m$, or there exists a $k \in \{1, \ldots, m\}$ such that $c_{ik} = 1$ and $c_{ij} = 0$ for $j < k$ and $c_{lk} = 0$ for $l \neq i$. (For more details on the echelon form see the Appendix.)

(8) (Hermite canonical form)
A $(m \times m)$: There exists a nonsingular $(m \times m)$ matrix B such that BA is an idempotent Hermite canonical form. (For the definition of a Hermite canonical form see the Appendix.)

(9) A $(m \times m)$: There exists a diagonalizable $(m \times m)$ matrix D and a nilpotent $(m \times m)$ matrix N such that $A = D + N$ and $DN = ND$.

Note: For results (1), (7) and (8) see Rao (1973, Chapter 1). The other results may be found in Horn & Johnson (1985, Chapter 3). Both references contain a number of further decomposition results.

6.2 Real Matrix Decompositions

6.2.1 Jordan Decompositions

(1) (Jordan decomposition of a real matrix with real eigenvalues)
A ($m \times m$) real with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ which are all real: There exists a real nonsingular ($m \times m$) matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_k \end{bmatrix}$$

is block diagonal with

$$\Lambda_i = \begin{bmatrix} \lambda_{n_i} & 1 & 0 & \ldots & 0 \\ 0 & \lambda_{n_i} & 1 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & & \lambda_{n_i} \end{bmatrix}, \quad i = 1, \ldots, k \geq n,$$

and $\{n_1, \ldots, n_k\} = \{1, \ldots, n\}$, that is, the same eigenvalue may appear on the principal diagonal of more than one Λ_i.

(2) (Real Jordan decomposition)
A ($m \times m$) real with distinct real eigenvalues $\lambda_1, \ldots, \lambda_p$ and distinct complex eigenvalues $\alpha_1 \pm \mathrm{i}\beta_1, \ldots, \alpha_q \pm \mathrm{i}\beta_q$: There exists a real nonsingular ($m \times m$) matrix T such that $A = T\Lambda T^{-1}$, where

$$\Lambda = \begin{bmatrix} \Lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \Lambda_k & & & \\ & & & \Gamma_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Gamma_l \end{bmatrix}$$

is block diagonal with

$$\Lambda_i = \begin{bmatrix} \lambda_{p_i} & 1 & 0 & \ldots & 0 \\ 0 & \lambda_{p_i} & 1 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & & \lambda_{p_i} \end{bmatrix}, \quad i = 1, \ldots, k \geq p,$$

$\{p_1, \ldots, p_k\} = \{1, \ldots, p\}$, and

$$\Gamma_i = \begin{bmatrix} \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} & I_2 & 0 & \ldots & 0 \\ 0 & \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} & I_2 & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & I_2 \\ 0 & 0 & \ldots & & \begin{bmatrix} \alpha_{q_i} & \beta_{q_i} \\ -\beta_{q_i} & \alpha_{q_i} \end{bmatrix} \end{bmatrix},$$

$i = 1, \ldots, l \geq q, \ \{q_1, \ldots, q_l\} = \{1, \ldots, q\}$.

Note: For proofs and further details see Horn & Johnson (1985, Chapter 3) or Barnett (1990, Chapter 8).

6.2.2 Other Real Block Diagonal and Diagonal Decompositions

(1) A $(m \times m)$ real normal $(A'A = AA')$: There exists a real orthogonal $(m \times m)$ matrix Q such that

$$A = Q \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_k \end{bmatrix} Q',$$

where the Λ_i are real numbers or real (2×2) matrices of the form

$$\Lambda_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}.$$

(2) (Spectral decomposition of a real symmetric matrix)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = Q\Lambda Q'$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ and Q is the real orthogonal $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i'.$$

(3) A $(m \times m)$ real skew-symmetric: There exists a real orthogonal $(m \times m)$ matrix Q such that

$$A = Q \begin{bmatrix} 0 & & & & & 0 \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \Lambda_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Lambda_k \end{bmatrix} Q',$$

where the Λ_i are real (2×2) matrices of the form

$$\Lambda_i = \begin{bmatrix} 0 & b_i \\ -b_i & 0 \end{bmatrix}.$$

(4) A $(m \times m)$ real orthogonal: There exists a real orthogonal $(m \times m)$ matrix Q such that

$$A = Q \begin{bmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_p & & & \\ & & & \Lambda_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Lambda_k \end{bmatrix} Q',$$

where the $\lambda_i = \pm 1$ and the Λ_i are real (2×2) matrices of the form

$$\Lambda_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}.$$

(5) (Singular value decomposition of a real matrix)
A $(m \times n)$ real, $\text{rk}(A) = r$, $\sigma_1, \dots, \sigma_r \neq 0$ singular values of A : There exist real orthogonal matrices U $(m \times m)$, V $(n \times n)$ such that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V',$$

where $D = \text{diag}(\sigma_1, \dots, \sigma_r)$ and some or all of the zero submatrices disappear when $r = m$ and/or $r = n$.

Note: The results (1) – (4) are taken from Horn & Johnson (1985, Chapter 2) and for (5) see Chapter 7 of the same reference.

6.2.3 Other Triangular and Miscellaneous Reductions

(1) (Schur decomposition of a real matrix with real eigenvalues)
A $(m\times m)$ real with real eigenvalues : There exists a real orthogonal $(m \times m)$ matrix Q and a real upper triangular matrix Λ with the eigenvalues of A on the principal diagonal such that $A = Q\Lambda Q'$.

(2) (Choleski decomposition of a real matrix)
A $(m \times m)$ real positive definite : There exists a unique real lower (upper) triangular $(m\times m)$ matrix B with positive diagonal elements such that $A = BB'$.

(3) (Gram–Schmidt triangular reduction of a real matrix)
A $(m \times n)$ real, $m \geq n$: There exists a real $(m \times n)$ matrix Q with orthogonal columns and a real upper triangular $(n \times n)$ matrix R such that $A = QR$.

(4) A $(m \times n)$ real, $m \geq n$: There exists a real orthogonal $(m \times m)$ matrix Q and a real upper triangular $(n \times n)$ matrix R such that

$$QA = \begin{bmatrix} R \\ O_{(m-n)\times n} \end{bmatrix},$$

where the zero matrix disappears if $n = m$.

(5) (Gram–Schmidt triangular reduction of a real square matrix)
A $(m \times m)$ real : There exists a real orthogonal $(m \times m)$ matrix Q and a real upper triangular $(m \times m)$ matrix R such that $A = QR$.

(6) (Gram–Schmidt triangular reduction of a real nonsingular matrix)
A $(m\times m)$ real nonsingular : There exists a unique real orthogonal $(m\times m)$ matrix Q and a unique upper triangular real $(m\times m)$ matrix R with positive principal diagonal such that $A = QR$.

(7) A $(m \times m)$ real : There exists a real orthogonal $(m \times m)$ matrix Q and an upper Hessenberg matrix R such that $A = QRQ'$.

(8) (Square root decomposition)
A $(m\times m)$ real positive (semi) definite: There exists a real positive (semi) definite $(m \times m)$ matrix B such that $A = BB$, that is, B is a square root of A.

(9) (Simultaneous diagonalization of a real positive definite and a real symmetric matrix)
A $(m \times m)$ real positive definite, B $(m\times m)$ real symmetric: There exists a real nonsingular $(m \times m)$ matrix T such that $A = TT'$ and $B = T\Lambda T'$, where Λ is a real diagonal matrix.

Note: These decomposition and factorization theorems can be found, for instance, in Horn & Johnson (1985) and Rao (1973). Algorithms for computing many of the matrix factors discussed in this chapter are described in Golub & Van Loan (1989).

7

Vectorization Operators

All matrices in this chapter are complex unless otherwise specified.

7.1 Definitions

vec denotes the column vectorizing operator which stacks the columns of a matrix in a column vector, that is, for an $(m \times n)$ matrix $A = [a_{ij}]$,

$$\text{vec}\, A = \text{vec}(A) = \text{col}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ a_{13} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1).$$

Related operators are

$$\begin{aligned} \text{rvec}(A) &\equiv [a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, \ldots a_{2n}, a_{31}, \ldots, a_{mn}] \ \ (1 \times mn) \\ &= [\text{vec}(A')]' \end{aligned}$$

which stacks the rows of A in a row vector and

$$\operatorname{row}(A) \equiv \operatorname{vec}(A') = \operatorname{rvec}(A)' = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \\ a_{21} \\ \vdots \\ a_{2n} \\ a_{31} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1)$$

which stacks the rows of A in a column vector. For example,

$$\operatorname{vec}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}$$

and

$$\operatorname{row}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{32} \end{bmatrix}.$$

The half-vectorization operator, vech, stacks only the columns from the principal diagonal of a square matrix downwards in a column vector, that is, for an $(m \times m)$ matrix $A = [a_{ij}]$,

$$\operatorname{vech} A = \operatorname{vech}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{33} \\ \vdots \\ a_{mm} \end{bmatrix} \quad (\tfrac{1}{2}m(m+1) \times 1).$$

For example,

$$\text{vech}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{22} \\ a_{32} \\ a_{33} \end{bmatrix}.$$

In the following sections results are given for vec and vech only. Similar results can be obtained for other vectorization operators from the results given. The following matrices are closely related to the vec and vech operators:

- K_{mn} or $K_{m,n}$ is an $(mn \times mn)$ commutation matrix defined such that for any $(m \times n)$ matrix A, $K_{mn}\text{vec}(A) = \text{vec}(A')$ (see Section 9.2 for its properties).
- D_m denotes the $(m^2 \times \frac{1}{2}m(m+1))$ duplication matrix defined such that $\text{vec}(A) = D_m\text{vech}(A)$ for any symmetric $(m \times m)$ matrix A (see Section 9.5).
- L_m denotes the $(\frac{1}{2}m(m+1) \times m^2)$ elimination matrix defined such that $\text{vech}(A) = L_m\text{vec}(A)$ for any $(m \times m)$ matrix A (see Section 9.6).

7.2 Rules for the vec Operator

(1) $A, B\ (m \times n)$: $\quad \text{vec}(A \pm B) = \text{vec}(A) \pm \text{vec}(B)$.

(2) $A\ (m \times n), c \in \mathbb{C}$: $\quad \text{vec}(cA) = c\ \text{vec}(A)$.

(3) $A\ (m \times n)$:

(a) $\text{vec}(A') = K_{mn}\text{vec}(A)$.

(b) $\text{vec}(A) = K_{nm}\text{vec}(A')$.

(4) $A\ (m \times m)$: $\quad L_m\text{vec}(A) = \text{vech}(A)$.

(5) $A\ (m \times n), B\ (n \times p)$:

$$\text{vec}(AB) = (I_p \otimes A)\text{vec}(B) = (B' \otimes I_m)\text{vec}(A) = (B' \otimes A)\text{vec}(I_n).$$

(6) $a\ (m \times 1), b\ (n \times 1)$: $\quad \text{vec}(ab') = b \otimes a$.

(7) $A\ (m \times n),\ B\ (n \times r),\ C\ (r \times s)$: $\quad \text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$.

(8) $A\ (m \times n),\ B\ (r \times s)$:

(a) $\text{vec}(A \otimes B) = (I_n \otimes K_{sm} \otimes I_r)(\text{vec}(A) \otimes \text{vec}(B))$.

(b) $\text{vec}(A' \otimes B) = (K_{ms,n} \otimes I_r)(\text{vec}(A) \otimes \text{vec}(B))$.

(c) $\text{vec}(A \otimes B') = (I_n \otimes K_{r,ms})(\text{vec}(A) \otimes \text{vec}(B))$.

(d) $\text{vec}(A' \otimes B) = (K_{ms,n}(I_n \otimes K_{ms}) \otimes I_r)\text{vec}(A \otimes B)$.

(e) $\text{vec}(A \otimes B') = (I_n \otimes K_{r,ms}(K_{ms} \otimes I_r))\text{vec}(A \otimes B)$.

(f) $\text{vec}(A) \otimes \text{vec}(B) = (I_n \otimes K_{ms} \otimes I_r)\text{vec}(A \otimes B)$.

(9) $A, B\ (m \times n)$:

(a) $\text{vec}(A \odot B) = \text{diag}(\text{vec}\, A)\text{vec}(B) = \text{diag}(\text{vec}\, B)\text{vec}(A)$.

(b) $\text{vec}(A \odot B) = (\text{vec}\, A) \odot \text{vec}(B)$.

(c) $\text{vec}(A)'\text{vec}(B) = \text{vec}(B)'\text{vec}(A)$.

(10) $A\ (m \times n),\ B\ (n \times m)$:

(a) $\text{vec}(A')'\text{vec}(B) = \text{tr}(AB) = \text{tr}(BA)$.

(b) $\text{vec}(A')'\text{vec}(B) = \text{vec}(B')'\text{vec}(A)$.

(11) $A\ (m \times n),\ B\ (n \times p),\ C\ (p \times q),\ D\ (q \times m)$:

$$\begin{aligned} \text{vec}(D')'(C' \otimes A)\text{vec}(B) &= \text{tr}(ABCD) \\ &= \text{tr}(DABC) \\ &= \text{tr}(CDAB) \\ &= \text{tr}(BCDA). \end{aligned}$$

(12) $A\ (m \times n), B\ (n \times p), C\ (p \times q), D\ (q \times m)$:

$$\begin{aligned} \text{vec}(D')'(C' \otimes A)\text{vec}(B) &= \text{vec}(A')'(D' \otimes B)\text{vec}(C) \\ &= \text{vec}(B')'(A' \otimes C)\text{vec}(D) \\ &= \text{vec}(C')'(B' \otimes D)\text{vec}(A) \\ &= \text{vec}(D)'(A \otimes C')\text{vec}(B'). \end{aligned}$$

(13) $A\ (m \times n),\ B\ (n \times p),\ C\ (p \times q),\ D\ (q \times m)$:

$$\begin{aligned} \text{vec}(D')'(C' \otimes A)\text{vec}(B) &= \text{vec}(D')'\text{vec}(ABC) \\ &= \text{vec}(A'D')'\text{vec}(BC) \\ &= \text{vec}(A'D'C')'\text{vec}(B). \end{aligned}$$

(14) $A\ (m \times m)$:

(a) A lower triangular $\Rightarrow$ $\text{vec}(A) = L'_m\text{vech}(A)$.

(b) A symmetric $\Rightarrow$ $D_m^+\, \text{vec}(A) = \text{vech}(A)$.

(c) A symmetric $\Rightarrow$ $\text{vec}(A) = D_m\text{vech}(A)$.

(15) $A\ (m \times m)$:

(a) $D'_m\text{vec}(A) = \text{vech}(A + A' - \text{dg}(A))$.

(b) $D_m^+\text{vec}(A) = \frac{1}{2}\text{vech}(A + A')$.

(c) $D_m D_m' \text{vec}(A) = \text{vec}(A + A' - \text{dg}(A))$.

(d) $\text{vec}(\text{dg}(A)) = L_m' L_m K_{mm} L_m' L_m \text{vec}(A)$.

(16) $A\ (m \times m),\ a, b\ (m \times 1),\ c \in \mathbb{C}$:

$$D_{m+1}' \text{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ a + b \\ D_m' \text{vec}(A) \end{bmatrix},$$

$$D_{m+1}^{+} \text{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ \frac{1}{2}(a + b) \\ D_m^{+} \text{vec}(A) \end{bmatrix}.$$

(17) $A, S, V\ (m \times m)$:

$$S = ASA' + V \text{ and } I_{m^2} - A \otimes A \text{ nonsingular}$$

$$\Rightarrow \quad \text{vec}(S) = (I_{m^2} - A \otimes A)^{-1} \text{vec}(V).$$

Note: A very rich source for results on the vec operator is Magnus (1988). Magnus & Neudecker (1988) also contains many of the results given here.

7.3 Rules for the vech Operator

(1) $A, B\ (m \times m)$: $\text{vech}(A \pm B) = \text{vech}(A) \pm \text{vech}(B)$.

(2) $A\ (m \times m), c \in \mathbb{C}$: $\text{vech}(cA) = c\,\text{vech}(A)$.

(3) $A, B\ (m \times m)$:

(a) $\text{vech}(A \odot B) = \text{diag}(\text{vech}\, A)\text{vech}(B) = \text{diag}(\text{vech}\, B)\text{vech}(A)$.

(b) $\text{vech}(A \odot B) = (\text{vech}\, A) \odot \text{vech}(B)$.

(4) $A\ (m \times m)$:

(a) $\text{vech}(A) = L_m \text{vec}(A)$.

(b) $\text{vech}(A + A') = 2D_m^{+} \text{vec}(A)$.

(c) $\text{vech}(A + A' - \text{dg}(A)) = D_m' \text{vec}(A)$.

(d) $D_m' D_m \text{vech}(A) = 2\text{vech}(A) - \text{vech}(\text{dg}(A))$.

(e) $\text{vech}(\text{dg}(A)) = L_m K_{mm} L_m' \text{vech}(A)$.

(5) $A\ (m \times m)$:

(a) A lower triangular $\Rightarrow$ $L_m' \text{vech}(A) = \text{vec}(A)$.

(b) A symmetric $\Rightarrow$ $\text{vech}(A) = D_m^{+} \text{vec}(A)$.

(c) A symmetric $\Rightarrow$ $D_m \text{vech}(A) = \text{vec}(A)$.

(6) A $(m \times n)$, B $(n \times p)$, C $(p \times m)$: $\quad \text{vech}(ABC) = L_m(C' \otimes A)\text{vec}(B)$.

(7) A, B $(m \times m)$: $\quad \text{vech}(A)' \text{ vech}(B) = \text{vech}(B)' \text{ vech}(A)$.

Note: Many results for the vech operator including the nontrivial ones given here are collected in Magnus (1988).

8

Vector and Matrix Norms

8.1 General Definitions

A function $||\cdot||$ attaching a nonnegative real number $||A||$ to an $(m \times n)$ matrix A is a **norm** if the following three conditions are satisfied for all complex $(m \times n)$ matrices A, B and complex numbers c :

(i) $||A|| > 0 \quad \text{if} \quad A \neq 0$,

(ii) $||cA|| = |c|_{\text{abs}}\, ||A||$,

(iii) $||A + B|| \leq ||A|| + ||B|| \qquad$ (triangle inequality).

Here $|c|_{\text{abs}}$ denotes the modulus of c, that is, $|c|_{\text{abs}} = \sqrt{c\bar{c}}$ with $\bar{c}$ being the complex conjugate of c. If (i) is replaced by

(i)′ $||A|| \geq 0$,

$||\cdot||$ is said to be a **seminorm**. In this case $||A|| = 0$ is possible even if $A \neq 0$. Instead of defining a norm for all complex $(m \times n)$ matrices, it may be defined for real matrices only. If (i), (ii) and (iii) hold for all real $(m \times n)$ matrices and real numbers c, $||\cdot||$ is called a **norm over the** field of **real numbers** ($\mathbb{R}$). In that case $|c|_{\text{abs}}$ is, of course, simply the absolute value of c. Since vectors are $(m \times 1)$ or $(1 \times n)$ matrices, norms are defined for vectors as well.

Let $||\cdot||_a$ and $||\cdot||_b$ be norms for $(m \times 1)$ and $(n \times 1)$ vectors, respectively. The norm $||\cdot||$ for $(m \times n)$ matrices is **compatible** with $||\cdot||_a$ and $||\cdot||_b$ if

$$||Ax||_a \leq ||A||\,||x||_b$$

for all $(m \times n)$ matrices A and $(n \times 1)$ vectors x. A norm $||\cdot||_a$ for $(m \times 1)$ vectors is **compatible** with a norm for $(m \times m)$ matrices $||\cdot||$ if

$$||Ax||_a \leq ||A||\,||x||_a$$

for all $(m \times m)$ matrices A and $(m \times 1)$ vectors x. An $(m \times m)$ matrix A is called an **isometry** for a norm $||\cdot||$ for $(m \times 1)$ vectors if $||Ax|| = ||x||$ for all $(m \times 1)$ vectors x.

A norm $||\cdot||$ for $(m \times m)$ matrices is called **multiplicative** or **submultiplicative** if

$$||AB|| \leq ||A||\,||B||$$

for all $(m \times m)$ matrices A, B. A multiplicative norm is said to be a **matrix norm**.

$$||A||_{lub} \equiv \sup\left\{\frac{||Ax||}{||x||} : x\ (m \times 1),\ x \neq 0\right\}$$

is the **matrix norm induced by a norm** $||\cdot||$ for $(m \times 1)$ vectors x. Alternatively it is called **sup norm** or **operator norm** or **lub** (least upper bound) **norm**.

A norm for $(m \times n)$ matrices A is **unitarily invariant** if $||UAV|| = ||A||$ for any two unitary matrices U $(m \times m)$ and V $(n \times n)$. A norm $||\cdot||$ for $(m \times n)$ matrices is **monotone** if for any two $(m \times n)$ matrices $A = [a_{ij}]$, $B = [b_{ij}]$ the following holds:

$$|a_{ij}|_{\text{abs}} \leq |b_{ij}|_{\text{abs}},\ i = 1, \ldots, m,\ j = 1, \ldots, n \ \Rightarrow\ ||A|| \leq ||B||.$$

A norm for $(m \times n)$ matrices is **absolute** if for all $(m \times n)$ matrices $A = [a_{ij}]$,

$$||A|| = ||\,[\ |a_{ij}|_{\text{abs}}\ ]\,||.$$

A matrix norm $||\cdot||$ is said to be a **minimal matrix norm**, if for any matrix norm $||\cdot||_a$ the following holds:

$$||A||_a \leq ||A|| \text{ for all } (m \times m) \text{ matrices } A \ \Rightarrow\ ||A|| = ||A||_a \text{ for all } A.$$

A matrix norm $||\cdot||$ is **self-adjoint** if $||A^H|| = ||A||$ for all $(m \times m)$ matrices A.

Related to norms are inner products. A function $< \cdot, \cdot >$ attaching a complex number $< A, B >$ to any two $(m \times n)$ matrices A, B is called an **inner product** if for any arbitrary $(m \times n)$ matrices A, B, C and $c \in \mathbb{C}$ the following conditions are satisfied:

(i) $< A, A > \ > 0 \quad \text{if} \quad A \neq 0,$

(ii) $< A + B, C > \ = \ < A, C > + < B, C >,$

(iii) $< cA, B > \ = c < A, B >,$

(iv) $< A, B > \ = \ \overline{< B, A >}.$

A function $< \cdot, \cdot >$ is an **inner product over the** field of **real numbers** if it attaches a real number to a pair of real matrices in such a way that for any real $(m \times n)$ matrices A, B, C and $c \in \mathbb{R}$ the following conditions are satisfied:

(i) $< A, A > \ > 0 \quad \text{if} \quad A \neq 0,$

(ii) $< A + B, C > \; = \; < A, C > + < B, C >$,

(iii) $< cA, B > \; = \; c < A, B >$,

(iv) $< A, B > \; = \; < B, A >$.

If $< \cdot, \cdot >$ is an inner product for $(m \times n)$ matrices, $||A|| \equiv < A, A >^{1/2}$ is a norm for $(m \times n)$ matrices. $|| \cdot ||$ is said to be a **norm derived from the inner product** $< \cdot, \cdot >$.

Note: More information on vector and matrix norms can be obtained from many books on matrices and linear algebra. Almost all results of this chapter can be found in Horn & Johnson (1985).

8.2 Specific Norms and Inner Products

Special norms for $(m \times n)$ matrices $A = [a_{ij}]$ are:

Euclidean norm, l_2-norm, Frobenius norm, Hilbert–Schmidt norm or **Schur norm:**

$$||A||_2 \equiv \sqrt{\operatorname{tr}(A^H A)} = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|_{\text{abs}}^2\right]^{1/2} \quad \text{for complex matrices.}$$

$$||A||_2 \equiv \sqrt{\operatorname{tr}(A'A)} = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|_{\text{abs}}^2\right]^{1/2} \quad \text{for real matrices.}$$

l_p-norm or **Minkowski p-norm:**

$$||A||_p \equiv \left[\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|_{\text{abs}}^p\right]^{1/p} \quad \text{for } 1 \le p < \infty.$$

Mahalanobis norm: $||A||_\Omega \equiv \left[\operatorname{tr}(A^H \Omega A)\right]^{1/2}$, where Ω is a positive definite $(m \times m)$ matrix.

Maximum norm: $||A||_\infty \equiv \max\{|a_{ij}|_{\text{abs}} : i = 1, \ldots, m,\ j = 1, \ldots, n\}$.

Column-sum norm: $$||A||_{col} \equiv \max\left\{\sum_{i=1}^{m} |a_{ij}|_{\text{abs}} : j = 1, \ldots, n\right\}.$$

Row-sum norm: $$||A||_{row} \equiv \max\left\{\sum_{j=1}^{n} |a_{ij}|_{\text{abs}} : i = 1, \ldots, m\right\}.$$

Spectral norm or **maximum singular value norm** or **operator norm:**

$$||A||_{spec} \equiv \sigma_{max}(A) \equiv \max\{\sqrt{\lambda} : \lambda \text{ is eigenvalue of } A^H A\},$$

that is, $\sigma_{max}(A)$ is the maximum singular value of A.

Special inner products for $(m \times n)$ matrices A:

$$< A, B > \equiv \operatorname{tr}(A^H B) \quad \text{for complex matrices } A, B.$$

$$< A, B > \equiv \operatorname{tr}(A'B) \quad \text{for real matrices } A, B.$$

8.3 Results for General Norms and Inner Products

(1) (Cauchy–Schwarz inequality)
$< \cdot, \cdot >$ an inner product for $(m \times n)$ matrices (over $\mathbb{R}$ or $\mathbb{C}$), A, B $(m \times n)$:

$$| < A, B > |_{\text{abs}}^2 \leq < A, A >< B, B > .$$

(2) $< \cdot, \cdot >$ an inner product for $(m \times n)$ matrices, A, B $(m \times n)$:

$$| < A, B > |_{\text{abs}}^2 = < A, A >< B, B > \Longleftrightarrow A = cB \text{ for some } c \in \mathbb{C}.$$

(3) $< \cdot, \cdot >$ an inner product over $\mathbb{R}$ for real $(m \times n)$ matrices, A, B $(m \times n)$ real:

$$| < A, B > |_{\text{abs}}^2 = < A, A >< B, B > \Longleftrightarrow A = cB \text{ for some } c \in \mathbb{R}.$$

(4) (Parallelogram identity)
$|| \cdot ||$ a norm for $(m \times n)$ matrices derived from an inner product, A, B $(m \times n)$:

$$||A + B||^2 + ||A - B||^2 = 2(||A||^2 + ||B||^2).$$

(5) (Pythagoras theorem)
$||\cdot||$ a norm for $(m \times n)$ matrices derived from an inner product $< \cdot, \cdot >$, A, B $(m \times n)$:

$$< A, B > = 0 \Rightarrow ||A + B||^2 = ||A||^2 + ||B||^2.$$

(6) (Polarization identity)
$||\cdot||$ a norm for $(m \times n)$ matrices derived from an inner product $< \cdot, \cdot >$, A, B $(m \times n)$:

$$\begin{aligned} \operatorname{Re} < A, B > &= \tfrac{1}{4}(||A + B||^2 - ||A - B||^2), \\ \operatorname{Re} < A, B > &= \tfrac{1}{2}(||A + B||^2 - ||A||^2 - ||B||^2), \end{aligned}$$

where Re denotes the real part of a complex number.

(7) $\|\cdot\|$ a norm for $(m \times n)$ matrices derived from an inner product, $A, B\ (m \times n)$:

$$\|A + B\|\,\|A - B\| \leq \|A\|^2 + \|B\|^2.$$

(8) $\|\cdot\|$ a seminorm for $(m \times n)$ matrices, $A, B\ (m \times n)$:

$$|\,\|A\| - \|B\|\,|_{\text{abs}} \leq \|A - B\|.$$

(9) $\|\cdot\|_a, \|\cdot\|_b$ norms for $(m \times n)$ matrices $A, c \in \mathbb{R}, c > 0$:

(a) $\|A\| \equiv c\|A\|_a$ is a norm for $(m \times n)$ matrices.

(b) $\|A\| \equiv \|A\|_a + \|A\|_b$ defines a norm for $(m \times n)$ matrices.

(c) $\|A\| \equiv \max(\|A\|_a, \|A\|_b)$ defines a norm for $(m \times n)$ matrices.

(10) $\|\cdot\|$ a norm for $(m \times 1)$ vectors x, $T\ (m \times m)$ nonsingular: $\|x\|_T \equiv \|Tx\|$ is a norm for $(m \times 1)$ vectors.

(11) $\|\cdot\|$ a seminorm for $(m \times 1)$ vectors x, $T\ (m \times m)$: $\|x\|_T \equiv \|Tx\|$ is a seminorm for $(m \times 1)$ vectors.

(12) $\|\cdot\|$ a norm for $(m \times m)$ matrices, $y\ (m \times 1), y \neq 0$: $\|x\|_y \equiv \|xy^H\|$ is a norm for $(m \times 1)$ vectors.

(13) $\|\cdot\|$ a norm for $(m \times m)$ matrices, $S, T\ (m \times m)$ nonsingular: $\|A\|_{S,T} \equiv \|SAT\|$ is a norm for $(m \times m)$ matrices.

(14) $\|\cdot\|$ a norm for $(m \times m)$ matrices, $S = [s_{ij}]\ (m \times m), s_{ij} \neq 0,\ i, j = 1, \ldots, m$: $\|A\|_{\odot} \equiv \|S \odot A\|$ is a norm for $(m \times m)$ matrices.

(15) $\|\cdot\|_a, \|\cdot\|_b$ norms for $(m \times n)$ matrices : There exist positive constants $c_1, c_2 \in \mathbb{R}$ such that $c_1\|A\|_a \leq \|A\|_b \leq c_2\|A\|_a$ for all $(m \times n)$ matrices A.

(16) $\|\cdot\|$ a norm for $(m \times n)$ matrices: $\|\cdot\|$ is monotone $\iff$ $\|\cdot\|$ is absolute.

(17) $\|\cdot\|$ an absolute norm for $(m \times n)$ matrices, $A, B\ (m \times n)$: $|A|_{\text{abs}} \leq |B|_{\text{abs}} \Rightarrow \|A\| \leq \|B\|$.

(18) $\|\cdot\|$ a norm for $(m \times m)$ matrices: There exists a norm for $(m \times 1)$ vectors which is compatible with $\|\cdot\| \Rightarrow \|A_1\| \cdots \|A_i\| \geq \rho(A_1 \cdots A_i) \equiv \max\{|\lambda|_{\text{abs}} : \lambda$ is an eigenvalue of $A_1 \cdots A_i\}$ for all $(m \times m)$ matrices $A_1, \ldots, A_i$ and $i = 1, 2, \ldots$

Note: The results of this section may be found in Horn & Johnson (1985, Chapter 5).

8.4 Results for Matrix Norms

8.4.1 General Matrix Norms

(1) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, $c \in \mathbb{R}$, $c \geq 1$:
$\|A\|_c \equiv c\|A\|$ is a matrix norm for $(m \times m)$ matrices.

(2) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, S $(m \times m)$ nonsingular:
$\|A\|_S \equiv \|S^{-1}AS\|$ is a matrix norm for $(m \times m)$ matrices.

(3) $\|\cdot\|_{(1)}, \dots, \|\cdot\|_{(n)}$ matrix norms for $(m \times m)$ matrices:
$\|A\| \equiv \max\{\|A\|_{(1)}, \dots, \|A\|_{(n)}\}$ is a matrix norm for $(m \times m)$ matrices.

(4) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices: $\|A\|_H \equiv \|A^H\|$ is a matrix norm for $(m \times m)$ matrices.

(5) $\|\cdot\|$ a norm for $(m \times m)$ matrices: There exists $c \in \mathbb{R}, c > 0$, such that $\|A\|_c \equiv c\|A\|$ is a matrix norm for $(m \times m)$ matrices.

(6) $\|\cdot\|$ a norm for $(m \times m)$ matrices:

$$\|A\|_{\max 1} \equiv \max\{\|AB\| : B\ (m \times m), \|B\| = 1\}$$

and

$$\|A\|_{\max 2} \equiv \max\{\|BA\| : B\ (m \times m), \|B\| = 1\}$$

are matrix norms for $(m \times m)$ matrices.

(7) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, S, T $(m \times m)$ nonsingular:
$\|A\|_{S,T} \equiv \|SAT\|$ is a norm for $(m \times m)$ matrices which may not be a matrix norm, that is, $\|\cdot\|_{S,T}$ may not be submultiplicative.

(8) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, $S = [s_{ij}]$ $(m \times m), s_{ij} \neq 0$, $i, j = 1, \dots, m$: $\|A\|_\odot \equiv \|S \odot A\|$ is a norm for $(m \times m)$ matrices which may not be a matrix norm.

(9) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, A $(m \times m)$:

(a) $\|I_m\| \geq 1$.

(b) $\|A^k\| \leq \|A\|^k$ for $k = 1, 2, \dots$

(c) $\|A\| \geq \rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda$ is eigenvalue of $A\}$.

(d) $\lim\limits_{i \to \infty} \|A^i\|^{1/i} = \rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda$ is eigenvalue of $A\}$.

(e) $A \neq 0$ idempotent $\Rightarrow$ $\|A\| \geq 1$.

(f) A singular $\Rightarrow$ $\|I_m - A\| \geq 1$.

(g) $\|I_m - A\| < 1$ $\Rightarrow$ A nonsingular.

(h) A nonsingular $\Rightarrow$ $\|A^{-1}\| \geq \dfrac{\|I_m\|}{\|A\|}$.

(i) $||I_m|| = 1$, $||A|| < 1 \ \Rightarrow \ \dfrac{1}{1 + ||A||} \le ||(I_m - A)^{-1}|| \le \dfrac{1}{1 - ||A||}$.

(j) $||A|| < 1$

$$\Rightarrow \ \frac{||I_m||}{||I_m|| + ||A||} \le ||(I_m - A)^{-1}|| \le \frac{||I_m|| - (||I_m|| - 1)||A||}{1 - ||A||}.$$

(10) $||\cdot||$ a matrix norm for $(m \times m)$ matrices, $A, B\ (m \times m)$:

(a) $||AB - I_m|| < 1 \ \Rightarrow \ A, B$ nonsingular.

(b) A nonsingular, B singular $\Rightarrow \ ||A - B|| \ge \dfrac{1}{||A^{-1}||}$.

(11) $||\cdot||$ a matrix norm for $(m \times m)$ matrices, $A, B\ (m \times m)$, A nonsingular, $A + B$ nonsingular:

(a) $||A^{-1} - (A + B)^{-1}|| \le ||A^{-1}||\,||(A + B)^{-1}||\,||B||$.

(b) $||A^{-1}B|| < 1 \ \Rightarrow \ \dfrac{||A^{-1} - (A + B)^{-1}||}{||A^{-1}||} \le \dfrac{||A^{-1}B||}{1 - ||A^{-1}B||}$.

(c) $||A^{-1}B|| < 1$ and $||B|| < \dfrac{1}{||A^{-1}||}$

$$\Rightarrow \ \frac{||A^{-1} - (A + B)^{-1}||}{||A^{-1}||} \le \frac{||A^{-1}||\,||B||}{1 - ||A^{-1}||\,||B||}.$$

(12) $||\cdot||$ a matrix norm for $(m \times m)$ matrices: $||\cdot||$ is unitarily invariant $\Rightarrow \ ||A||_{spec} \le ||A||$ for all $(m \times m)$ matrices A.

(13) $A\ (m \times m)$: There exists a matrix norm $||\cdot||$ such that $||A|| < 1 \ \Rightarrow$ $\lim_{i \to \infty} A^i = 0$, i.e., A is convergent.

(14) $A\ (m \times m), p(x) = \sum_{i=0}^{\infty} p_i x^i$ a power series: There exists a matrix norm $||\cdot||$ such that $\sum_{i=0}^{n} |p_i|_{abs} ||A||^i$ converges for $n \to \infty \ \Rightarrow \ p(A) = \sum_{i=0}^{\infty} p_i A^i$ exists.

(15) $A\ (m \times m), \epsilon \in \mathbb{R}, \epsilon > 0, \rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda \text{ is eigenvalue of } A\}$: There exists a matrix norm $||\cdot||$ such that $\rho(A) \le ||A|| \le \rho(A) + \epsilon$.

(16) $A\ (m \times m)$:

$$\begin{aligned} &\inf\{||A|| : ||\cdot|| \text{ is a matrix norm for } (m \times m) \text{ matrices}\} \\ &\quad = \rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda \text{ is eigenvalue of } A\}. \end{aligned}$$

Note: The foregoing results may be found in Horn & Johnson (1985, Chapter 5).

8.4.2 Induced Matrix Norms

(1) $\|\cdot\|$ a norm for $(m \times 1)$ vectors:

$$\|A\|_{lub} \equiv \max\left\{\frac{\|Ax\|}{\|x\|} \; : \; x\ (m \times 1),\ x \neq 0\right\}$$

is a matrix norm for $(m \times m)$ matrices.

(2) $\|\cdot\|$ a norm for $(m \times 1)$ vectors with induced matrix norm $\|\cdot\|_{lub}$, $A\ (m \times m)$:

$$\begin{aligned}
\|A\|_{lub} &= \max\{\|Ax\| : x\ (m \times 1),\ \|x\| = 1\} \\
&= \max\{\|Ax\| : x\ (m \times 1),\ \|x\| \leq 1\} \\
&= \max\left\{\frac{\|Ax\|}{\|x\|} : x\ (m \times 1),\ \|x\|_a = 1\right\} \\
&\qquad \text{for any norm } \|\cdot\|_a \text{ for } (m \times 1) \text{ vectors }.
\end{aligned}$$

(3) $\|\cdot\|_{lub}$ an induced matrix norm for $(m \times m)$ matrices: $\|I_m\|_{lub} = 1$.

(4) $\|\cdot\|_{lub}^{(a)}, \|\cdot\|_{lub}^{(b)}$ induced matrix norms:

$$\max\left\{\frac{\|A\|_{lub}^{(a)}}{\|A\|_{lub}^{(b)}} : A\ (m \times m),\ A \neq 0\right\}$$
$$= \max\left\{\frac{\|A\|_{lub}^{(b)}}{\|A\|_{lub}^{(a)}} : A\ (m \times m),\ A \neq 0\right\}.$$

(5) $\|\cdot\|_{lub}^{(a)}, \|\cdot\|_{lub}^{(b)}$ matrix norms induced by norms $\|\cdot\|_a, \|\cdot\|_b$, respectively, for $(m \times 1)$ vectors: $\|A\|_{lub}^{(a)} = \|A\|_{lub}^{(b)}$ for every $(m \times m)$ matrix $A \iff$ there exists a $c \in \mathbb{R}, c > 0$, such that $\|x\|_a = c\|x\|_b$ for all $(m \times 1)$ vectors x.

(6) $\|\cdot\|_{lub}^{(a)}, \|\cdot\|_{lub}^{(b)}$ induced matrix norms: $\|A\|_{lub}^{(a)} \leq \|A\|_{lub}^{(b)}$ for all $(m \times m)$ matrices $A \iff \|A\|_{lub}^{(a)} = \|A\|_{lub}^{(b)}$ for all $(m \times m)$ matrices A.

(7) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices: There exists an induced matrix norm $\|\cdot\|_{lub}$ such that $\|A\| \geq \|A\|_{lub}$ for every $(m \times m)$ matrix A.

(8) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices, $\|\cdot\|_{lub}$ an induced matrix norm for $(m \times m)$ matrices: $\|A\| \leq \|A\|_{lub}$ for all $(m \times m)$ matrices $A \iff \|A\| = \|A\|_{lub}$ for all $(m \times m)$ matrices A.

(9) $\|\cdot\|$ a matrix norm for $(m \times m)$ matrices: $\|\cdot\|$ is a minimal matrix norm $\iff \|\cdot\|$ is an induced matrix norm.

(10) $\|\cdot\|$ a norm for $(m \times 1)$ vectors with induced matrix norm $\|\cdot\|_{lub}$: $\|\cdot\|$ is an absolute norm $\iff$ $\|D\|_{lub} = \max_{i=1,\ldots,m} |d_i|_{\text{abs}}$ for every diagonal matrix $D = \text{diag}(d_1, \ldots, d_m)$.

(11) $\|\cdot\|$ a norm for $(m \times 1)$ vectors with induced matrix norm $\|\cdot\|_{lub}$, S $(m \times m)$ nonsingular, $\|x\|_{S^{-1}} \equiv \|S^{-1}x\|$, $\|A\|_{lub}^{(S)} \equiv \|S^{-1}AS\|_{lub}$: $\|\cdot\|_{lub}^{(S)}$ is the matrix norm induced by $\|\cdot\|_{S^{-1}}$.

Note: The foregoing results on induced matrix norms may be found in Chapter 5 of Horn & Johnson (1985).

8.5 Properties of Special Norms

8.5.1 General Results

l_1-norm

(1) $\|\cdot\|_1$ is a matrix norm.

(2) $\|\cdot\|_1$ is not derived from an inner product.

(3) x $(m \times 1)$: $\|x\|_1 = \max\{|x^H y|_{\text{abs}} : y\ (m \times 1), \|y\|_\infty = 1\}$.

(4) $A\ (m \times m)$: $\|A^H\|_1 = \|A\|_1$.

Euclidean norm

(5) $\|\cdot\|_2$ is a matrix norm for $(m \times m)$ matrices.

(6) $\|\cdot\|_2$ is derived from the inner product $\text{tr}(A^H B)$ for complex $(m \times n)$ matrices A, B or $\text{tr}(A'B)$ for real $(m \times n)$ matrices A, B.

(7) $A\ (m \times n)$: $\|A^H\|_2 = \|A\|_2$.

(8) U $(m \times m)$ unitary, A $(m \times m)$: $\|UA\|_2 = \|A\|_2$.

(9) U $(m \times m)$ unitary, V $(n \times n)$ unitary, A $(m \times n)$: $\|UAV\|_2 = \|A\|_2$.

(10) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:

$$A \text{ normal} \Rightarrow \|A\|_2^2 = \sum_{i=1}^{m} |\lambda_i|_{\text{abs}}^2.$$

Maximum norm

(11) $\|\cdot\|_\infty$ is not a matrix norm.

(12) $\|A\| \equiv m\|A\|_\infty$ is a matrix norm for $(m \times m)$ matrices.

(13) $A\ (m \times m)$: $\quad ||A||_\infty = \lim_{p\to\infty} ||A||_p$.

Spectral norm

(14) $||\cdot||_{spec}$ is a matrix norm for $(m \times m)$ matrices induced by the vector norm $||\cdot||_2$.

(15) $U\ (m \times m)$ unitary, $V\ (n \times n)$ unitary, $A\ (m \times n)$:
$||UAV||_{spec} = ||A||_{spec}$.

(16) $||\cdot||_{spec}$ is the only induced unitarily invariant matrix norm.

(17) $||\cdot||$ a matrix norm for $(m \times m)$ matrices: $\quad ||\cdot||$ is unitarily invariant $\Rightarrow\ ||A||_{spec} \le ||A||$ for every $(m \times m)$ matrix A.

(18) $A\ (m \times m)$: $\quad ||A^H||_{spec} = ||A||_{spec}$.

(19) $||\cdot||_{spec}$ is the only induced, self-adjoint matrix norm.

(20) $A\ (m \times m)$:

$$\begin{aligned}
||A||_{spec} &= \max\{||Ax||_2 : x\ (m \times 1),\ ||x||_2 = 1\} \\
&= \max\{||Ax||_2 : x\ (m \times 1),\ ||x||_2 \le 1\} \\
&= \max\left\{\frac{||Ax||_2}{||x||_2} : x\ (m \times 1),\ x \neq 0\right\} \\
&= \max\left\{|y^H Ax|_{\text{abs}} : x, y\ (m \times 1),\ ||x||_2 = ||y||_2 = 1\right\} \\
&= \max\left\{|y^H Ax|_{\text{abs}} : x, y\ (m \times 1),\ ||x||_2 = ||y||_2 \le 1\right\}.
\end{aligned}$$

(21) $A\ (m \times m)$: $\quad ||AA^H||_{spec} = ||A^H A||_{spec} = ||A||^2_{spec}$.

(22) $A\ (m \times n), B\ (m \times p)$: $\quad ||A : B||^2_{spec} \le ||A||^2_{spec} + ||B||^2_{spec}$.

Column-sum norm

(23) $||\cdot||_{col}$ is a matrix norm for $(m \times m)$ matrices.

(24) $||\cdot||_{col}$ is induced by $||\cdot||_1$, that is,

$$||A||_{col} = \max\left\{\frac{||Ax||_1}{||x||_1} : x\ (m \times 1),\ x \neq 0\right\}.$$

Row-sum norm

(25) $||\cdot||_{row}$ is a matrix norm for $(m \times m)$ matrices.

(26) $||\cdot||_{row}$ is induced by $||\cdot||_\infty$, that is,

$$||A||_{row} = \max\left\{\frac{||Ax||_\infty}{||x||_\infty} : x\ (m \times 1),\ x \neq 0\right\}.$$

Note: The foregoing properties of special norms may be found in Horn & Johnson (1985, Chapter 5).

8.5.2 Inequalities

(1) $A\ (m \times n)$:

(a) $\|A\|_1 \le \sqrt{mn}\,\|A\|_2$.

(b) $\|A\|_1 \le mn\,\|A\|_\infty$.

(c) $\|A\|_1 \le n\,\|A\|_{col}$.

(d) $\|A\|_1 \le m\,\|A\|_{row}$.

(2) $A\ (m \times m)$: $\quad \|A\|_1 \le m^{3/2}\,\|A\|_{spec}$.

(3) (Hölder's inequality)
$A, B\ (m \times m), 1 \le p, q < \infty, \frac{1}{p} + \frac{1}{q} = 1$: $\quad \|AB\|_1 \le \|A\|_p\,\|B\|_q$.

(4) $A\ (m \times n)$:

(a) $\|A\|_2 \le \|A\|_1$.

(b) $\|A\|_2 \le \sqrt{mn}\,\|A\|_\infty$.

(c) $\|A\|_2 \le \sqrt{n}\,\|A\|_{col}$.

(d) $\|A\|_2 \le \sqrt{m}\,\|A\|_{row}$.

(e) $\|A\|_2 \le \sqrt{\min(m, n)}\,\|A\|_{spec}$.

(5) (Cauchy–Schwarz inequality)
$A, B\ (m \times n)$:

(a) $\|A^H B\|_2 \le \|A\|_2\,\|B\|_2$.

(b) $|\mathrm{tr}(A^H B)|_{\mathrm{abs}} \le \|A\|_2\,\|B\|_2$.

(6) $A, B\ (m \times m)$:

(a) $\|AB\|_2 \le \|A\|_{spec}\,\|B\|_2$.

(b) $\|AB\|_2 \le \|A\|_2\,\|B\|_{spec}$.

(7) $A\ (m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$: $\quad \|A\|_2^2 \ge \sum_{i=1}^{m} |\lambda_i|_{\mathrm{abs}}^2$.

(8) $A\ (m \times n)$:

(a) $\|A\|_\infty \le \|A\|_1$.

(b) $\|A\|_\infty \le \|A\|_2$.

(c) $\|A\|_\infty \le \|A\|_{col}$.

(d) $\|A\|_\infty \le \|A\|_{row}$.

(e) $\|A\|_\infty \le \|A\|_{spec}$.

(9) A $(m \times n)$:

(a) $\|A\|_{col} \leq \|A\|_1$.

(b) $\|A\|_{col} \leq \sqrt{n}\,\|A\|_2$.

(c) $\|A\|_{col} \leq m\,\|A\|_\infty$.

(d) $\|A\|_{col} \leq m\,\|A\|_{row}$.

(10) A $(m \times m)$: $\|A\|_{col} \leq \sqrt{m}\,\|A\|_{spec}$.

(11) A $(m \times n)$:

(a) $\|A\|_{row} \leq \|A\|_1$.

(b) $\|A\|_{row} \leq \sqrt{m}\,\|A\|_2$.

(c) $\|A\|_{row} \leq n\,\|A\|_\infty$.

(d) $\|A\|_{row} \leq n\,\|A\|_{col}$.

(12) A $(m \times m)$: $\|A\|_{row} \leq \sqrt{m}\,\|A\|_{spec}$.

(13) A $(m \times n)$:

(a) $\|A\|_{spec} \leq \|A\|_1$.

(b) $\|A\|_{spec} \leq \|A\|_2$.

(c) $\|A\|_{spec} \leq \sqrt{mn}\,\|A\|_\infty$.

(d) $\|A\|_{spec} \leq \sqrt{n}\,\|A\|_{col}$.

(e) $\|A\|_{spec} \leq \sqrt{m}\,\|A\|_{row}$.

(14) A $(m \times m)$: $\|A\|_{spec}^2 \leq \|A\|_{col}\,\|A\|_{row}$.

Note: The foregoing inequalities can be derived from results presented in Horn & Johnson (1985, Chapter 5), where also further references are given.

9

Properties of Special Matrices

9.1 Circulant Matrices

Definition: An $(m \times m)$ matrix

$$\operatorname{circ}(a_1, \ldots, a_m) \equiv \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{m-1} & a_m \\ a_m & a_1 & a_2 & \cdots & a_{m-2} & a_{m-1} \\ a_{m-1} & a_m & a_1 & \cdots & a_{m-3} & a_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_3 & a_4 & a_5 & \cdots & a_1 & a_2 \\ a_2 & a_3 & a_4 & \cdots & a_m & a_1 \end{bmatrix}$$

is a circulant matrix.

Properties

(1) $a_1, \ldots, a_m, b_1, \ldots, b_m, c_1, c_2 \in \mathbb{C}$:

$$\begin{aligned} &c_1 \operatorname{circ}(a_1, \ldots, a_m) \pm c_2 \operatorname{circ}(b_1, \ldots, b_m) \\ &\quad = \operatorname{circ}(c_1 a_1 \pm c_2 b_1, \ldots, c_1 a_m \pm c_2 b_m). \end{aligned}$$

(2) $A = \operatorname{circ}(a_1, \ldots, a_m)$:

(a) $A' = \operatorname{circ}(a_1, a_m, a_{m-1}, \ldots, a_2)$.

(b) $\bar{A} = \operatorname{circ}(\bar{a}_1, \ldots, \bar{a}_m)$.

(3) $A = \operatorname{circ}(a_1, \ldots, a_m)$, $a(x) = a_1 + a_2 x + \cdots + a_m x^{m-1}$, $\theta = \exp(2\pi i/m)$:

(a) $\operatorname{tr}(A) = m a_1$.

(b) $\det(A) = \displaystyle\prod_{j=0}^{m-1} a(\theta^j)$.

(c) $\lambda_j = a(\theta^j)$, $j = 0, \ldots, m-1$, are the eigenvalues of A.

(4) A $(m \times m)$ circulant:

(a) A^H is circulant.

(b) A^i is circulant for $i = 0, 1, 2, \ldots$

(c) A^{-1} is circulant if A is nonsingular.

(d) A^+ is circulant.

(e) $\mathrm{rk}(A^i) = \mathrm{rk}(A)$ for $i = 1, 2, \ldots$

(f) A is normal, that is, $A^H A = AA^H$.

(g) A is simple.

(h) A is a Toeplitz matrix.

(i) A is unitarily similar to a diagonal matrix.

(5) A, B $(m \times m)$ circulant:

(a) AB is circulant.

(b) $AB = BA$.

(c) $A \odot B$ is circulant.

(6) A, B $(m \times m)$ circulant with eigenvalues $\lambda_1(A), \ldots, \lambda_m(A)$ and $\lambda_1(B)$, $\ldots$, $\lambda_m(B)$, respectively:

(a) $\lambda_1(A)+\lambda_1(B), \ldots, \lambda_m(A)+\lambda_m(B)$ are the eigenvalues of $A+B$.

(b) $\lambda_1(A)-\lambda_1(B), \ldots, \lambda_m(A)-\lambda_m(B)$ are the eigenvalues of $A-B$.

(c) $\lambda_1(A)\lambda_1(B), \ldots, \lambda_m(A)\lambda_m(B)$ are the eigenvalues of AB.

(7) A $(m \times m)$ circulant with eigenvalues $\lambda_1, \ldots, \lambda_m$:

(a) A is Hermitian $\iff$ $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, m$.

(b) A is Hermitian positive definite $\iff$ $\lambda_i \in \mathbb{R}$, $\lambda_i > 0$, $i = 1, \ldots, m$.

(c) A is unitary $\iff$ $|\lambda_i|_{\text{abs}} = 1$, $i = 1, \ldots, m$.

(d) $A = F^H \mathrm{diag}(\lambda_1, \ldots, \lambda_m) F$, where F is a Fourier matrix (see the Appendix for a definition).

(e) $\lim\limits_{n \to \infty} A^n$ exists $\iff$ $\lambda_i = 1$ or $|\lambda_i|_{\text{abs}} < 1$, $i = 1, \ldots, m$.

(f) $\lim\limits_{n \to \infty} \frac{1}{n}(I_m + A + \cdots + A^{n-1})$ exists $\iff$ $|\lambda_i|_{\text{abs}} \leq 1$, $i = 1, \ldots, m$.

(8) $C = \mathrm{circ}(0, 1, 0, \ldots, 0) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (m \times m):$

- (a) $C^m = I_m$.
- (b) $C' = C^{m-1}$.
- (c) $C^{-1} = C^{m-1}$.
- (d) $C^{-1} = C'$.
- (e) $A = \mathrm{circ}(a_1, \ldots, a_m) \iff A = a_1 I_m + a_2 C + \cdots + a_m C^{m-1}$.
- (f) $A\ (m \times m)$ is circulant $\iff AC = CA$.

(9) $I_m = \mathrm{circ}(1, 0, \ldots, 0)$.

Note: For proofs and more on circulant matrices see Davis (1979).

9.2 Commutation Matrices

Definition: The $(mn \times mn)$ commutation matrix K_{mn} or $K_{m,n}$ is defined such that, for any $(m \times n)$ matrix A, $K_{mn}\mathrm{vec}(A) = \mathrm{vec}(A')$. For instance,

$$K_{32} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

so that, for a (3×2) matrix $A = [a_{ij}]$,

$$K_{32}\,\mathrm{vec}\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}\right) = K_{32} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{32} \end{bmatrix} = \mathrm{vec}(A').$$

9.2.1 General Properties

(1) $H_{kl} = [h_{ij}]\ (m \times n)$ with $h_{kl} = 1$ and $h_{ij} = 0$ if $i \neq k$ or $j \neq l$:

$$K_{mn} = \sum_{k=1}^{m} \sum_{l=1}^{n} H_{kl} \otimes H'_{kl}.$$

(2) K_{mn} is a real matrix.

(3) K_{mn} is a permutation matrix.

(4) $K_{m1} = I_m$.

(5) $K_{1n} = I_n$.

(6) $K'_{mn} = K_{nm}$.

(7) $K_{mn}^{-1} = K_{nm}$.

(8)
$$\begin{aligned} K_{m\cdot n,p} &= K_{m,np} K_{n,pm} \\ &= K_{n,pm} K_{m,np} \\ &= (I_m \otimes K_{n,p})(K_{m,p} \otimes I_n) \\ &= (I_n \otimes K_{m,p})(K_{n,p} \otimes I_m). \end{aligned}$$

(9) $K_{m\,n,p} K_{n\,p,m} K_{p\,m,n} = I_{m\,n\,p}$.

(10) $\mathrm{rk}(K_{mn}) = mn$.

(11) $\mathrm{tr}(K_{mn}) = 1 +$ greatest common divisor of $m - 1$ and $n - 1$.

(12) $\mathrm{tr}(K_{mm}) = m$.

(13) $\det(K_{mn}) = (-1)^{mn(m-1)(n-1)/4}$.

(14) $\det(K_{mm}) = (-1)^{m(m-1)/2}$.

(15) The eigenvalues of K_{mm} are 1 and -1 with multiplicities $\frac{1}{2}m(m+1)$ and $\frac{1}{2}m(m-1)$, respectively.

(16) $(I_{m^2} + K_{mm})' = I_{m^2} + K_{mm}$.

(17) $\frac{1}{2}(I_{m^2} + K_{mm})$ is idempotent.

(18) $\mathrm{rk}(I_{m^2} + K_{mm}) = \frac{1}{2}m(m+1)$.

(19) $\mathrm{tr}(I_{m^2} + K_{mm}) = m(m+1)$.

(20) $K_{mm}(I_{m^2} + K_{mm}) = I_{m^2} + K_{mm}$.

(21) $(I_{m^2} + K_{mm})K_{mm} = I_{m^2} + K_{mm}$.

(22) $(I_{m^2} - K_{mm})' = I_{m^2} - K_{mm}$.

(23) $\frac{1}{2}(I_{m^2} - K_{mm})$ is idempotent.

(24) $\mathrm{rk}(I_{m^2} - K_{mm}) = \frac{1}{2}m(m-1)$.

(25) $\operatorname{tr}(I_{m^2} - K_{mm}) = m(m-1)$.

9.2.2 Kronecker Products

(1) $A\ (m \times n)$:

(a) $K_{mn}(A' \otimes A)$ is symmetric.

(b) $[K_{mn}(A' \otimes A)]^2 = AA' \otimes A'A$.

(c) $\operatorname{rk}(A) = r \ \Rightarrow\ \operatorname{rk}[K_{mn}(A' \otimes A)] = r^2$.

(2) $A\ (m \times m)$:

(a) $$\begin{aligned}(I_{m^2} + K_{mm})(A \otimes A)(I_{m^2} + K_{mm}) &= 2(I_{m^2} + K_{mm})(A \otimes A)\\ &= 2(A \otimes A)(I_{m^2} + K_{mm}).\end{aligned}$$

(b) $$\begin{aligned}(I_{m^2} - K_{mm})(A \otimes A)(I_{m^2} - K_{mm}) &= 2(I_{m^2} - K_{mm})(A \otimes A)\\ &= 2(A \otimes A)(I_{m^2} - K_{mm}).\end{aligned}$$

(3) $A\ (m \times n), a\ (p \times 1), b\ (1 \times p)$:
$$\begin{aligned}K_{pm}(A \otimes a) &= a \otimes A,\\ K_{mp}(a \otimes A) &= A \otimes a,\\ (A \otimes b)K_{np} &= b \otimes A,\\ (b \otimes A)K_{pn} &= A \otimes b.\end{aligned}$$

(4) $A\ (m \times m), a\ (m \times 1)$:

(a) $$\begin{aligned}(I_{m^2} + K_{mm})(A \otimes a) &= (I_{m^2} + K_{mm})(a \otimes A)\\ &= A \otimes a + a \otimes A.\end{aligned}$$

(b) $$\begin{aligned}(I_{m^2} - K_{mm})(A \otimes a) &= -(I_{m^2} - K_{mm})(a \otimes A)\\ &= A \otimes a - a \otimes A.\end{aligned}$$

(5) $A\ (m \times n), B\ (p \times q)$:

(a) $K_{pm}(A \otimes B) = (B \otimes A)K_{qn}$.

(b) $K_{pm}(A \otimes B)K_{nq} = B \otimes A$.

(c) $\operatorname{vec}(A \otimes B) = (I_n \otimes K_{qm} \otimes I_p)[\operatorname{vec}(A) \otimes \operatorname{vec}(B)]$.

(6) $A, B\ (m \times n)$: $\quad \operatorname{tr}[K_{mn}(A' \otimes B)] = \operatorname{tr}(A'B)$.

(7) $A, B\ (m \times m)$:

(a) $$\begin{aligned}&(I_{m^2} + K_{mm})(A \otimes B)(I_{m^2} + K_{mm})\\ &\quad = (I_{m^2} + K_{mm})(B \otimes A)(I_{m^2} + K_{mm}).\end{aligned}$$

(b) $$\begin{aligned}&(I_{m^2} - K_{mm})(A \otimes B)(I_{m^2} - K_{mm})\\ &\quad = (I_{m^2} - K_{mm})(B \otimes A)(I_{m^2} - K_{mm}).\end{aligned}$$

(c) $$\begin{aligned}&(I_{m^2} + K_{mm})(A \otimes B + B \otimes A)(I_{m^2} + K_{mm})\\ &\quad = 2(I_{m^2} + K_{mm})(A \otimes B)(I_{m^2} + K_{mm})\\ &\quad = 2(I_{m^2} + K_{mm})(A \otimes B + B \otimes A)\end{aligned}$$

$$= 2(A \otimes B + B \otimes A)(I_{m^2} + K_{mm}).$$

(d) $(I_{m^2} - K_{mm})(A \otimes B + B \otimes A)(I_{m^2} - K_{mm})$
$$= 2(I_{m^2} - K_{mm})(A \otimes B)(I_{m^2} - K_{mm})$$
$$= 2(I_{m^2} - K_{mm})(A \otimes B + B \otimes A)$$
$$= 2(A \otimes B + B \otimes A)(I_{m^2} - K_{mm}).$$

(8) $A\ (m \times n), B\ (p \times q), C\ (q \times s), D\ (n \times r)$:

$$K_{mp}(BC \otimes AD) = (A \otimes B)K_{nq}(C \otimes D).$$

(9) $A\ (m \times n), B\ (p \times q), C\ (r \times s)$:

$$\begin{aligned} A \otimes B \otimes C &= K_{m\ p,r}(C \otimes A \otimes B)K_{s,q\ n} \\ &= K_{m,p\ r}(B \otimes C \otimes A)K_{s \cdot q,n}. \end{aligned}$$

(10) $A\ (m \times n)$ real, $\text{rk}(A) = r$: $\lambda_1, \ldots, \lambda_r > 0$ are eigenvalues of $A'A \Rightarrow \lambda_1, \ldots, \lambda_r$ and $\pm\sqrt{\lambda_i \lambda_j}$, $i < j$, are the nonzero eigenvalues of $K_{mn}(A' \otimes A)$.

9.2.3 Relations With Duplication and Elimination Matrices

(1) $K_{mm} = 2D_m D_m^+ - I_{m^2}$.

(2) $K_{mm} D_m = D_m$.

(3) $D_m^+ K_{mm} = D_m^+$.

(4) $I_{m^2} + K_{mm} = D_m L_m (I_{m^2} + K_{mm})$.

(5) $D_m^+ = \frac{1}{2} L_m (I_{m^2} + K_{mm})$.

(6) $L_m K_{mm} L_m' = 2I_{\frac{1}{2}m(m+1)} - D_m' D_m$.

(7) $D_m = (I_{m^2} + K_{mm})L_m'(L_m(I_{m^2} + K_{mm})L_m')^{-1}$.

(8) $L_m K_{mm} L_m'$ is idempotent diagonal.

(9) $\text{rk}(L_m K_{mm} L_m') = m$.

(10) $\text{tr}(L_m K_{mm} L_m') = m$.

(11) $L_m' L_m K_{mm} L_m' L_m$ is idempotent diagonal.

(12) $\text{rk}(L_m' L_m K_{mm} L_m' L_m) = m$.

(13) $\text{tr}(L_m' L_m K_{mm} L_m' L_m) = m$.

(14) $L_m' L_m K_{mm} L_m' = K_{mm} L_m' L_m K_{mm} L_m'$.

(15) $K_{mm} = D_m D_m' + L_m' L_m K_{mm} L_m' L_m - I_{m^2}$.

(16) $A\ (m \times m)$: $L_m' L_m K_{mm} L_m' L_m \text{vec}(A) = \text{vec}(\text{dg}(A))$.

Note: A substantial collection of results for commutation matrices including the ones given here may be found in Magnus (1988).

9.3 Convergent Matrices

Definition: A $(m \times m)$ is said to be convergent or stable if $A^n \to 0$ as $n \to \infty$.

(1) $A\ (m \times m), B\ (n \times n),\ c \in \mathbb{C}$:

(a) A is convergent, $|c|_{\text{abs}} < 1 \ \Rightarrow\ cA$ is convergent.

(b) A, B are convergent $\Rightarrow$ $A \otimes B$ is convergent.

(c) A, B are convergent $\Rightarrow$ $A \oplus B$ is convergent.

(2) $A\ (m \times m)$ with eigenvalues $\lambda_1, \dots, \lambda_m$:

$$A \text{ is convergent} \iff \max_{i=1,\dots,m} |\lambda_i|_{\text{abs}} < 1.$$

(3) $A = \text{diag}(a_1, \dots, a_m)$:
A is convergent $\iff$ $|a_i|_{\text{abs}} < 1,\ i = 1, \dots, m.$

(4) $A = [a_{ij}]\ (m \times m)$ triangular:
A is convergent $\iff$ $|a_{ii}|_{\text{abs}} < 1,\ i = 1, \dots, m.$

(5) $A\ (m \times m)$:

(a) A is convergent $\iff \displaystyle\sum_{i=1}^{n} A^i$ converges for $n \to \infty$.

(b) $\|A\| < 1$ for some matrix norm $\|\cdot\|$ $\Rightarrow$ A is convergent.

(c) A is convergent $\Rightarrow$ $I_m \pm A$ is nonsingular.

(6) $A\ (m \times m)$ convergent:

(a) $\displaystyle (I_m + A)^{-1} = \sum_{i=0}^{\infty} (-A)^i.$

(b) $\displaystyle (I_m - A)^{-1} = \sum_{i=0}^{\infty} A^i.$

(c) $\displaystyle (I_{m^2} + A \otimes A)^{-1} = \sum_{i=0}^{\infty} (-A)^i \otimes A^i.$

(d) $\displaystyle (I_{m^2} - A \otimes A)^{-1} = \sum_{i=0}^{\infty} A^i \otimes A^i.$

(7) (Stein's theorem)
A $(m \times m)$: A is convergent $\iff$ given a positive definite $(m \times m)$ matrix V, there exists a positive definite matrix Q such that $Q - A^H QA = V$.

Note: These results follow straightforwardly from definitions or may be found in Lancaster & Tismenetsky (1985) or Horn & Johnson (1985).

9.4 Diagonal Matrices

Definition: $A = [a_{ij}]$ $(m \times m)$ is a diagonal matrix or, briefly, A is diagonal, if $a_{ij} = 0$ for $i \neq j$, that is,

$$A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix}.$$

Properties

(1) A, B $(m \times m)$ diagonal, $c \in \mathbb{C}$:

(a) cA is diagonal.

(b) $A \pm B$ is diagonal.

(c) AB is diagonal.

(d) $AB = BA$.

(e) $A \odot B$ is diagonal.

(2) A $(m \times m)$ diagonal, B $(n \times n)$ diagonal:

(a) $A \otimes B$ is diagonal.

(b) $A \oplus B$ is diagonal.

(3) A $(m \times m)$ diagonal:

(a) $\mathrm{rk}(A)$ = number of nonzero diagonal elements.

(b) $A' = A$.

(c) $A^H = A$, if A is real.

(4) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm})$:

(a) $\mathrm{tr}(A) = a_{11} + \cdots + a_{mm}$.

(b) $\det(A) = \prod_{i=1}^{m} a_{ii}$.

(c) $a_{ii} \neq 0,\ i = 1, \ldots, m \ \Rightarrow\ A^{-1} = \mathrm{diag}(a_{11}^{-1}, \ldots, a_{mm}^{-1})$.

(d) $A^{+} = \mathrm{diag}(a_{11}^{+}, \ldots, a_{mm}^{+})$, where $a_{ii}^{+} = \begin{cases} a_{ii}^{-1} & \text{if } a_{ii} \neq 0 \\ 0 & \text{if } a_{ii} = 0 \end{cases}$

(e) $A^{adj} = [a_{ij}^{adj}]$, where $a_{ij}^{adj} = \begin{cases} 0 & \text{for } i \neq j \\ \prod\limits_{\substack{n=1 \\ n \neq i}}^{m} a_{nn} & \text{for } i = j \end{cases}$

(5) A $(m \times m)$ diagonal:

(a) A is idempotent
$\Rightarrow$ all diagonal elements of A are either 0 or 1.

(b) A is nilpotent $\Rightarrow$ $A = O_{m \times m}$.

(c) A is orthogonal $\Rightarrow$ $A = I_m$.

(d) A is unitary $\Rightarrow$ $A = I_m$.

(6) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm})$:

(a) A is positive definite $\Rightarrow$ $0 < a_{ii} \in \mathbb{R},\ i = 1, \ldots, m$.

(b) A is positive semidefinite $\Rightarrow$ $0 \leq a_{ii} \in \mathbb{R},\ i = 1, \ldots, m$.

(c) A is negative definite $\Rightarrow$ $0 > a_{ii} \in \mathbb{R},\ i = 1, \ldots, m$.

(d) A is negative semidefinite $\Rightarrow$ $0 \geq a_{ii} \in \mathbb{R},\ i = 1, \ldots, m$.

(7) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm})$: a_{ii} is an eigenvalue of A and the ith column of I_m is a corresponding eigenvector for $i = 1, \ldots, m$.

(8) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm}), B = \mathrm{diag}(b_{11}, \ldots, b_{mm})$:

(a) $D_m^{+}(A \otimes B)D_m$ is a $(\frac{1}{2}m(m+1) \times \frac{1}{2}m(m+1))$ diagonal matrix with diagonal elements $\frac{1}{2}(a_{ii}b_{jj} + a_{jj}b_{ii}),\ 1 \leq j \leq i \leq m$.

(b) $L_m(A \otimes B)L_m'$ is a $(\frac{1}{2}m(m+1) \times \frac{1}{2}m(m+1))$ diagonal matrix with diagonal elements $a_{ii}b_{jj},\ i \leq j$.

(c) $L_m(A \otimes B)D_m$ is a $(\frac{1}{2}m(m+1) \times \frac{1}{2}m(m+1))$ diagonal matrix with diagonal elements $a_{ii}b_{jj},\ i \leq j$.

(9) $A = \mathrm{diag}(a_{11}, \ldots, a_{mm}),\ B = \mathrm{diag}(b_{11}, \ldots, b_{mm})$:

(a) $\det[L_m(A \otimes B)L_m'] = \prod\limits_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

(b) $\det[L_m(A \otimes B)D_m] = \prod\limits_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

Note: Many of the results of this section follow easily by considering the individual elements of the matrices involved. The others may be found in Magnus (1988) or in Horn & Johnson (1985). Note that a diagonal matrix is a special case of a triangular matrix. For transforming matrices to diagonal form see Chapter 6.

9.5 Duplication Matrices

Definition: The $(m^2 \times \frac{1}{2}m(m+1))$ duplication matrix D_m is defined such that $\text{vec}(A) = D_m \text{vech}(A)$ for any symmetric $(m \times m)$ matrix A. For example.

$$D_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so that

$$\text{vec}\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{12} \\ a_{22} \end{bmatrix} = D_2 \begin{bmatrix} a_{11} \\ a_{12} \\ a_{22} \end{bmatrix} = D_2 \text{vech}\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}.$$

9.5.1 General Properties

(1) D_m is a real matrix.

(2) $\text{rk}(D_m) = \frac{1}{2}m(m+1)$.

(3) $D_m^+ = (D_m' D_m)^{-1} D_m'$.

(4) $D_m^+ D_m = I_{\frac{1}{2}m(m+1)}$.

(5) $D_m^+ D_m^{+\prime} = (D_m' D_m)^{-1}$.

(6) $\det(D_m' D_m) = 2^{m(m-1)/2}$.

(7) $\det(D_m D_m') = 0$.

(8) $\text{tr}(D_m' D_m) = m^2$.

(9) $\text{tr}(D_m' D_m)^{-1} = \frac{1}{4}m(m+3)$.

(10) $D_{m+1}' D_{m+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2I_m & 0 \\ 0 & 0 & D_m' D_m \end{bmatrix}$.

(11) $D^{+}_{m+1}D^{+'}_{m+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}I_m & 0 \\ 0 & 0 & (D'_m D_m)^{-1} \end{bmatrix}$.

(12) $D'_m D_m$ is a $(\frac{1}{2}m(m+1) \times \frac{1}{2}m(m+1))$ diagonal matrix with 1 (m times) and 2 ($\frac{1}{2}m(m-1)$ times) on the diagonal.

(13) $D_m D'_m$ is $(m^2 \times m^2)$ with eigenvalues 0 ($\frac{1}{2}m(m-1)$ times), 1 (m times) and 2 ($\frac{1}{2}m(m-1)$ times).

9.5.2 Relations With Commutation and Elimination Matrices

(1) $D_m D^{+}_m = \frac{1}{2}(I_{m^2} + K_{mm})$.

(2) $K_{mm}D_m = D_m$.

(3) $D^{+}_m K_{mm} = D^{+}_m$.

(4) $L_m D_m = I_{\frac{1}{2}m(m+1)}$.

(5) $D_m L_m(I_{m^2} + K_{mm}) = I_{m^2} + K_{mm}$.

(6) $D^{+}_m = \frac{1}{2}L_m(I_{m^2} + K_{mm})$.

(7) $D'_m D_m = 2I_{\frac{1}{2}m(m+1)} - L_m K_{mm} L'_m$.

(8) $D_m = (I_{m^2} + K_{mm})L'_m[L_m(I_{m^2} + K_{mm})L'_m]^{-1}$.

9.5.3 Expressions With vec and vech Operators

(1) A $(m \times m)$:

$$D^{+}_m \text{vec}(A) = \tfrac{1}{2}\text{vech}(A + A'),$$

$$D^{+}_m \text{vec}(A) = L_m \text{vec}(A) = \text{vech}(A), \text{ if } A \text{ is symmetric},$$

$$D'_m D_m \text{vech}(A) = 2\text{vech}(A) - \text{vech}[\text{dg}(A)],$$

$$D'_m \text{vec}(A) = \text{vech}[A + A' - \text{dg}(A)],$$

$$D_m D'_m \text{vec}(A) = \text{vec}[A + A' - \text{dg}(A)].$$

(2) A $(m \times m), a, b$ $(m \times 1)$, $c \in \mathbb{C}$:

$$D'_{m+1}\text{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ a + b \\ D'_m \text{vec}(A) \end{bmatrix},$$

$$D^{+}_{m+1}\text{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ \frac{1}{2}(a + b) \\ D^{+}_m \text{vec}(A) \end{bmatrix}.$$

(3) A $(m \times m)$ symmetric nonsingular, $c \in \mathbb{C}$:

(a) $\det(D^{+}_m[A \otimes A + c\ \text{vec}(A)\text{vec}(A)']D_m) = (1 + cm)(\det A)^{m+1}$.

(b) $(D_m^+[A \otimes A + c\ \mathrm{vec}(A)\mathrm{vec}(A)']D_m)^{-1}$

$$= D_m^+ \left[A^{-1} \otimes A^{-1} - \frac{c}{1+cm}\mathrm{vec}(A^{-1})\mathrm{vec}(A^{-1})'\right] D_m.$$

9.5.4 Duplication Matrices and Kronecker Products

(1) $A\ (m \times m)$:

(a) $\det[D_m'(A \otimes A)D_m] = 2^{m(m-1)/2}(\det A)^{m+1}$.

(b) $[D_m'(A \otimes A)D_m]^{-1} = D_m^+(A^{-1} \otimes A^{-1})D_m^{+'}$, if $\det(A) \neq 0$.

(c) $\det[D_m^+(A \otimes A)D_m] = (\det A)^{m+1}$.

(d) $D_m D_m^+(A \otimes A)D_m = (A \otimes A)D_m$.

(e) $D_m^+(A \otimes A)D_m D_m^+ = D_m^+(A \otimes A)$.

(f) $[D_m^+(A \otimes A)D_m]^s = D_m^+(A^s \otimes A^s)D_m, \quad s = 0, 1, 2, \ldots$

(g) $[D_m^+(A \otimes A)D_m]^{-s} = D_m^+(A^{-s} \otimes A^{-s})D_m, \ s = 1, 2, \ldots,$ if $\det(A) \neq 0$.

(h) $[D_m^+(A \otimes A)D_m]^{1/2} = D_m^+(A^{1/2} \otimes A^{1/2})D_m$, if $A^{1/2}$ exists.

(2) $A\ (m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$: $\quad \lambda_i\lambda_j,\ 1 \leq j \leq i \leq m$, are the eigenvalues of $D_m^+(A \otimes A)D_m$.

(3) $A, B\ (m \times m)$:

$$D_m^+(A \otimes B)D_m = D_m^+(B \otimes A)D_m,$$

$$D_m^+(A \otimes B)D_m = \tfrac{1}{2}D_m^+(A \otimes B + B \otimes A)D_m,$$

$$\mathrm{tr}[D_m^+(A \otimes B)D_m] = \tfrac{1}{2}[\mathrm{tr}(A)\mathrm{tr}(B) + \mathrm{tr}(AB)],$$

$$\det(B) = 0 \ \Rightarrow\ \det[D_m^+(A \otimes B)D_m] = 0.$$

(4) $A = \mathrm{diag}\,(a_{11}, \ldots, a_{mm}),\ B = \mathrm{diag}\,(b_{11}, \ldots, b_{mm})$: $\quad D_m^+(A \otimes B)D_m$ is a diagonal matrix with diagonal elements $\frac{1}{2}(a_{ii}b_{jj} + a_{jj}b_{ii}),\ 1 \leq j \leq i \leq m$.

(5) $A = [a_{ij}], B = [b_{ij}]\ (m \times m)$ upper (lower) triangular: $D_m^+(A \otimes B)D_m$ is upper (lower) triangular with diagonal elements $\frac{1}{2}(a_{ii}b_{jj} + a_{jj}b_{ii}),\ 1 \leq j \leq i \leq m$.

(6) $A\ (m \times m), B\ (m \times m)$ nonsingular, $\lambda_1, \ldots, \lambda_m$ eigenvalues of AB^{-1} :

$$\det[D_m^+(A \otimes B)D_m] = 2^{-m(m-1)/2}\det(A)(\det B)^m \prod_{i>j}(\lambda_i + \lambda_j).$$

(7) $A, B\ (m \times m)$ with $A \otimes B + B \otimes A$ nonsingular:

$$[D_m^+(A \otimes B)D_m]^{-1} = 2D_m^+(A \otimes B + B \otimes A)^{-1}D_m.$$

(8) $A\ (m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:

(a) $D_m D_m^+(I_m \otimes A + A \otimes I_m)D_m = (I_m \otimes A + A \otimes I_m)D_m$.

(b) $\mathrm{tr}[D_m^+(I_m \otimes A + A \otimes I_m)D_m] = (m+1)\mathrm{tr}(A)$.

(c) $[D_m^+(I_m \otimes A + A \otimes I_m)D_m]^{-1} = D_m^+(I_m \otimes A + A \otimes I_m)^{-1}D_m$, if $I_m \otimes A + A \otimes I_m$ is nonsingular.

(d) $\det[D_m^+(I_m \otimes A + A \otimes I_m)D_m] = 2^m \det(A) \prod_{i>j}(\lambda_i + \lambda_j)$.

(e) $\lambda_i + \lambda_j$, $1 \le i \le j \le m$, are the eigenvalues of $D_m^+(I_m \otimes A + A \otimes I_m)D_m$.

(9) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$, $i \in \mathbb{N}, i > 1$:

$$\det\left(D_m^+ \sum_{j=0}^{i-1}(A^{i-1-j} \otimes A^j)D_m\right) = i^m(\det A)^{i-1} \prod_{k>l} \mu_{kl}$$

where

$$\mu_{kl} = \begin{cases} (\lambda_k^i - \lambda_l^i)/(\lambda_k - \lambda_l) & \text{if} \quad \lambda_k \neq \lambda_l \\ i\lambda_k^{i-1} & \text{if} \quad \lambda_k = \lambda_l \end{cases}.$$

(10) $A = [a_{ij}], B = [b_{ij}]$ $(m \times m)$ lower triangular :

$$\det[D_m^+(A \otimes A \pm B \otimes B)D_m] = \prod_{i \ge j}(a_{ii}a_{jj} \pm b_{ii}b_{jj}).$$

(11) A $(m \times m)$ nonsingular, B $(m \times m)$, $\lambda_1, \ldots, \lambda_m$ eigenvalues of BA^{-1} :

$$\det[D_m^+(A \otimes A \pm B \otimes B)D_m] = (\det A)^{m+1} \prod_{i \ge j}(1 \pm \lambda_i\lambda_j).$$

(12) A, B $(m \times m)$ symmetric, a, b $(m \times 1), \alpha, \beta \in \mathbb{C}$:

(a) $$D'_{m+1}\left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix}\right) D_{m+1} =$$

$$\begin{bmatrix} \alpha\beta & \alpha b' + \beta a' & (a' \otimes b')D_m \\ \alpha b + \beta a & \alpha B + \beta A + ab' + ba' & (a' \otimes B + b' \otimes A)D_m \\ D'_m(a \otimes b) & D'_m(a \otimes B + b \otimes A) & D'_m(A \otimes B)D_m \end{bmatrix}.$$

(b) $$D^+_{m+1}\left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix}\right) D^{+\prime}_{m+1} =$$

$$\begin{bmatrix} \alpha\beta & \frac{1}{2}(\alpha b' + \beta a') & (a' \otimes b')D_m^{+\prime} \\ \frac{1}{2}(\alpha b + \beta a) & \frac{1}{4}(\alpha B + \beta A + ab' + ba') & \frac{1}{2}(a' \otimes B + b' \otimes A)D_m^{+\prime} \\ D_m^+(a \otimes b) & \frac{1}{2}D_m^+(a \otimes B + b \otimes A) & D_m^+(A \otimes B)D_m^{+\prime} \end{bmatrix}.$$

9.5.5 Duplication Matrices, Elimination Matrices and Kronecker Products

(1) $A\ (m \times m)$:

(a) $L_m(A \otimes A)D_m = D_m^+(A \otimes A)D_m$.

(b) $L_m(I_m \otimes A + A \otimes I_m)D_m = D_m^+(I_m \otimes A + A \otimes I_m)D_m$.

(2) $A = \operatorname{diag}(a_{11}, \ldots, a_{mm}), B = \operatorname{diag}(b_{11}, \ldots, b_{mm})$: $L_m(A \otimes B)D_m$ is a diagonal matrix with diagonal elements $a_{ii}b_{jj},\ 1 \leq i \leq j \leq m$.

(3) $A = [a_{ij}], B = [b_{ij}]\ (m \times m)$ upper (lower) triangular:

(a) $L_m(A \otimes B)D_m$ is upper (lower) triangular with diagonal elements $a_{ii}b_{jj},\ 1 \leq i \leq j \leq m$.

(b) $\det[L_m(A \otimes B)D_m] = \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}$.

(4) $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]\ (m \times m)$ lower triangular:

(a) $L_m(A \otimes B')D_m = L_m(A \otimes B')L'_m$.

(b) $$\begin{aligned} \det[L_m(AB' \otimes B'A)L'_m] &= \det[L_m(AB \otimes B'A)D_m] \\ &= \det[L_m(AB' \otimes A'B')D_m] \\ &= (\det A)^{m+1}(\det B)^{m+1}. \end{aligned}$$

(c) $\det[L_m(A \otimes B'C)D_m] = \prod_{i=1}^{m} (b_{ii}c_{ii})^i a_{ii}^{m-i+1}$.

(d) $\det[L_m(AB' \otimes C')D_m] = \prod_{i=1}^{m} c_{ii}^i (a_{ii}b_{ii})^{m-i+1}$.

(5) $A, B\ (m \times m), \det(B) \neq 0$:

$$\det[L_m(A \otimes B)D_m] = \det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)})$$

and

$$\det[L_m(A \otimes B)D_m^{+\prime}] = 2^{-m(m-1)/2}\det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)}),$$

where

$$C_{(n)} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$$

are the principal submatrices of $C = [c_{ij}] = AB^{-1}$.

Note: Magnus (1988) provides a very extensive collection of results on duplication matrices. The results of this section are taken from that source.

9.6 Elimination Matrices

Definition: The $(\frac{1}{2}m(m+1) \times m^2)$ elimination matrix L_m is defined such that $\mathrm{vech}(A) = L_m \mathrm{vec}(A)$ for any $(m \times m)$ matrix A. For example,

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

so that

$$\mathrm{vech}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{22} \end{bmatrix} = L_2 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{bmatrix} = L_2 \mathrm{vec}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

9.6.1 General Properties

(1) L_m is a real matrix.

(2) $\mathrm{rk}(L_m) = \frac{1}{2}m(m+1)$.

(3) $L_m^+ = L_m'$.

(4) $L_m L_m' = I_{\frac{1}{2}m(m+1)}$.

(5) $L_m' L_m$ is idempotent diagonal with diagonal elements 1 and 0.

(6) $\mathrm{rk}(L_m' L_m) = \frac{1}{2}m(m+1)$.

(7) $\mathrm{tr}(L_m' L_m) = \frac{1}{2}m(m+1)$.

(8) $\det(L_m L_m') = 1$.

(9) $\det(L_m' L_m) = 0$.

9.6.2 Relations With Commutation and Duplication Matrices

(1) $L_m K_{mm} L_m'$ is idempotent diagonal.

(2) $\mathrm{rk}(L_m K_{mm} L_m') = m$.

(3) $\mathrm{tr}(L_m K_{mm} L_m') = m$.

(4) $L_m' L_m K_{mm} L_m' L_m$ is idempotent diagonal.

(5) $\mathrm{rk}(L_m' L_m K_{mm} L_m' L_m) = m$.

(6) $\operatorname{tr}(L'_m L_m K_{mm} L'_m L_m) = m.$

(7) $L'_m L_m K_{mm} L'_m = K_{mm} L'_m L_m K_{mm} L'_m.$

(8) $L_m D_m = I_{\frac{1}{2}m(m+1)}.$

(9) $D_m^+ = \frac{1}{2} L_m (I_{m^2} + K_{mm}).$

(10) $D_m L_m (I_{m^2} + K_{mm}) = I_{m^2} + K_{mm}.$

(11) $L_m K_{mm} L'_m + D'_m D_m = 2I_{\frac{1}{2}m(m+1)}.$

(12) $D_m L_m (I_{m^2} + K_{mm}) L'_m = (I_{m^2} + K_{mm}) L'_m.$

(13) $L'_m L_m K_{mm} L'_m L_m = I_{m^2} + K_{mm} - D_m D'_m.$

9.6.3 Expressions With vec and vech Operators

(1) A $(m \times m)$:

(a) $L_m \operatorname{vec}(A) = D_m^+ \operatorname{vec}(A) = \operatorname{vech}(A)$, if A is symmetric.

(b) $\operatorname{vec}(A) = L'_m \operatorname{vech}(A)$, if A is lower triangular.

(c) $L_m K_{mm} L'_m \operatorname{vech}(A) = \operatorname{vech}[\operatorname{dg}(A)]$.

(d) $L'_m L_m K_{mm} L'_m L_m \operatorname{vec}(A) = \operatorname{vec}[\operatorname{dg}(A)]$.

(2) A $(m \times m)$ nonsingular, lower triangular, $c \in \mathbb{C}$:

(a) $\det(L_m[A' \otimes A + c \operatorname{vec}(A)\operatorname{vec}(A')']L'_m) = (1 + cm)(\det A)^{m+1}$.

(b) $\{L_m[A' \otimes A + c \operatorname{vec}(A)\operatorname{vec}(A')']L'_m\}^{-1}$

$$= L_m \left[A'^{-1} \otimes A^{-1} - \frac{c}{1+cm} \operatorname{vec}(A^{-1})\operatorname{vec}(A'^{-1})' \right] L'_m.$$

9.6.4 Elimination Matrices and Kronecker Products

(1) $A = \operatorname{diag}(a_{11}, \ldots, a_{mm})$, $B = \operatorname{diag}(b_{11}, \ldots, b_{mm})$: $L_m(A \otimes B)L'_m$ is diagonal with diagonal elements $a_{ii}b_{jj}$, $i \le j$.

(2) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ upper (lower) triangular:

(a) $L_m(A \otimes B)L'_m$ is upper (lower) triangular with diagonal elements $a_{ii}b_{jj}$, $1 \le i \le j \le m$.

(b) $\det[L_m(A \otimes B)L'_m] = \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}$.

(3) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ lower triangular :

(a) $\det[L_m(A' \otimes B)L_m'] = \prod_{i=1}^{m} b_{ii}^i a_{ii}^{m-i+1}$.

(b) $a_{ii}b_{jj}$, $1 \le i \le j \le m$, are the eigenvalues of $L_m(A' \otimes B)L_m'$.

(c) $[L_m(A' \otimes B)L_m']^s = L_m[(A')^s \otimes B^s]L_m'$ for

$$\begin{cases} s = 0, 1, 2, \ldots \\ s = \ldots -2, -1, \text{ if } A, B \text{ nonsingular} \\ s = \frac{1}{2}, \text{ if lower triangular } A^{1/2}, B^{1/2} \text{ exist} \end{cases}$$

(d) $\det[L_m(AB' \otimes B'A)L_m'] = \det[L_m(AB' \otimes A'B)L_m']$
$= (\det A)^{m+1}(\det B)^{m+1}$.

(e) $L_m' L_m(A' \otimes B)L_m' = (A' \otimes B)L_m'$.

(4) A, B $(m \times m)$: A, B are strictly lower triangular $\Rightarrow$ $L_m(A' \otimes B)L_m'$ is nilpotent.

(5) $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$, $D = [d_{ij}]$ $(m \times m)$ lower triangular :

$$\det[L_m(AB' \otimes C'D)L_m'] = \prod_{i=1}^{m} (c_{ii}d_{ii})^i (a_{ii}b_{ii})^{m-i+1}.$$

$$\det[L_m(A' \otimes B + C' \otimes D)L_m'] = \prod_{i \ge j} (b_{ii}a_{jj} + d_{ii}c_{jj}).$$

(6) A, B, C, D $(m \times m)$ nonsingular, lower triangular :

$$[L_m(AB' \otimes C'D)L_m']^{-1} = L_m[(B')^{-1} \otimes D^{-1}]L_m' L_m[A^{-1} \otimes (C')^{-1}]L_m'.$$

(7) $A = [a_{ij}]$ $(m \times m)$ lower triangular, $n \in \mathbb{N}$:

$$\det\left(L_m \left(\sum_{j=0}^{n-1} (A')^{n-1-j} \otimes A^j \right) L_m' \right) = n^m (\det A)^{n-1} \prod_{k>l} \mu_{kl},$$

where

$$\mu_{kl} = \begin{cases} (a_{kk}^n - a_{ll}^n)/(a_{kk} - a_{ll}) & \text{if } a_{kk} \ne a_{ll} \\ n a_{kk}^{n-1} & \text{if } a_{kk} = a_{ll} \end{cases}$$

(8) $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$:

$$\det[L_m(A \otimes B)L_m'] = \prod_{n=1}^{m} \det(A_{(n)})\det(B^{(n)}),$$

where

$$A_{(n)} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad B^{(n)} = \begin{bmatrix} b_{nn} & \cdots & b_{nm} \\ \vdots & \ddots & \vdots \\ b_{mn} & \cdots & b_{mm} \end{bmatrix},$$

$n = 1, \ldots, m$.

9.6.5 Elimination Matrices, Duplication Matrices and Kronecker Products

(1) $A\ (m \times m)$:

(a) $L_m(A \otimes A)D_m = D_m^+(A \otimes A)D_m$.

(b) $L_m(I_m \otimes A + A \otimes I_m)D_m = D_m^+(I_m \otimes A + A \otimes I_m)D_m$.

(2) $A = \operatorname{diag}(a_{11}, \ldots, a_{mm}),\ B = \operatorname{diag}(b_{11}, \ldots, b_{mm})$: $L_m(A \otimes B)D_m$ is a diagonal matrix with diagonal elements $a_{ii}b_{jj},\ 1 \leq i \leq j \leq m$.

(3) $A = [a_{ij}], B = [b_{ij}]\ (m \times m)$ upper (lower) triangular:

(a) $L_m(A \otimes B)D_m$ is upper (lower) triangular.

(b) $\det[L_m(A \otimes B)D_m] = \prod_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

(4) $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]\ (m \times m)$ lower triangular:

(a) $L_m(A \otimes B')D_m = L_m(A \otimes B')L'_m$.

(b)
$$\begin{aligned}\det[L_m(AB' \otimes B'A)L'_m] &= \det[L_m(AB \otimes B'A)D_m] \\ &= \det[L_m(AB' \otimes A'B')D_m] \\ &= (\det A)^{m+1}(\det B)^{m+1}.\end{aligned}$$

(c) $\det[L_m(A \otimes B'C')D_m] = \prod_{i=1}^{m} (b_{ii}c_{ii})^{i} a_{ii}^{m-i+1}$.

(d) $\det[L_m(AB' \otimes C')D_m] = \prod_{i=1}^{m} c_{ii}^{i} (a_{ii}b_{ii})^{m-i+1}$.

(5) $A, B\ (m \times m), \det(B) \neq 0$:

$$\det[L_m(A \otimes B)D_m] = \det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)})$$

and

$$\det[L_m(A \otimes B)D_m^{+\prime}] = 2^{-m(m-1)/2}\det(A)(\det B)^m \prod_{n=1}^{m-1} \det(C_{(n)}),$$

where

$$C_{(n)} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$$

are the principal submatrices of $C = [c_{ij}] = AB^{-1}$.

Note: The results of this section are taken from Magnus (1988) where also further results on elimination matrices may be found.

9.7 Hermitian Matrices

Definitions: An $(m \times m)$ matrix A is Hermitian if $A^H = A$, that is, the ijth element a_{ij} is the complex conjugate of the jith element, $\bar{a}_{ji} = a_{ij}$, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \bar{a}_{12} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{1m} & \bar{a}_{2m} & \cdots & a_{mm} \end{bmatrix},$$

where the $a_{ii}, i = 1, \ldots, m$, are real.

An $(m \times m)$ matrix A is skew-Hermitian if $A^H = -A$, that is, the ijth element a_{ij} is -1 times the complex conjugate of the jith element, $-\bar{a}_{ji} = a_{ij}$, that is,

$$A = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1m} \\ -\bar{a}_{12} & 0 & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ -\bar{a}_{1m} & -\bar{a}_{2m} & \cdots & 0 \end{bmatrix}.$$

9.7.1 General Results

(1) A $(m \times m)$ real: A is Hermitian $\iff$ A is symmetric.

(2) A, B $(m \times m)$ Hermitian, $c \in \mathbb{R}$:

(a) cA is Hermitian.

(b) $A \pm B$ is Hermitian.

(c) $AB = BA \Rightarrow AB$ is Hermitian.

(d) $A \odot B$ is Hermitian.

(3) A $(m \times m)$ Hermitian, B $(n \times n)$ Hermitian:

(a) $A \otimes B$ is Hermitian.

(b) $A \oplus B$ is Hermitian.

(4) $A = [a_{ij}]$ $(m \times m)$ Hermitian:

(a) A' is Hermitian.

(b) A^H is Hermitian.

(c) A^{-1} is Hermitian, if A is nonsingular.

(d) A^i is Hermitian for $i = 1, 2, \ldots$

(e) $\bar{A}$ is Hermitian.

(f) $|A|_{\text{abs}}$ is Hermitian.

(g) $\begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix}$ is Hermitian for $i = 1, 2, \dots, m$.

(h) $A^H A = AA^H$, that is, A is normal.

(5) A $(m \times m)$:

(a) $A + A^H$ is Hermitian.

(b) $A - A^H$ is skew-Hermitian.

(c) AA^H is Hermitian.

(d) $A^H A$ is Hermitian.

(e) There exist unique Hermitian $(m \times m)$ matrices B, C such that $A = B + \mathrm{i}C$.

(6) A $(m \times m)$:

(a) A is Hermitian $\iff$ $\mathrm{i}A$ is skew-Hermitian.

(b) A is Hermitian $\iff$ $x^H A x \in \mathbb{R}$ for all $(m \times 1)$ vectors x.

(c) A is Hermitian $\iff$ $B^H AB$ Hermitian for all $(m \times m)$ matrices B.

(d) A is Hermitian $\iff$ $A^H A = AA^H$ and all eigenvalues of A are real numbers.

(e) A is Hermitian $\iff$ there exists a unitary $(m \times m)$ matrix U and a real diagonal $(m \times m)$ matrix Λ such that $A = U \Lambda U^H$.

(7) (Householder transformation)

x $(m \times 1)$: $\quad I_m - \dfrac{2xx^H}{x^H x}$ is Hermitian.

(8) A $(m \times m)$, B $(m \times n)$: $\quad A$ is Hermitian $\Rightarrow$ $B^H AB$ is Hermitian.

(9) A, B $(m \times m)$ Hermitian: $\quad AB = BA \iff$ there exists a unitary $(m \times m)$ matrix U such that UAU^H and UBU^H are diagonal matrices.

(10) A $(m \times m)$: $\quad A$ is similar to a real $(m \times m)$ matrix $B \iff$ there exist Hermitian $(m \times m)$ matrices H and C, one of which is nonsingular, such that $A = HC$.

(11) A, B $(m \times m)$ Hermitian:

(a) $[\mathrm{tr}(AB)]^2 \le \mathrm{tr}(A^2 B^2)$.

(b) $\mathrm{rk}(A) \ge \dfrac{(\mathrm{tr}\, A)^2}{\mathrm{tr}(A^2)}$, if $\mathrm{tr}(A^2) \ne 0$.

(c) $\mathrm{rk}(A \odot B) \leq \mathrm{rk}(A)\mathrm{rk}(B)$.

(12) A $(m \times m), b$ $(m \times 1), c \in \mathbb{R}$:

$$B = \left[\begin{array}{cc} c & b^H \\ b & A \end{array} \right] \text{ is Hermitian } \Rightarrow \det(B) = c\det(A) - b^H A^{adj} b.$$

Note: Many of these results are elementary. The others can be found in Horn & Johnson (1985).

9.7.2 Eigenvalues of Hermitian Matrices

Notation: $\lambda(A)$ denotes an eigenvalue of the matrix A.

$\lambda_{min}(A)$ is the smallest eigenvalue of the matrix A.

$\lambda_{max}(A)$ is the largest eigenvalue of the matrix A.

(1) A $(m \times m)$ Hermitian:

(a) All eigenvalues of A are real numbers.

(b) A is positive definite $\Longleftrightarrow$ all eigenvalues of A are real and greater than 0.

(c) A is positive semidefinite $\Longleftrightarrow$ all eigenvalues of A are real and greater than or equal to 0.

(2) (Rayleigh–Ritz theorem)
A $(m \times m)$ Hermitian:

$$\lambda_{min}(A) = \min\left\{ \frac{x^H Ax}{x^H x} \; : \; x \; (m \times 1), \; x \neq 0 \right\},$$

$$\lambda_{max}(A) = \max\left\{ \frac{x^H Ax}{x^H x} \; : \; x \; (m \times 1), \; x \neq 0 \right\}.$$

(3) (Courant–Fischer theorem)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, $1 < i < m$:

$$\lambda_i = \min_{\substack{y_1, \ldots, y_{m-i} \\ (m \times 1)}} \max\left\{ \frac{x^H Ax}{x^H x} : \right.$$
$$\left. x \; (m \times 1), \; x \neq 0, \; x^H y_j = 0, \; j = 1, \ldots, m - i \right\}$$

and

$$\lambda_i = \max_{\substack{y_1, \ldots, y_{i-1} \\ (m \times 1)}} \min\left\{ \frac{x^H Ax}{x^H x} : \right.$$
$$\left. x \; (m \times 1), \; x \neq 0, \; x^H y_j = 0, \; j = 1, \ldots, i - 1 \right\}.$$

(4) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\operatorname{tr}(X^H AX) \ : \ X \ (m \times n), \ X^H X = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $X = [v_1, \ldots, v_n]$.

$$\max\{\operatorname{tr}(X^H AX) \ : \ X \ (m \times n), \ X^H X = I_n\} = \lambda_m + \cdots + \lambda_{m-n+1}.$$

The maximizing matrix is $X = [v_m, \ldots, v_{m-n+1}]$.

9.7.3 Eigenvalue Inequalities

(1) A $(m \times m)$ Hermitian, $x \neq 0$ $(m \times 1)$:

$$\lambda_{min}(A) \leq \frac{x^H Ax}{x^H x} \leq \lambda_{max}(A).$$

(2) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A), \ldots, \lambda_m(A)$:

$$\min\left\{\frac{x^H Ax}{x^H x} \ : \ x \ (m \times 1), \ x \neq 0\right\}$$
$$\leq \lambda_i(A) \leq \max\left\{\frac{x^H Ax}{x^H x} \ : \ x \ (m \times 1), \ x \neq 0\right\},$$

$i = 1, \ldots, m$.

(3) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$:

$$X^H X = I_n \quad \Rightarrow \quad \sum_{i=1}^{n} \lambda_i(A) \leq \operatorname{tr}(X^H AX) \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A).$$

(4) $A = [a_{ij}]$ $(m \times m)$ Hermitian:

$$\lambda_{min}(A) \leq a_{ii} \leq \lambda_{max}(A), \quad i = 1, \ldots, m.$$

(5) $A = [a_{ij}]$ $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\sum_{i=1}^{n} \lambda_i(A) \leq \sum_{i=1}^{n} a_{ii} \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A), \quad n = 1, \ldots, m.$$

(6) $A = [a_{ij}]$ $(m \times m)$ Hermitian with eigenvalues $0 < \lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\prod_{i=1}^{n} \lambda_i(A) \leq \prod_{i=1}^{n} a_{ii}, \quad n = 1, \ldots, m.$$

(7) (Inclusion principle)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, $A_{(n)}$ $(n \times n)$ a principal submatrix of A with eigenvalues $\lambda_1(A_{(n)}) \leq \cdots \leq \lambda_n(A_{(n)})$:

$$\lambda_i(A) \leq \lambda_i(A_{(n)}) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n,$$

$$\lambda_{min}(A) \leq \lambda_{min}(A_{(n)}) \leq \lambda_{max}(A_{(n)}) \leq \lambda_{max}(A).$$

(8) A $(m \times m)$ Hermitian, B $(m \times m)$ positive semidefinite:

$$\lambda_{min}(A + B) \geq \lambda_{min}(A),$$

$$\lambda_{max}(A + B) \geq \lambda_{max}(A).$$

(9) A, B $(m \times m)$ Hermitian:

$$\lambda_{min}(A + B) \geq \lambda_{min}(A) + \lambda_{min}(B),$$

$$\lambda_{max}(A + B) \leq \lambda_{max}(A) + \lambda_{max}(B).$$

(10) A, B $(m \times m)$ Hermitian, $0 < a, b \in \mathbb{R}$:

$$\lambda_{min}(aA + bB) \geq a\lambda_{min}(A) + b\lambda_{min}(B),$$

$$\lambda_{max}(aA + bB) \leq a\lambda_{max}(A) + b\lambda_{max}(B).$$

(11) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, B $(m \times m)$ Hermitian positive semidefinite, $\lambda_1(A + B) \leq \cdots \leq \lambda_m(A + B)$ eigenvalues of $A + B$:

$$\lambda_i(A + B) \geq \lambda_i(A), \quad i = 1, \ldots, m.$$

(12) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, $\boldsymbol{x}$ $(m \times 1)$, $\lambda_1(A \pm \boldsymbol{x}\boldsymbol{x}^H) \leq \cdots \leq \lambda_m(A \pm \boldsymbol{x}\boldsymbol{x}^H)$ eigenvalues of $A \pm \boldsymbol{x}\boldsymbol{x}^H$:

$$\lambda_i(A \pm \boldsymbol{x}\boldsymbol{x}^H) \leq \lambda_{i+1}(A) \leq \lambda_{i+2}(A \pm \boldsymbol{x}\boldsymbol{x}^H), \quad i = 1, 2, \ldots, m-2;$$

$$\lambda_i(A) \leq \lambda_{i+1}(A \pm \boldsymbol{x}\boldsymbol{x}^H) \leq \lambda_{i+2}(A), \quad i = 1, 2, \ldots, m-2.$$

(13) A, B $(m \times m)$ Hermitian, $\text{rk}(B) \leq r$, $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ eigenvalues of A, $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

$$\lambda_i(A + B) \leq \lambda_{i+r}(A) \leq \lambda_{i+2r}(A + B), \quad i = 1, 2, \ldots, m-2r;$$

$$\lambda_i(A) \leq \lambda_{i+r}(A + B) \leq \lambda_{i+2r}(A), \quad i = 1, 2, \ldots, m-2r.$$

(14) (Poincare's separation theorem)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$ such that $X^H X = I_n, n \leq m$, $\lambda_1(X^H AX) \leq \cdots \leq \lambda_n(X^H AX)$ eigenvalues of $X^H AX$:

$$\lambda_i(A) \leq \lambda_i(X^H AX) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n.$$

(15) A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A + B$:

(a) $\displaystyle\sum_{i=1}^{n}[\lambda_i(A) + \lambda_i(B)] \leq \sum_{i=1}^{n} \lambda_i(A+B), \quad n = 1, \ldots, m.$

(b) $\lambda_1(B) \leq \lambda_n(A+B) - \lambda_n(A) \leq \lambda_m(B), \quad n = 1, \ldots, m.$

(c) $|\lambda_n(A+B) - \lambda_n(A)|_{\text{abs}}$
$\leq \max\{|\lambda_j(B)|_{\text{abs}} : j = 1, \ldots, m\}, \quad n = 1, \ldots, m.$

(d) $\lambda_n(A+B) \leq \min\{\lambda_i(A) + \lambda_j(B) : i + j = n + m\}$,
$n = 1, \ldots, m.$

(16) (Weyl's theorem)
A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A + B$:

$$\lambda_i(A+B) \geq \begin{cases} \lambda_i(A) + \lambda_1(B) \\ \lambda_{i-1}(A) + \lambda_2(B) \\ \vdots \\ \lambda_1(A) + \lambda_i(B) \end{cases}$$

and

$$\lambda_i(A+B) \leq \begin{cases} \lambda_i(A) + \lambda_m(B) \\ \lambda_{i+1}(A) + \lambda_{m-1}(B) \\ \vdots \\ \lambda_m(A) + \lambda_i(B) \end{cases}$$

$i = 1, \ldots, m.$

(17) A, B $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A-B) \leq \cdots \leq \lambda_m(A-B)$ eigenvalues of $A - B$:

$$\lambda_1(A-B) > 0 \quad \Rightarrow \quad \lambda_i(A) \geq \lambda_i(B), \ i = 1, \ldots, m.$$

(18) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, x $(m \times 1)$, $c \in \mathbb{R}$,

$$B = \begin{bmatrix} A & x \\ x^H & c \end{bmatrix}$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_{m+1}(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{i+1}(B), \quad i = 1, \ldots, m.$$

(19) A $(m \times m)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, D $(p \times p)$ Hermitian, C $(m \times p)$ and

$$B = \begin{bmatrix} A & C \\ C^H & D \end{bmatrix} (n \times n)$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_n(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{n-m+i}(B), \quad i = 1, \ldots, m.$$

9.7.4 Decompositions of Hermitian Matrices

(1) (Spectral decomposition)
A $(m \times m)$ Hermitian with eigenvalues $\lambda_1, \ldots, \lambda_m$: $\quad A = U\Lambda U^H$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ and U is the unitary $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i^H.$$

(2) (Congruent canonical form)
A $(m \times m)$ Hermitian: There exists a nonsingular $(m \times m)$ matrix T such that $A = T\Lambda T^H$, where $\Lambda = \text{diag}(d_1, \ldots, d_m)$ and the d_i are $+1, -1$ or 0 corresponding to the positive, negative and zero eigenvalues of A, respectively.

(3) (Simultaneous diagonalization of two Hermitian matrices)
A, B $(m \times m)$ Hermitian, $AB = BA$: There exists a unitary $(m \times m)$ matrix U such that $A = UDU^H$ and $B = U\Lambda U^H$, where D and Λ are diagonal matrices.

(4) (Simultaneous diagonalization of a positive definite and a Hermitian matrix)
A $(m \times m)$ positive definite, B $(m \times m)$ Hermitian: There exists a nonsingular $(m \times m)$ matrix T such that $A = TT^H$ and $B = T\Lambda T^H$, where Λ is a diagonal matrix.

Note: Most results of this section are given in Horn & Johnson (1985, Chapter 4) or follow easily from results given there. For the results on eigenvalues see also Chapter 5 and for the decomposition theorems see Chapter 6.

9.8 Idempotent Matrices

Definition: An $(m \times m)$ matrix A is idempotent if $A^2 = A$.

(1) A $(m \times m)$ idempotent, B $(n \times n)$ idempotent:

(a) $A \otimes B$ is idempotent.

(b) $A \oplus B$ is idempotent.

(2) A $(m \times m)$ idempotent:

(a) $A^i = A$, $i = 1, 2, \ldots$

(b) A' is idempotent.

(c) A^H is idempotent.

(d) $I_m - A$ is idempotent.

(e) All eigenvalues of A are 0 or 1.

(f) $\text{rk}(A) = \text{tr}(A)$.

(g) A is simple.

(3) A $(m \times m)$ idempotent:

(a) A is nonsingular $\Rightarrow$ $A = I_m$.

(b) A is Hermitian $\Rightarrow$ $A^+ = A$.

(c) A is Hermitian $\Rightarrow$ A is positive semidefinite.

(d) A is diagonal $\Rightarrow$ the diagonal elements of A are 0 or 1.

(e) $\text{rk}(A) = r$ $\Rightarrow$ there exists a nonsingular matrix Q such that

$$Q^{-1}AQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where the zero submatrices disappear if $r = m$.

(4) A $(m \times n)$:

(a) AA^+ and A^+A are idempotent.

(b) $I_m - AA^+$ and $I_n - A^+A$ are idempotent.

(c) $A(A^HA)^+A^H$ is idempotent.

(d) $A^H(AA^H)^+A$ is idempotent.

(e) $I_m - A(A^HA)^+A^H$ is idempotent.

(f) $I_n - A^H(AA^H)^+A$ is idempotent.

(5) A $(m \times m)$: A idempotent $\Longleftrightarrow$ $\text{rk}(A) + \text{rk}(I_m - A) = m$.

(6) A_i $(m \times m)$ idempotent, $i = 1, \ldots, n$:

(a) $A_iA_j = O_{m\times m}$, $i \neq j \Rightarrow A_1 + \cdots + A_n$ is idempotent.

(b) $A_1 + \cdots + A_n$ is idempotent
$\Rightarrow \mathrm{rk}(A_1 + \cdots + A_n) = \mathrm{rk}(A_1) + \cdots + \mathrm{rk}(A_n)$.

(7) A_i $(m_i \times m_i)$ idempotent for $i = 1, \ldots, r$:

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix} \text{ is idempotent.}$$

Note: Most of these results follow immediately from the defining property of idempotency. A number of results is given in Lancaster & Tismenetsky (1985) and other matrix books.

9.9 Nonnegative, Positive and Stochastic Matrices

In this section all matrices are real unless otherwise noted.

9.9.1 Definitions

$A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times n)$ real:

A is **nonnegative** ($A \geq 0$ or $0 \leq A$) if $a_{ij} \geq 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

A is **positive** ($A > 0$ or $0 < A$) if $a_{ij} > 0$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

$A \leq B$ (or $B \geq A$) if $a_{ij} \leq b_{ij}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

$A < B$ (or $B > A$) if $a_{ij} < b_{ij}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

$|A|_{\text{abs}} \equiv [\, |a_{ij}|_{\text{abs}} \,]$.

$A = [a_{ij}]$ $(m \times m)$ real:

A is **stochastic** if $0 \leq a_{ij} \leq 1$, $i,j = 1, \ldots, m$, and $\sum_{j=1}^{m} a_{ij} = 1$, $i = 1, \ldots, m$.

A is **doubly stochastic** if $0 \leq a_{ij} \leq 1$, $i,j = 1, \ldots, m$, $\sum_{j=1}^{m} a_{ij} = 1$, $i = 1, \ldots, m$, and $\sum_{i=1}^{m} a_{ij} = 1$, $j = 1, \ldots, m$.

$\rho(A) \equiv \max\{|\lambda|_{\text{abs}} : \lambda \text{ is eigenvalue of } A\}$ is the **spectral radius** of A.

9.9.2 General Results

All results for nonnegative matrices are also valid for stochastic and doubly stochastic matrices because they are special nonnegative matrices.

(1) $A, B\ (m \times n), a, b \in \mathbb{R}$:

(a) $A \geq 0, B \geq 0, a, b \geq 0 \ \Rightarrow\ aA + bB \geq 0.$

(b) $A \geq 0 \ \Rightarrow\ A' \geq 0.$

(c) $A > 0 \ \Rightarrow\ A' > 0.$

(2) $A, B\ (m \times m)$:

(a) $A \geq 0 \ \Rightarrow\ A^i \geq 0,\ i = 0, 1, 2 \ldots$

(b) $A > 0 \ \Rightarrow\ A^i > 0,\ i = 1, 2 \ldots$

(c) $0 \leq A \leq B \ \Rightarrow\ 0 \leq A^i \leq B^i,\ i = 0, 1, 2 \ldots$

(3) $A, B\ (m \times m), C, D\ (m \times n)$:

$$0 \leq A \leq B,\ 0 \leq C \leq D \quad \Rightarrow \quad 0 \leq AC \leq BD.$$

(4) $A\ (m \times n), \boldsymbol{x}\ (n \times 1)$:

(a) $A > 0,\ \boldsymbol{x} \geq 0,\ \boldsymbol{x} \neq 0 \ \Rightarrow\ A\boldsymbol{x} > 0.$

(b) $A \geq 0,\ \boldsymbol{x} > 0,\ A\boldsymbol{x} = 0 \ \Rightarrow\ A = 0.$

(5) $A\ (m \times m), D = \mathrm{diag}(d_1, \ldots, d_m)$:
$A \geq 0,\ d_i > 0,\ i = 1, \ldots, m \ \Rightarrow\ D^{-1}AD \geq 0.$

(6) $A\ (m \times m)$ nonsingular, $A \geq 0$: $\quad A^{-1} \geq 0 \iff A$ is a generalized permutation matrix.

(7) $A\ (m \times m), A \geq 0$:

(a) A is tridiagonal $\Rightarrow$ all eigenvalues of A are real.

(b) A is stochastic $\iff A \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$

(8) $A, B\ (m \times m)$:

(a) A, B are stochastic $\Rightarrow AB$ is stochastic.

(b) A, B are doubly stochastic $\Rightarrow AB$ is doubly stochastic.

(9) $A\ (m \times m)$: $\quad A$ is doubly stochastic $\iff$ for some $n \in \mathbb{N}$, there exist $(m \times m)$ permutation matrices $P_1, \ldots, P_n$ and $c_1, \ldots, c_n \in \mathbb{R}$ with $c_1 + \cdots + c_n = 1$ such that $A = c_1 P_1 + \cdots + c_n P_n$.

Note: More on nonnegative matrices can be found in Minc (1988), Seneta (1973), Berman, Neumann & Stern (1989), Berman & Plemmons (1979).

Many of the results of this section can also be found in Horn & Johnson (1985, Chapter 8).

9.9.3 Results Related to the Spectral Radius

(1) $A = [a_{ij}]$ $(m \times m)$:

(a) $A > 0 \;\Rightarrow\; \rho(A) > 0$.

(b) $A \geq 0, \sum_{j=1}^{m} a_{ij} > 0,\ i = 1, \ldots, m \;\Rightarrow\; \rho(A) > 0$.

(c) A is stochastic $\;\Rightarrow\; \rho(A) > 0$.

(d) A is doubly stochastic $\;\Rightarrow\; \rho(A) > 0$.

(e) $A \geq 0, A^i > 0$ for some $i \geq 1 \;\Rightarrow\; \rho(A) > 0$.

(2) A, B $(m \times m)$:

(a) $0 \leq A \leq B \;\Rightarrow\; \rho(A) \leq \rho(B)$.

(b) $0 < A < B \;\Rightarrow\; \rho(A) < \rho(B)$.

(3) $A = [a_{ij}]$ $(m \times m)$, $A \geq 0$: $\quad \rho(A_{(n)}) \leq \rho(A),\ n = 1, \ldots, m$, where

$$A_{(n)} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}$$

are principal submatrices of A.

(4) $A = [a_{ij}]$ $(m \times m)$, $A \geq 0$:

(a) $\max_{i=1,\ldots,m} a_{ii} \leq \rho(A)$.

(b) $\min_{i=1,\ldots,m} \sum_{j=1}^{m} a_{ij} \leq \rho(A) \leq \max_{i=1,\ldots,m} \sum_{j=1}^{m} a_{ij}$.

(c) $\min_{j=1,\ldots,m} \sum_{i=1}^{m} a_{ij} \leq \rho(A) \leq \max_{j=1,\ldots,m} \sum_{i=1}^{m} a_{ij}$.

(d) $\rho(A) = \max_{\substack{x=(x_1,\ldots,x_m)' \geq 0 \\ x \neq 0}} \ \min_{\substack{1 \leq i \leq m \\ x_i \neq 0}} \ \frac{1}{x_i} \sum_{j=1}^{m} a_{ij} x_j$.

(5) $A = [a_{ij}] \geq 0$ $(m \times m)$, $x = (x_1, \ldots, x_m)' > 0$ $(m \times 1)$:

$$\min_{i=1,\ldots,m} \frac{1}{x_i} \sum_{j=1}^{m} a_{ij} x_j \leq \rho(A) \leq \max_{i=1,\ldots,m} \frac{1}{x_i} \sum_{j=1}^{m} a_{ij} x_j,$$

$$\min_{j=1,\ldots,m} x_j \sum_{i=1}^{m} \frac{a_{ij}}{x_i} \leq \rho(A) \leq \max_{j=1,\ldots,m} x_j \sum_{i=1}^{m} \frac{a_{ij}}{x_i}.$$

(6) $A\ (m \times m), A \geq 0,\ x\ (m \times 1), x > 0,\ a, b \in \mathbb{R}, a, b \geq 0:$

(a) $ax < Ax < bx \ \Rightarrow\ a < \rho(A) < b.$

(b) $ax \leq Ax \leq bx \ \Rightarrow\ a \leq \rho(A) \leq b.$

(c) $Ax = ax \ \Rightarrow\ a = \rho(A).$

(7) $A\ (m \times m)$, λ eigenvalue of A with eigenvector $x \neq 0$: $A > 0$, $|\lambda|_{\mathsf{abs}} = \rho(A) \ \Rightarrow\ |\lambda|_{\mathsf{abs}}$ is eigenvalue of A with eigenvector $|x|_{\mathsf{abs}}$.

(8) $A\ (m \times m), A \geq 0:$
$\rho(A)$ is eigenvalue of A with eigenvector $x \geq 0, x \neq 0$.

(9) $A\ (m \times m), A > 0:$

(a) $\rho(A)$ is a simple eigenvalue of A.

(b) $\rho(A)$ is eigenvalue of A with eigenvector $x > 0$.

(c) λ is eigenvalue of A, $\lambda \neq \rho(A) \ \Rightarrow\ |\lambda|_{\mathsf{abs}} < \rho(A)$.

(10) (Hopf's theorem)
$A = [a_{ij}]\ (m \times m), A > 0, \lambda_{m-1}$ eigenvalue of A with second largest modulus, $M = \max\{a_{ij} : i, j = 1, \ldots, m\}$, $\mu = \min\{a_{ij} : i, j = 1, \ldots, m\}$:

$$\frac{|\lambda_{m-1}|_{\mathsf{abs}}}{\rho(A)} \leq \frac{M - \mu}{M + \mu} < 1.$$

(11) $A\ (m \times m), A > 0, x > 0\ (m \times 1)$ eigenvector of A corresponding to eigenvalue $\rho(A), y > 0\ (m \times 1)$ eigenvector of A' corresponding to eigenvalue $\rho(A), x'y = 1$: $\lim_{i \to \infty} [\rho(A)^{-1} A]^i = xy'$.

(12) $A\ (m \times m), A > 0:$

(a) $\lim_{i \to \infty} [\rho(A)^{-1} A]^i > 0$.

(b) $\mathbf{rk} \left\{ \lim_{i \to \infty} [\rho(A)^{-1} A]^i \right\} = 1$.

Note: The results of this subsection are given in Horn & Johnson (1985, Chapter 8).

9.10 Orthogonal Matrices

Definition: An $(m \times m)$ matrix A is orthogonal if it is nonsingular and $A^{-1} = A'$.

9.10.1 General Results

(1) $A\ (m \times m)$:

(a) A is orthogonal $\iff AA' = I_m \iff A'A = I_m$.

(b) A is orthogonal $\iff A'$ is orthogonal.

(c) A is orthogonal $\iff A^H$ is orthogonal.

(d) A is orthogonal $\iff \bar{A}$ is orthogonal.

(e) A is orthogonal $\iff A^{-1}$ is orthogonal.

(2) $A\ (m \times m)$ orthogonal:

(a) A^i is orthogonal for $i = 1, 2, \ldots$

(b) $|\det(A)|_{\text{abs}} = 1$.

(c) $\text{rk}(A) = m$.

(3) $A\ (m \times m)$ real:

(a) A is orthogonal $\Rightarrow A$ is normal.

(b) A is orthogonal $\iff A$ is unitary.

(c) A is orthogonal
$\iff x'A'Ay = x'y$ for all real $(m \times 1)$ vectors x, y.

(d) A is skew-symmetric $\Rightarrow (I_m - A)(I_m + A)^{-1}$ is orthogonal.

(4) (a) The $(m \times m)$ unit matrix I_m is orthogonal.

(b) The $(mn \times mn)$ commutation matrix K_{mn} is orthogonal.

(c) A permutation matrix is orthogonal.

(d) For $\theta \in \mathbb{R}$, $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

(5) $A_i\ (m \times m)$ real $i = 1, 2, \ldots$: A_i is orthogonal and $\lim\limits_{i\to\infty} A_i$ exists
$\Rightarrow \lim\limits_{i\to\infty} A_i$ is real orthogonal.

(6) $A, B\ (m \times m)$: A, B are orthogonal $\Rightarrow AB$ is orthogonal.

(7) $A\ (m \times m)$ orthogonal, $B\ (n \times n)$ orthogonal:

(a) $A \otimes B$ is orthogonal.

(b) $A \oplus B$ is orthogonal.

(8) $x\ (m \times 1), A\ (m \times m)$: A is real orthogonal $\Rightarrow \|Ax\|_2 = \|x\|_2$.

(9) $A\ (m \times n), B\ (n \times n)$: B is real orthogonal $\Rightarrow \|AB\|_2 = \|A\|_2$.

(10) $A\ (m \times m), B\ (m \times n)$: A is real orthogonal $\Rightarrow \|AB\|_2 = \|B\|_2$.

(11) $A\ (m \times m)$ real orthogonal:

(a) λ is eigenvalue of $A \Rightarrow |\lambda|_{\text{abs}} = 1$.

(b) λ is eigenvalue of $A \Rightarrow 1/\lambda$ is eigenvalue of A.

(12) A $(m \times m)$ real with eigenvalues $\lambda_1, \ldots, \lambda_m$:
A is normal. $|\lambda_i|_{\text{abs}} = 1, i = 1, \ldots, m \iff A$ is orthogonal.

(13) A $(m \times m)$ real: A is similar to a real orthogonal matrix $\Rightarrow A^{-1}$ is similar to A'.

Note: The results of this subsection may be found in Horn & Johnson (1985) or follow easily from (3).

9.10.2 Decompositions of Orthogonal Matrices

(1) A $(m \times m)$ real orthogonal: There exists a real orthogonal $(m \times m)$ matrix Q such that

$$A = Q \begin{bmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_p & & & \\ & & & \Lambda_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Lambda_k \end{bmatrix} Q',$$

where the $\lambda_i = \pm 1$ and the Λ_i are real (2×2) matrices of the form

$$\Lambda_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}.$$

(2) A $(m \times m)$ orthogonal : There exists a real orthogonal $(m \times m)$ matrix Q and a real skew-symmetric $(m \times m)$ matrix S such that $A = Q\exp(\mathrm{i}S)$.

Note: For the results of this subsection see Chapter 6.

9.11 Partitioned Matrices

Definition: An $(m \times n)$ matrix

$$A = [A_{ij}] = \begin{bmatrix} A_{11} & \ldots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \ldots & A_{pq} \end{bmatrix}$$

consisting of $(m_i \times n_j)$ submatrices A_{ij}, $i = 1, \ldots, p$, $j = 1, \ldots, q$, is said to be a partitioned matrix or a block matrix. Special cases are

$$A = [A_1, \ldots, A_n] = [A_1 : \ldots : A_n],$$

where the A_i are blocks of columns, and

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix},$$

where the A_i are blocks of rows of A.

9.11.1 General Results

(1) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p)$:

(a) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}' = \begin{bmatrix} A' & C' \\ B' & D' \end{bmatrix}$.

(b) $\overline{\begin{bmatrix} A & B \\ C & D \end{bmatrix}} = \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{bmatrix}$.

(c) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^H = \begin{bmatrix} A^H & C^H \\ B^H & D^H \end{bmatrix}$.

(d) $\left|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right|_{\text{abs}} = \begin{bmatrix} |A|_{\text{abs}} & |B|_{\text{abs}} \\ |C|_{\text{abs}} & |D|_{\text{abs}} \end{bmatrix}$.

(2) $A\ (m \times n), B\ (m \times p), C\ (q \times n), D\ (q \times p), E\ (n \times r), F\ (n \times s), G\ (p \times r), H\ (p \times s)$:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}.$$

(3) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

$$\operatorname{tr} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \operatorname{tr}(A) + \operatorname{tr}(D).$$

(4) $A\ (m \times n), B\ (p \times q)$:

$$\operatorname{rk} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \operatorname{rk}(A) + \operatorname{rk}(B).$$

(5) $A\ (m \times m), B\ (n \times n)$: $\lambda(A)$ and $\lambda(B)$ are eigenvalues of A and B, respectively, with associated eigenvectors $v(A)$ and $v(B)$ $\Rightarrow$ $\lambda(A)$

and $\lambda(B)$ are eigenvalues of

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

with eigenvectors

$$\begin{bmatrix} v(A) \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ v(B) \end{bmatrix},$$

respectively.

(6) A $(n \times n)$ Hermitian with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_n(A)$, D $(p \times p)$ Hermitian, C $(n \times p)$ and

$$B = \begin{bmatrix} A & C \\ C^H & D \end{bmatrix} (m \times m)$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{m-n+i}(B) \quad \text{for } i = 1, \ldots, n.$$

(7) A_i $(m_i \times m_i)$, $i = 1, \ldots, r$:

(a) $$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}^i = \begin{bmatrix} A_1^i & & 0 \\ & \ddots & \\ 0 & & A_r^i \end{bmatrix}, \quad i = 0, 1, 2, \ldots$$

(b) A_i is idempotent, $i = 1, \ldots, r$

$$\Rightarrow \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix} \text{ is idempotent.}$$

9.11.2 Determinants of Partitioned Matrices

(1) $\det\begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix} = (-1)^m.$

(2) A $(m \times m)$, B $(m \times n)$, C $(n \times n)$:

$$\det\begin{bmatrix} A & B \\ O_{n\times m} & C \end{bmatrix} = \det(A)\det(C),$$

$$\det\begin{bmatrix} A & O_{m\times n} \\ B & C \end{bmatrix} = \det(A)\det(C).$$

(3) $A, B, C\ (m \times m)$:

$$\det \begin{bmatrix} A & B \\ C & O_{m \times m} \end{bmatrix} = (-1)^m \det(B)\det(C),$$

$$\det \begin{bmatrix} O_{m \times m} & B \\ C & A \end{bmatrix} = (-1)^m \det(B)\det(C).$$

(4) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

(a) A nonsingular $\Rightarrow \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B)$.

(b) D nonsingular $\Rightarrow \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(D)\det(A - BD^{-1}C)$.

(5) $A\ (m \times m)$ nonsingular, $a, b\ (m \times 1)$, $c \in \mathbb{C}$:

$$\det \begin{bmatrix} A & a \\ b' & c \end{bmatrix} = c \det(A) - b' A^{adj} a = \det(A)(c - b' A^{-1} a).$$

(6) (Fischer's inequality)
$A\ (m \times m), B\ (m \times n), C\ (n \times n)$:

$$D = \begin{bmatrix} A & B \\ B^H & C \end{bmatrix} \text{ positive definite } \Rightarrow \det(D) \leq \det(A)\det(C).$$

9.11.3 Partitioned Inverses

(1) $A_i\ (m_i \times m_i)$ nonsingular, $i = 1, \ldots, r$:

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} & & 0 \\ & \ddots & \\ 0 & & A_r^{-1} \end{bmatrix}.$$

(2) $A\ (m \times m), B\ (m \times n), C\ (n \times m), D\ (n \times n)$:

(a) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, A and $(D - CA^{-1}B)$ nonsingular

$$\Rightarrow \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}.$$

(b) $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, D and $(A - BD^{-1}C)$ nonsingular

$$\Rightarrow \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}.$$

(3) A $(m \times m)$ symmetric, B $(m \times n)$, C $(m \times p)$, D $(n \times n)$ symmetric, E $(p \times p)$ symmetric :

$$\begin{bmatrix} A & B & C \\ B' & D & 0 \\ C' & 0 & E \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} F & -FBD^{-1} & -FCE^{-1} \\ -D^{-1}B'F & D^{-1} + D^{-1}B'FBD^{-1} & D^{-1}B'FCE^{-1} \\ -E^{-1}C'F & E^{-1}C'FBD^{-1} & E^{-1} + E^{-1}C'FCE^{-1} \end{bmatrix},$$

if all inverses exist and $F = (A - BD^{-1}B' - CE^{-1}C')^{-1}$.

(4) $m, n \in \mathbb{N}, m > n, A$ $(m \times n), B$ $(m \times (m - n))$:

$$\mathrm{rk}(A) = n,\ \mathrm{rk}(B) = m - n,\ A^H B = 0$$
$$\Rightarrow\ [A : B]^{-1} = \begin{bmatrix} (A^H A)^{-1} A^H \\ (B^H B)^{-1} B^H \end{bmatrix}.$$

(5) $m, n \in \mathbb{N}, m < n, A$ $(m \times n), B$ $((n - m) \times n)$:

$$\mathrm{rk}(A) = m,\ \mathrm{rk}(B) = n - m,\ AB^H = 0$$
$$\Rightarrow \begin{bmatrix} A \\ B \end{bmatrix}^{-1} = [A^H(AA^H)^{-1} : B^H(BB^H)^{-1}].$$

9.11.4 Partitioned Generalized Inverses

(1) A $(m \times m)$ nonsingular, B $(m \times n)$, C $(r \times m)$, D $(r \times n)$:

$$\mathrm{rk}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = m,\ D = CA^{-1}B$$

$\Rightarrow \begin{bmatrix} A^{-1} & O_{m\times r} \\ O_{n\times m} & O_{n\times r} \end{bmatrix}$ is a generalized inverse of $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

(2) A $(m_i \times n_i)$, $i = 1, \ldots, r$:

(a) $\begin{bmatrix} A_1^- & & 0 \\ & \ddots & \\ 0 & & A_r^- \end{bmatrix}$ is a generalized inverse of $\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}$.

(b) $$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}^+ = \begin{bmatrix} A_1^+ & & 0 \\ & \ddots & \\ 0 & & A_r^+ \end{bmatrix}.$$

(3) $A\ (m \times n)$:

$$\begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix}^+ = \begin{bmatrix} 0 & 0 \\ 0 & A^+ \end{bmatrix}, \quad \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} A^+ & 0 \\ 0 & 0 \end{bmatrix}.$$

(4) $A\ (m \times n)$, $B\ (m \times p)$:

$$[A : B]^+ = \begin{bmatrix} A^+ - A^+B(C^+ + D) \\ C^+ + D \end{bmatrix},$$

where $C = (I_m - AA^+)B$ and

$$\begin{aligned} D &= (I_p - C^+C)[I_p + (I_p - C^+C)B^H(A^+)^H A^+B(I_p - C^+C)]^{-1} \\ &\quad \times B^H(A^+)^H A^+(I_m - BC^+). \end{aligned}$$

(5) $A\ (m \times n)$, $B\ (p \times n)$:

$$\begin{bmatrix} A \\ B \end{bmatrix}^+ = [A^+ - TBA^+ : T],$$

where $T = E^+ + (I_n - E^+B)A^+(A^+)^H B^H K(I_p - EE^+)$ with $E = B(I_n - A^+A)$ and $K = [I_p + (I_p - EE^+)BA^+(A^+)^H B^H(I_p - EE^+)]^{-1}$.

9.11.5 Partitioned Matrices Related to Duplication Matrices

(1) D_n $(n^2 \times \frac{1}{2}n(n+1))$ duplication matrix:

$$D'_{m+1}D_{m+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2I_m & 0 \\ 0 & 0 & D'_m D_m \end{bmatrix},$$

$$D^+_{m+1}(D^+_{m+1})' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}I_m & 0 \\ 0 & 0 & (D'_m D_m)^{-1} \end{bmatrix}.$$

(2) $A\ (m \times m),\ a, b\ (m \times 1),\ c \in \mathbb{C}$:

$$D'_{m+1}\operatorname{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ a+b \\ D'_m \operatorname{vec}(A) \end{bmatrix},$$

$$D^+_{m+1}\operatorname{vec}\begin{bmatrix} c & b' \\ a & A \end{bmatrix} = \begin{bmatrix} c \\ \frac{1}{2}(a+b) \\ D^+_m \operatorname{vec}(A) \end{bmatrix}.$$

(3) $A, B\ (m \times m)$ symmetric, $a, b\ (m \times 1), \alpha, \beta \in \mathbb{C}$:

(a) $D'_{m+1}\left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix}\right) D_{m+1} =$

$$\begin{bmatrix} \alpha\beta & \alpha b' + \beta a' & (a' \otimes b')D_m \\ \alpha b + \beta a & \alpha B + \beta A + ab' + ba' & (a' \otimes B + b' \otimes A)D_m \\ D'_m(a \otimes b) & D'_m(a \otimes B + b \otimes A) & D'_m(A \otimes B)D_m \end{bmatrix}.$$

(b) $D^+_{m+1}\left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix}\right) D^{+'}_{m+1} =$

$$\begin{bmatrix} \alpha\beta & \frac{1}{2}(\alpha b' + \beta a') & (a' \otimes b')D^{+'}_m \\ \frac{1}{2}(\alpha b + \beta a) & \frac{1}{4}(\alpha B + \beta A + ab' + ba') & \frac{1}{2}(a' \otimes B + b' \otimes A)D^{+'}_m \\ D^+_m(a \otimes b) & \frac{1}{2}D^+_m(a \otimes B + b \otimes A) & D^+_m(A \otimes B)D^{+'}_m \end{bmatrix}.$$

Note: Results on partitioned matrices including a number of the foregoing ones are given, for instance, in Magnus (1988) and Magnus & Neudecker (1988). Many of the foregoing results are immediate consequences of definitions.

9.12 Positive Definite, Negative Definite and Semidefinite Matrices

Definitions: A Hermitian $(m \times m)$ matrix A is

- positive definite if $x^H Ax > 0$ for all $(m \times 1)$ vectors $x \neq 0$;
- positive semidefinite if $x^H Ax \geq 0$ for all $(m \times 1)$ vectors x;
- negative definite if $x^H Ax < 0$ for all $(m \times 1)$ vectors $x \neq 0$;
- negative semidefinite if $x^H Ax \leq 0$ for all $(m \times 1)$ vectors x.

Note that a definite matrix is always Hermitian according to this definition. A real symmetric $(m \times m)$ matrix A is

- positive definite if $x'Ax > 0$ for all real $(m \times 1)$ vectors $x \neq 0$;

- positive semidefinite if $x'Ax \geq 0$ for all real $(m \times 1)$ vectors x;
- negative definite if $x'Ax < 0$ for all real $(m \times 1)$ vectors $x \neq 0$;
- negative semidefinite if $x'Ax \leq 0$ for all real $(m \times 1)$ vectors x.

Without special notice, a real definite matrix is always understood to be symmetric.

9.12.1 General Properties

(1) $A, B\ (m \times m)$:

(a) A positive definite, B positive semidefinite $\Rightarrow$ $A + B$ positive definite.

(b) A negative definite, B negative semidefinite $\Rightarrow$ $A + B$ negative definite.

(c) A, B positive semidefinite $\Rightarrow$ $A + B$ positive semidefinite.

(d) A, B negative semidefinite $\Rightarrow$ $A + B$ negative semidefinite.

(2) $A\ (m \times m), c \in \mathbb{R}$:

(a) A positive (semi) definite, $c > 0$ $\Rightarrow$ cA is positive (semi) definite.

(b) A negative (semi) definite, $c > 0$ $\Rightarrow$ cA negative (semi) definite.

(c) A positive (semi) definite, $c < 0$ $\Rightarrow$ cA negative (semi) definite.

(d) A negative (semi) definite, $c < 0$ $\Rightarrow$ cA positive (semi) definite.

(e) A positive (semi) definite $\iff$ $-A$ negative (semi) definite.

(f) $A = [a_{ij}]$ positive (semi) definite

$$\iff \det \begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix} > (\geq)\ 0 \text{ for } i = 1, \dots, m.$$

(g) $A = [a_{ij}]$ negative (semi) definite

$$\iff \det \begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix} \begin{cases} > (\geq)\ 0, & i \text{ even} \\ < (\leq)\ 0, & i \text{ odd} \end{cases}$$

for $i = 1, \dots, m$.

(h) A positive (semi) definite $\Rightarrow$ A^i positive (semi) definite for $i = 1, 2, \dots$

(3) $A, B\ (m \times m)$:

(a) (Schur product theorem)
A, B positive (semi) definite $\Rightarrow$ $A \odot B$ positive (semi) definite.

(b) A, B positive semidefinite $\Rightarrow$ $\text{rk}(A \odot B) \leq \text{rk}(A)\text{rk}(B)$.

(4) $A\ (m \times m), B\ (n \times n)$: A, B positive (semi) definite $\Rightarrow$ $A \oplus B$ positive (semi) definite.

(5) $A\ (m \times m)$ positive (semi) definite:

(a) $\det(A) > (\geq)\ 0$.

(b) $\text{tr}(A) > (\geq)\ 0$.

(6) $A\ (m \times m)$ negative (semi) definite : $\text{tr}(A) < (\leq)\ 0$.

(7) $A\ (m \times m)$ positive (negative) definite:

(a) $\text{rk}(A) = m$.

(b) A is nonsingular.

(c) A^{-1} is positive (negative) definite.

(8) $A\ (m \times m)$ Hermitian:
A is idempotent $\Rightarrow$ A is positive semidefinite.

(9) $A = \text{diag}(a_{11}, \ldots, a_{mm})$:

(a) A positive (semi) definite $\iff$ $a_{ii} > (\geq)\ 0,\ i = 1, \ldots, m$.

(b) A negative (semi) definite $\iff$ $a_{ii} < (\leq)\ 0,\ i = 1, \ldots, m$.

(10) $A\ (m \times m), B\ (m \times n)$:

(a) A positive (negative) semidefinite $\Rightarrow$ $B^H AB$ positive (negative) semidefinite.

(b) A positive (negative) definite, $\text{rk}(B) = n$ $\Rightarrow$ $B^H AB$ positive (negative) definite.

(11) (Fischer's inequality)
$A\ (m \times m), B\ (m \times n), C\ (n \times n)$:

$$D = \begin{bmatrix} A & B \\ B^H & C \end{bmatrix} \text{ positive definite } \Rightarrow \det(D) \leq \det(A)\det(C).$$

(12) (Aitken's integral)
$A\ (m \times m)$ real positive definite, $x = (x_1, \ldots, x_m)'\ (m \times 1)$ real:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}x'Ax} \mathrm{d}x_1 \cdots \mathrm{d}x_m = (\sqrt{2\pi})^m (\det A)^{-1/2}.$$

9.12.2 Eigenvalue Results

(1) A $(m \times m)$ Hermitian (or real symmetric):

(a) A is positive (semi) definite $\iff$ all eigenvalues of A are positive (nonnegative).

(b) A is negative (semi) definite $\iff$ all eigenvalues of A are negative (nonpositive).

(2) $A = [a_{ij}]$ $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$:

$$\prod_{i=1}^{n} \lambda_i \leq \prod_{i=1}^{n} a_{ii}, \quad n = 1, \ldots, m.$$

(3) A $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, X $(m \times n)$:

$$X^H X = I_n \quad \Rightarrow \quad \prod_{i=1}^{n} \lambda_i \leq \det(X^H AX) \leq \prod_{i=1}^{n} \lambda_{m-n+i}.$$

(4) A $(m \times m)$ real positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, X $(m \times n)$ real:

$$X'X = I_n \quad \Rightarrow \quad \prod_{i=1}^{n} \lambda_i \leq \det(X'AX) \leq \prod_{i=1}^{n} \lambda_{m-n+i}.$$

(5) $A = [a_{ij}]$ $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$:

$$A_{(n)} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad \Rightarrow \quad \prod_{i=1}^{n} \lambda_i \leq \det(A_{(n)}) \leq \prod_{i=1}^{n} \lambda_{m-n+i}.$$

(6) A $(m \times m)$ Hermitian (or real symmetric), B $(m \times m)$ positive semidefinite:

$$\lambda_{min}(A + B) \geq \lambda_{min}(A),$$

$$\lambda_{max}(A + B) \geq \lambda_{max}(A).$$

(7) A $(m \times m)$ Hermitian (or real symmetric) with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, B $(m \times m)$ positive semidefinite, $\lambda_1(A + B) \leq \cdots \leq \lambda_m(A + B)$ eigenvalues of $A + B$:

$$\lambda_i(A + B) \geq \lambda_i(A), \quad i = 1, \ldots, m.$$

(8) A $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, $n \in \{1, \ldots, m\}$:

$$\lambda_1 \cdot \lambda_2 \cdots \lambda_n = \min\{(x_1^H A x_1) \cdots (x_n^H A x_n) : x_i \ (m \times 1),\ X = [x_1, \ldots, x_n] \ (m \times n),\ X^H X = I_n\}.$$

(9) A $(m \times m)$ positive definite with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B^H AB) \ : \ B \ (m \times n), B^H B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$\max\{\det(B^H AB) \ : \ B \ (m \times n), B^H B = I_n\} = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

(10) A $(m \times m)$ real positive definite with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and associated orthonormal $(m \times 1)$ eigenvectors $v_1, \ldots, v_m$:

$$\min\{\det(B'AB) \ : \ B \ (m \times n) \text{ real}, \ B'B = I_n\} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

The minimizing matrix is $B = [v_1, \ldots, v_n]$.

$\max\{\det(B'AB) : B \ (m \times n) \text{ real}, \ B'B = I_n\} = \lambda_m \cdot \lambda_{m-1} \cdots \lambda_{m-n+1}.$

The maximizing matrix is $B = [v_m, \ldots, v_{m-n+1}]$.

9.12.3 Decomposition Theorems for Definite Matrices

(1) A $(m \times m)$:

(a) A is positive definite $\Rightarrow$ A is similar to I_m.

(b) A is positive semidefinite, $\text{rk}(A) = n$ $\Rightarrow$ A is similar to $\begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}$.

(c) A is negative definite $\Rightarrow$ A is similar to $-I_m$.

(d) A is negative semidefinite, $\text{rk}(A) = n$ $\Rightarrow$ A is similar to $\begin{bmatrix} -I_n & 0 \\ 0 & 0 \end{bmatrix}$.

(2) A $(m \times m)$:

(a) (Choleski decomposition)
A is positive definite $\Longleftrightarrow$ there exists a lower triangular $(m \times m)$ matrix B with real positive elements on the principal diagonal such that $A = BB^H$.

(b) (Choleski decomposition of a real matrix)
A is real positive definite $\iff$ there exists a real lower triangular $(m \times m)$ matrix B with positive elements on the principal diagonal such that $A = BB'$.

(c) (Square root decomposition)
A is positive (semi) definite $\iff$ there exists a positive (semi) definite $(m \times m)$ matrix B such that $A = BB$, that is, B is a square root of A.

(d) (Square root decomposition of a real matrix)
A is real positive (semi) definite $\iff$ there exists a real positive (semi) definite $(m \times m)$ matrix B such that $A = BB$, that is, B is a real square root of A.

(e) A is positive definite $\iff$ there exists a nonsingular $(m \times m)$ matrix B such that $A = B^H B$.

(f) A is real positive definite $\iff$ there exists a real nonsingular $(m \times m)$ matrix B such that $A = B'B$.

(3) $A\ (m \times m)$: A is positive semidefinite, $\mathrm{rk}(A) = r \Rightarrow$ there exists an $(m \times r)$ matrix B with $\mathrm{rk}(B) = r$ such that $A = BB^H$ and $A^+ = B(B^H B)^{-2} B^H$.

(4) (Simultaneous diagonalization of a positive definite and a Hermitian matrix)
$A\ (m \times m)$ positive definite, $B\ (m \times m)$ Hermitian: There exists a nonsingular $(m \times m)$ matrix T such that $A = TT^H$ and $B = T\Lambda T^H$, where Λ is a diagonal matrix.

(5) (Simultaneous diagonalization of a positive definite and a symmetric matrix)
$A\ (m \times m)$ positive definite, $B\ (m \times m)$ symmetric: There exists a nonsingular $(m \times m)$ matrix T such that $A = TT^H$ and $B = T\Lambda T'$, where Λ is a real diagonal matrix with nonnegative diagonal elements.

Note: Many results on positive and negative (semi) definite matrices including proofs can be found in Horn & Johnson (1985, Chapter 7) and other books on matrices. See also Chapter 4 for results related to traces and determinants of definite matrices. For Aitken's integral see Searle (1982, p. 340) and for the decomposition results see Chapter 6. Since positive (negative) definite and semidefinite matrices are Hermitian, all results for the latter also hold for the former matrices (see Section 9.7).

9.13 Symmetric Matrices

Definitions: An $(m \times m)$ matrix A is symmetric if $A' = A$, that is, the ijth element a_{ij} is equal to the jith element, $a_{ji} = a_{ij}$, so that

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{12} & a_{22} & \dots & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{mm} \end{bmatrix}.$$

An $(m \times m)$ matrix A is skew-symmetric if $A' = -A$, that is, the ijth element a_{ij} is -1 times the jith element, $a_{ij} = -a_{ij}$, so that

$$A = \begin{bmatrix} 0 & a_{12} & \dots & a_{1m} \\ -a_{12} & 0 & \dots & a_{2m} \\ \vdots & & \ddots & \vdots \\ -a_{1m} & -a_{2m} & \dots & 0 \end{bmatrix}.$$

9.13.1 General Properties

(1) A $(m \times m)$ real: A is symmetric $\iff$ A is Hermitian.

(2) A, B $(m \times m)$ symmetric, $c \in \mathbb{C}$:

(a) cA is symmetric.

(b) $A \pm B$ is symmetric.

(c) $AB = BA \Rightarrow AB$ is symmetric.

(d) $A \odot B$ is symmetric.

(3) A $(m \times m)$ symmetric, B $(n \times n)$ symmetric:

(a) $A \otimes B$ is symmetric.

(b) $A \oplus B$ is symmetric.

(4) $A = [a_{ij}]$ $(m \times m)$ symmetric:

(a) A' is symmetric.

(b) A^H is symmetric.

(c) A^{-1} is symmetric, if A is nonsingular.

(d) A^{adj} is symmetric.

(e) A^i is symmetric for $i = 1, 2, \dots$

(f) $\bar{A}$ is symmetric.

(g) $|A|_{\text{abs}}$ is symmetric.

(h) $\begin{bmatrix} a_{11} & \dots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix}$ is symmetric for $i = 1, 2, \dots, m$.

(5) $A\ (m \times m)$:

(a) $A + A'$ is symmetric.

(b) $A - A'$ is skew-symmetric.

(c) AA' is symmetric.

(d) $A'A$ is symmetric.

(6) $A\ (m \times m)$:

(a) A is symmetric $\iff$ $B'AB$ is symmetric for all $(m \times m)$ matrices B.

(b) A is symmetric $\iff$ there exists a real orthogonal $(m \times m)$ matrix Q and a real diagonal $(m \times m)$ matrix Λ such that $A = Q\Lambda Q'$.

(7) $A\ (m \times m), B\ (m \times n)$: A is symmetric $\Rightarrow$ $B'AB$ is symmetric.

(8) $A, B\ (m \times m)$ real symmetric:

(a) $[\text{tr}(AB)]^2 \leq \text{tr}(A^2B^2)$.

(b) $\text{rk}(A) \geq \dfrac{(\text{tr } A)^2}{\text{tr}(A^2)}$, if $\text{tr}(A^2) \neq 0$.

(9) $A\ (m \times m)$ real, $b\ (m \times 1)$ real, $c \in \mathbb{R}$:

$$B = \begin{bmatrix} c & b' \\ b & A \end{bmatrix} \text{ is symmetric} \quad \Rightarrow \det(B) = c \det(A) - b'A^{adj}b.$$

(10) $A\ (m \times m), B\ (n \times n), C\ (m \times n)$:

$$A, B \text{ are symmetric} \quad \Rightarrow \begin{bmatrix} A & C \\ C' & B \end{bmatrix} \text{ is symmetric.}$$

Note: Many of these results are elementary. Most others can be found in Horn & Johnson (1985). More rules for symmetric matrices are given in the section on positive definite and semidefinite matrices which are symmetric by definition if they are real (see Section 9.12).

9.13.2 Symmetry and Duplication Matrices

(1) $A\ (m \times m)$ symmetric, $D_m\ (m^2 \times \frac{1}{2}m(m+1))$ duplication matrix:

(a) $\text{vec}(A) = D_m \text{vech}(A)$.

(b) $D_m^+ \text{vec}(A) = \text{vech}(A)$.

(2) A $(m \times m)$ symmetric nonsingular, $c \in \mathbb{C}$:

(a) $\det(D_m^+[A \otimes A + c\ \text{vec}(A)\text{vec}(A)']D_m) = (1 + cm)\det(A)^{m+1}$.

(b) $(D_m^+[A \otimes A + c\ \text{vec}(A)\text{vec}(A)']D_m)^{-1}$

$$= D_m^+ \left[A^{-1} \otimes A^{-1} - \frac{c}{1+cm} \text{vec}(A^{-1})\text{vec}(A^{-1})' \right] D_m.$$

(3) A, B $(m \times m)$ symmetric, a, b $(m \times 1), \alpha, \beta \in \mathbb{C}$:

(a) $D'_{m+1} \left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix} \right) D_{m+1} =$

$$\begin{bmatrix} \alpha\beta & \alpha b' + \beta a' & (a' \otimes b')D_m \\ \alpha b + \beta a & \alpha B + \beta A + ab' + ba' & (a' \otimes B + b' \otimes A)D_m \\ D'_m(a \otimes b) & D'_m(a \otimes B + b \otimes A) & D'_m(A \otimes B)D_m \end{bmatrix}.$$

(b) $D_{m+1}^+ \left(\begin{bmatrix} \alpha & a' \\ a & A \end{bmatrix} \otimes \begin{bmatrix} \beta & b' \\ b & B \end{bmatrix} \right) D_{m+1}^{+'} =$

$$\begin{bmatrix} \alpha\beta & \frac{1}{2}(\alpha b' + \beta a') & (a' \otimes b')D_m^{+'} \\ \frac{1}{2}(\alpha b + \beta a) & \frac{1}{4}(\alpha B + \beta A + ab' + ba') & \frac{1}{2}(a' \otimes B + b' \otimes A)D_m^{+'} \\ D_m^+(a \otimes b) & \frac{1}{2}D_m^+(a \otimes B + b \otimes A) & D_m^+(A \otimes B)D_m^{+'} \end{bmatrix}.$$

Note: These results are given in Magnus (1988) (see also Section 9.5).

9.13.3 Eigenvalues of Symmetric Matrices

Notation: $\lambda(A)$ denotes an eigenvalue of the matrix A.

$\lambda_{min}(A)$ is the smallest eigenvalue of the matrix A.

$\lambda_{max}(A)$ is the largest eigenvalue of the matrix A.

(1) A $(m \times m)$ real symmetric:

(a) All eigenvalues of A are real numbers.

(b) A is positive definite $\iff$ all eigenvalues of A are real and greater than 0.

(c) A positive semidefinite $\iff$ all eigenvalues of A are real and greater than or equal to 0.

(2) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1, \ldots, \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$:

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i'.$$

(3) (Rayleigh–Ritz theorem)
A $(m \times m)$ real symmetric:

$$\lambda_{min}(A) = \min\left\{\frac{x'Ax}{x'x} \ : \ x\ (m \times 1) \text{ real}, \ x \neq 0\right\},$$

$$\lambda_{max}(A) = \max\left\{\frac{x'Ax}{x'x} \ : \ x\ (m \times 1) \text{ real}, \ x \neq 0\right\}.$$

(4) (Courant–Fischer theorem)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$, $1 < i < m$:

$$\lambda_i = \min_{\substack{y_1, \ldots, y_{m-i} \\ (m \times 1)\text{real}}} \max\left\{\frac{x'Ax}{x'x} : \right.$$

$$\left. x\ (m \times 1) \text{ real}, \ x \neq 0, \ x'y_j = 0, \ j = 1, \ldots, m-i\right\}$$

and

$$\lambda_i = \max_{\substack{y_1, \ldots, y_{i-1} \\ (m \times 1)\text{real}}} \min\left\{\frac{x'Ax}{x'x} : \right.$$

$$\left. x\ (m \times 1) \text{ real}, \ x \neq 0, \ x'y_j = 0, \ j = 1, \ldots, i-1\right\}.$$

(5) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_m$ and associated orthonormal eigenvectors $v_1, \ldots, v_m$, $n \in \{1, \ldots, m\}$:

$$\min\{\text{tr}(X'AX) \ : \ X\ (m \times n) \text{ real}, \ X'X = I_n\} = \lambda_1 + \cdots + \lambda_n.$$

The minimizing matrix is $X = [v_1, \ldots, v_n]$.

$$\max\{\text{tr}(X'AX) \ : \ X\ (m \times n) \text{ real}, \ X'X = I_n\} = \lambda_m + \cdots + \lambda_{m-n+1}.$$

The maximizing matrix is $X = [v_m, \ldots, v_{m-n+1}]$.

Note: For these results see Chapter 5.

9.13.4 Eigenvalue Inequalities

(1) A $(m \times m)$ real symmetric, $x \neq 0$ $(m \times 1)$ real:

$$\lambda_{min}(A) \leq \frac{x'Ax}{x'x} \leq \lambda_{max}(A).$$

(2) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A), \ldots, \lambda_m(A)$:

$$\min\left\{\frac{x'Ax}{x'x} \;:\; x\ (m \times 1) \text{ real, } x \neq 0\right\}$$
$$\leq \lambda_i(A) \leq \max\left\{\frac{x'Ax}{x'x} \;:\; x\ (m \times 1) \text{ real, } x \neq 0\right\},$$

$i = 1, \ldots, m$.

(3) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$ real:

$$X'X = I_n \quad \Rightarrow \quad \sum_{i=1}^{n} \lambda_i(A) \leq \operatorname{tr}(X'AX) \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A).$$

(4) $A = [a_{ij}]$ $(m \times m)$ real symmetric:

$$\lambda_{min}(A) \leq a_{ii} \leq \lambda_{max}(A), \quad i = 1, \ldots, m.$$

(5) $A = [a_{ij}]$ $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\sum_{i=1}^{n} \lambda_i(A) \leq \sum_{i=1}^{n} a_{ii} \leq \sum_{i=1}^{n} \lambda_{m-n+i}(A), \quad n = 1, \ldots, m.$$

(6) $A = [a_{ij}]$ $(m \times m)$ real symmetric with eigenvalues $0 < \lambda_1(A) \leq \cdots \leq \lambda_m(A)$:

$$\prod_{i=1}^{n} \lambda_i(A) \leq \prod_{i=1}^{n} a_{ii}, \quad n = 1, \ldots, m.$$

(7) (Inclusion principle)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, $A_{(n)}$ $(n \times n)$ principal submatrix of A with eigenvalues $\lambda_1(A_{(n)}) \leq \cdots \leq \lambda_n(A_{(n)})$:

$$\lambda_i(A) \leq \lambda_i(A_{(n)}) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n,$$

$$\lambda_{min}(A) \leq \lambda_{min}(A_{(n)}) \leq \lambda_{max}(A_{(n)}) \leq \lambda_{max}(A).$$

(8) A $(m \times m)$ real symmetric, B $(m \times m)$ real positive semidefinite:

$$\lambda_{min}(A + B) \geq \lambda_{min}(A),$$

$$\lambda_{max}(A + B) \geq \lambda_{max}(A).$$

(9) A, B $(m \times m)$ real symmetric:

$$\lambda_{min}(A+B) \geq \lambda_{min}(A) + \lambda_{min}(B),$$

$$\lambda_{max}(A+B) \leq \lambda_{max}(A) + \lambda_{max}(B).$$

(10) A, B $(m \times m)$ real symmetric, $0 < a, b \in \mathbb{R}$:

$$\lambda_{min}(aA+bB) \geq a\lambda_{min}(A) + b\lambda_{min}(B),$$

$$\lambda_{max}(aA+bB) \leq a\lambda_{max}(A) + b\lambda_{max}(B).$$

(11) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, B $(m \times m)$ real positive semidefinite, $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

$$\lambda_i(A+B) \geq \lambda_i(A), \quad i = 1, \ldots, m.$$

(12) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, $\boldsymbol{x}$ $(m \times 1)$ real, $\lambda_1(A \pm \boldsymbol{x}\boldsymbol{x}') \leq \cdots \leq \lambda_m(A \pm \boldsymbol{x}\boldsymbol{x}')$ eigenvalues of $A \pm \boldsymbol{x}\boldsymbol{x}'$:

$$\lambda_i(A \pm \boldsymbol{x}\boldsymbol{x}') \leq \lambda_{i+1}(A) \leq \lambda_{i+2}(A \pm \boldsymbol{x}\boldsymbol{x}'), \quad i = 1, 2, \ldots, m-2;$$

$$\lambda_i(A) \leq \lambda_{i+1}(A \pm \boldsymbol{x}\boldsymbol{x}') \leq \lambda_{i+2}(A), \quad i = 1, 2, \ldots, m-2.$$

(13) A, B $(m \times m)$ real symmetric, $\mathrm{rk}(B) \leq r$, $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ eigenvalues of A, $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

$$\lambda_i(A+B) \leq \lambda_{i+r}(A) \leq \lambda_{i+2r}(A+B), \quad i = 1, 2, \ldots, m-2r;$$

$$\lambda_i(A) \leq \lambda_{i+r}(A+B) \leq \lambda_{i+2r}(A), \quad i = 1, 2, \ldots, m-2r.$$

(14) (Poincare's separation theorem)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, X $(m \times n)$ real such that $X'X = I_n, n \leq m$, $\lambda_1(X'AX) \leq \cdots \leq \lambda_n(X'AX)$ eigenvalues of $X'AX$:

$$\lambda_i(A) \leq \lambda_i(X'AX) \leq \lambda_{m-n+i}(A), \quad i = 1, \ldots, n.$$

(15) A, B $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A+B$:

(a) $\displaystyle\sum_{i=1}^{n}[\lambda_i(A) + \lambda_i(B)] \leq \sum_{i=1}^{n} \lambda_i(A+B), \quad n = 1, \ldots, m.$

(b) $\lambda_1(B) \leq \lambda_n(A+B) - \lambda_n(A) \leq \lambda_m(B), \quad n = 1, \ldots, m.$

(c) $|\lambda_n(A+B) - \lambda_n(A)|_{\text{abs}} \leq \max\{|\lambda_j(B)|_{\text{abs}} \ : \ j = 1, \ldots, m\}$,
$n = 1, \ldots, m$.

(d) $\lambda_n(A+B) \leq \min\{\lambda_i(A) + \lambda_j(B) \ : \ i + j = n + m\}$,
$n = 1, \ldots, m$.

(16) (Weyl's theorem)
A, B $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A+B) \leq \cdots \leq \lambda_m(A+B)$ eigenvalues of $A + B$:

$$\lambda_i(A+B) \geq \begin{cases} \lambda_i(A) + \lambda_1(B) \\ \lambda_{i-1}(A) + \lambda_2(B) \\ \vdots \\ \lambda_1(A) + \lambda_i(B) \end{cases}$$

and

$$\lambda_i(A+B) \leq \begin{cases} \lambda_i(A) + \lambda_m(B) \\ \lambda_{i+1}(A) + \lambda_{m-1}(B) \\ \vdots \\ \lambda_m(A) + \lambda_i(B) \end{cases}$$

$i = 1, \ldots, m$.

(17) A, B $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$ and $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$, respectively; $\lambda_1(A-B) \leq \cdots \leq \lambda_m(A-B)$ eigenvalues of $A - B$:

$$\lambda_1(A-B) > 0 \quad \Rightarrow \quad \lambda_i(A) \geq \lambda_i(B), \ i = 1, \ldots, m.$$

(18) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, x $(m \times 1)$ real, $c \in \mathbb{R}$,

$$B = \begin{bmatrix} A & x \\ x' & c \end{bmatrix}$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_{m+1}(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{i+1}(B), \quad i = 1, \ldots, m.$$

(19) A $(m \times m)$ real symmetric with eigenvalues $\lambda_1(A) \leq \cdots \leq \lambda_m(A)$, D $(p \times p)$ real symmetric, C $(m \times p)$ real and

$$B = \begin{bmatrix} A & C \\ C' & D \end{bmatrix} (n \times n)$$

with eigenvalues $\lambda_1(B) \leq \cdots \leq \lambda_n(B)$:

$$\lambda_i(B) \leq \lambda_i(A) \leq \lambda_{n-m+i}(B) \quad \text{for } i = 1, \ldots, m.$$

Note: For these results see Chapter 5.

9.13.5 Decompositions of Symmetric and Skew-Symmetric Matrices

(1) (Spectral decomposition)
A $(m \times m)$ real symmetric with eigenvalues $\lambda_1, \ldots, \lambda_m$: $A = Q\Lambda Q'$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ and Q is the real orthogonal $(m \times m)$ matrix whose columns are the orthonormal eigenvectors $v_1, \ldots, v_m$ of A associated with $\lambda_1, \ldots, \lambda_m$. In other words,

$$A = \sum_{i=1}^{m} \lambda_i v_i v_i'.$$

(2) A $(m \times m)$ symmetric, $\text{rk}(A) = r$: There exists a nonsingular $(m \times m)$ matrix T such that

$$A = T \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} T',$$

where the zero submatrices disappear if $r = m$.

(3) A $(m \times m)$ real skew-symmetric: There exists a real orthogonal $(m \times m)$ matrix Q such that

$$A = Q \begin{bmatrix} 0 & & & & & 0 \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \Lambda_1 & & \\ & & & & \ddots & \\ 0 & & & & & \Lambda_k \end{bmatrix} Q',$$

where the Λ_i are real (2×2) matrices of the form

$$\Lambda_i = \begin{bmatrix} 0 & b_i \\ -b_i & 0 \end{bmatrix}.$$

(4) (Simultaneous diagonalization of a real positive definite and a real symmetric matrix)
A $(m \times m)$ real positive definite, B $(m \times m)$ real symmetric: There exists a real nonsingular $(m \times m)$ matrix T such that $A = TT'$ and $B = T\Lambda T'$, where Λ is a real diagonal matrix.

Note: For more details on matrix decomposition results see Chapter 6.

9.14 Triangular Matrices

Definitions: An $(m \times m)$ matrix

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ a_{m1} & \cdots & \cdots & a_{mm} \end{bmatrix} = [a_{ij}] ,$$

with $a_{ij} = 0$ for $j > i$, is a lower triangular matrix. An $(m \times m)$ matrix

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1m} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{mm} \end{bmatrix} = [a_{ij}] ,$$

with $a_{ij} = 0$ for $i > j$, is upper triangular. A is triangular if it is upper triangular or lower triangular. A is strictly upper (lower) triangular if $a_{ij} = 0$ for $i \geq j$ $(i \leq j)$, that is, all elements on the principal diagonal in addition to the elements below or above that diagonal are zero. A is strictly triangular if it is strictly upper or strictly lower triangular.

9.14.1 Properties of General Triangular Matrices

(1) A, B $(m \times m)$ upper (lower) triangular, $c \in \mathbb{C}$:

(a) $A \pm B$ is upper (lower) triangular.

(b) cA is upper (lower) triangular.

(c) AB is upper (lower) triangular.

(d) $A \odot B$ is upper (lower) triangular.

(2) A $(m \times m)$ upper (lower) triangular, B $(n \times n)$ upper (lower) triangular:

(a) $A \otimes B$ is upper (lower) triangular.

(b) $A \oplus B$ is upper (lower) triangular.

(3) A $(m \times m)$ upper (lower) triangular:

(a) A' is lower (upper) triangular.

(b) A^H is lower (upper) triangular.

(c) $\bar{A}$ is upper (lower) triangular.

(d) A^{-1} is upper (lower) triangular, if A is nonsingular.

(e) A^{adj} is upper (lower) triangular.

(4) $A = [a_{ij}]$ $(m \times m)$ triangular:

(a) $\text{tr}(A) = a_{11} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}$.

(b) $\det(A) = a_{11} \cdots a_{mm} = \prod_{i=1}^{m} a_{ii}$.

(c) $\text{rk}(A) \geq$ number of nonzero diagonal elements.

(d) $a_{11}, \ldots, a_{mm}$ are the eigenvalues of A.

(e) A is symmetric $\Rightarrow$ $A = \text{diag}(a_{11}, \ldots, a_{mm})$.

(f) A is Hermitian $\Rightarrow$ A is real diagonal.

Note: Most of these results are simple consequences of the definitions. See Magnus (1988) for more details.

9.14.2 Triangularity, Elimination and Duplication Matrices

(1) A $(m \times m)$: A lower triangular $\Rightarrow L'_m \text{vech}(A) = \text{vec}(A)$.

(2) $A = [a_{ij}], B = [b_{ij}]$ $(m \times m)$ upper (lower) triangular:

(a) $D_m^+(A \otimes B)D_m$ is upper (lower) triangular with diagonal elements $\frac{1}{2}(a_{ii}b_{jj} + a_{jj}b_{ii})$, $1 \leq j \leq i \leq m$.

(b) $\det[D_m^+(A \otimes B)D_m] = \frac{1}{2} \prod_{i \geq j}(a_{ii}b_{jj} + a_{jj}b_{ii})$.

(c) $L_m(A \otimes B)L'_m$ is upper (lower) triangular with diagonal elements $a_{ii}b_{jj}$, $1 \leq i \leq j \leq m$.

(d) $\det[L_m(A \otimes B)L'_m] = \prod_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

(e) $L_m(A \otimes B)D_m$ is upper (lower) triangular with diagonal elements $a_{ii}b_{jj}$, $1 \leq i \leq j \leq m$.

(f) $\det[L_m(A \otimes B)D_m] = \prod_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

(3) $A = [a_{ij}], B = [b_{ij}]$ $(m \times m)$ lower triangular:

(a) $\det[L_m(A' \otimes B)L'_m] = \prod_{i=1}^{m} b_{ii}^{i} a_{ii}^{m-i+1}$.

(b) $a_{ii}b_{jj}$, $1 \leq i \leq j \leq m$, are the eigenvalues of $L_m(A' \otimes B)L'_m$.

(c) $[L_m(A' \otimes B)L'_m]^s = L_m[(A')^s \otimes B^s]L'_m$ for

$$\begin{cases} s &= 0,1,2,\dots \\ s &= \dots -2,-1, \text{ if } A, B \text{ nonsingular} \\ s &= \frac{1}{2}, \text{ if lower triangular } A^{1/2},\ B^{1/2} \text{ exist} \end{cases}.$$

(d) $\det[D_m^+(A \otimes A \pm B \otimes B)D_m] = \prod_{i \geq j}(a_{ii}a_{jj} \pm b_{ii}b_{jj})$.

(e) $$\begin{aligned} \det[L_m(AB' \otimes B'A)L_m'] &= \det[L_m(AB' \otimes A'B)L_m'] \\ &= (\det A)^{m+1}(\det B)^{m+1}. \end{aligned}$$

(f) $L_m' L_m(A' \otimes B)L_m' = (A' \otimes B)L_m'$.

(4) $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}], D = [d_{ij}]$ $(m \times m)$ lower triangular :

$$\det[L_m(AB' \circledast C'D)L_m'] = \prod_{i=1}^{m}(c_{ii}d_{ii})^i(a_{ii}b_{ii})^{m-i+1},$$

$$\det[L_m(A' \circledast B + C' \circledast D)L_m'] = \prod_{i \geq j}(b_{ii}a_{jj} + d_{ii}c_{jj}).$$

(5) A, B, C, D $(m \times m)$ nonsingular, lower triangular:

$$[L_m(AB' \circledast C'D)L_m']^{-1} = L_m[(B')^{-1} \otimes D^{-1}]L_m' L_m[A^{-1} \otimes (C')^{-1}]L_m'.$$

(6) $A = [a_{ij}]$ $(m \times m)$ lower triangular, $n \in \mathbb{N}$:

$$\det\left(L_m\left(\sum_{j=0}^{n-1}(A')^{n-1-j} \otimes A^j\right)L_m'\right) = n^m(\det A)^{n-1}\prod_{k>l}\mu_{kl},$$

where

$$\mu_{kl} = \begin{cases} (a_{kk}^n - a_{ll}^n)/(a_{kk} - a_{ll}) & \text{if } a_{kk} \neq a_{ll} \\ n a_{kk}^{n-1} & \text{if } a_{kk} = a_{ll} \end{cases}.$$

(7) $A = [a_{ij}], B = [b_{ij}], C = [c_{ij}]$ $(m \times m)$ lower triangular:

(a) $L_m(A \circledast B')D_m = L_m(A \otimes B')L_m'$.

(b) $$\begin{aligned} \det[L_m(AB \otimes B'A)D_m] &= \det[L_m(AB' \circledast A'B')D_m] \\ &= \det[L_m(AB' \otimes B'A)L_m'] \\ &= (\det A)^{m+1}(\det B)^{m+1}. \end{aligned}$$

(c) $\det[L_m(A \otimes B'C)D_m] = \prod_{i=1}^{m}(b_{ii}c_{ii})^i a_{ii}^{m-i+1}$.

(d) $\det[L_m(AB' \otimes C')D_m] = \prod_{i=1}^{m} c_{ii}^i (a_{ii}b_{ii})^{m-i+1}$.

Note: The foregoing results may be found in Magnus (1988) which is a rich source of results for triangular matrices and their relationship to duplication and elimination matrices.

9.14.3 Properties of Strictly Triangular Matrices

(1) A $(m \times m)$ strictly triangular:

(a) A is singular.

(b) $A^m = O_{m\times m}$.

(c) $\mathrm{tr}(A) = 0$.

(d) $\det(A) = 0$.

(2) A, B $(m \times m)$ strictly upper (lower) triangular, $c \in \mathbb{C}$:

(a) $A \pm B$ is strictly upper (lower) triangular.

(b) cA is strictly upper (lower) triangular.

(c) AB is strictly upper (lower) triangular.

(d) $A \odot B$ is strictly upper (lower) triangular.

(e) $L_m(A' \otimes B)L_m'$ is nilpotent.

(3) A $(m \times m)$ strictly upper (lower) triangular, B $(n \times n)$ strictly upper (lower) triangular:

(a) $A \otimes B$ is strictly upper (lower) triangular.

(b) $A \oplus B$ is strictly upper (lower) triangular.

(4) A $(m \times m)$ strictly upper (lower) triangular:

(a) A' is strictly lower (upper) triangular.

(b) A^H is strictly lower (upper) triangular.

(c) $\bar{A}$ is strictly upper (lower) triangular.

(d) A^{adj} is strictly upper (lower) triangular.

Note: These results are easy implications of the definitions.

9.15 Unitary Matrices

Definition: An $(m \times m)$ matrix A is unitary if it is nonsingular and $A^{-1} = A^H$.

(1) A $(m \times m)$:

(a) A is unitary $\iff$ the columns of A are orthonormal vectors.

(b) A is unitary $\iff$ the rows of A are orthonormal vectors.

(c) A is unitary $\iff$ $x^H A^H A y = x^H y$ for all $(m \times 1)$ vectors x, y.

(2) A $(m \times m)$: A is unitary $\iff$ $AA^H = A^H A = I_m$.

$\iff A'$ is unitary.
$\iff A^H$ is unitary.
$\iff \bar{A}$ is unitary.
$\iff A^{-1}$ is unitary.
$\iff A^i$ is unitary for $i = 1, 2, \ldots$

(3) A $(m \times m)$ real: A is orthogonal $\iff$ A is unitary.

(4) A $(m \times m)$ unitary:

(a) $|\det(A)|_{\text{abs}} = 1$.

(b) $\text{rk}(A) = m$.

(c) A is normal, that is, $A^H A = AA^H$.

(d) λ is eigenvalue of $A \Rightarrow |\lambda|_{\text{abs}} = 1$.

(e) x $(m \times 1)$ $\Rightarrow \|Ax\|_2 = \|x\|_2$.

(f) B $(n \times m)$ $\Rightarrow \|BA\|_2 = \|B\|_2$.

(g) B $(m \times n)$ $\Rightarrow \|AB\|_2 = \|B\|_2$.

(5) A_i $(m \times m), i = 1, 2, \ldots$:
A_i is unitary for $i = 1, 2, \ldots$ and $\lim_{i\to\infty} A_i$ exists $\Rightarrow$ $\lim_{i\to\infty} A_i$ is unitary.

(6) A, B $(m \times m)$: A, B are unitary $\Rightarrow$ AB is unitary.

(7) A $(m \times m)$ unitary, B $(n \times n)$ unitary:

(a) $A \otimes B$ is unitary.

(b) $A \oplus B$ is unitary.

(8) A $(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$:
A is normal, $|\lambda_i|_{\text{abs}} = 1, i = 1, \ldots, m \iff A$ is unitary.

(9) (a) The $(m \times m)$ unit matrix I_m is unitary.

(b) The $(mn \times mn)$ commutation matrix K_{mn} is unitary.

(c) A permutation matrix is unitary.

(d) For $\theta \in \mathbb{R}$, $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is unitary.

(e) (Householder transformation)
For x $(m \times 1)$, $I_m - \dfrac{2xx^H}{x^H x}$ is unitary.

(10) A, B $(m \times m)$:

(a) A is similar to a unitary matrix $\Rightarrow$ A^{-1} is similar to A^H.

(b) A, B are unitarily equivalent $\iff$ A, B have the same singular values.

(c) A, B are unitarily equivalent $\iff$ $A^H A$ and $B^H B$ are similar.

(11) A $(m \times m)$ unitary : There exists a real orthogonal $(m \times m)$ matrix Q and a real symmetric $(m \times m)$ matrix S such that $A = Q\exp(\mathrm{i}S)$.

Note: The results of this subsection follow either directly from the definitions or can be found in Horn & Johnson (1985). A number of results on unitary matrices are also given in Lancaster & Tismenetsky (1985).

10

Vector and Matrix Derivatives

In this chapter all vectors and matrices are assumed to be real unless otherwise stated. Differentiable always means continuously differentiable of sufficiently high order so that all expressions are well-defined.

10.1 Notation

Let $f(\boldsymbol{x})$ be a differentiable real valued function of the real $(m \times 1)$ vector $\boldsymbol{x} = (x_1, \dots, x_m)'$.

$$\frac{\partial f}{\partial \boldsymbol{x}} \equiv \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_m} \end{bmatrix} \quad \text{or} \quad \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} \equiv \begin{bmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_m} \end{bmatrix} \qquad (m \times 1)$$

and

$$\frac{\partial f}{\partial \boldsymbol{x}'} \equiv \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right) \quad \text{or} \quad \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}'} \equiv \left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \dots, \frac{\partial f(\boldsymbol{x})}{\partial x_m} \right) \qquad (1 \times m)$$

are vectors of first order partial derivatives. $\partial f(\boldsymbol{x})/\partial \boldsymbol{x}$ is sometimes called the **gradient vector** of $f(\boldsymbol{x})$.

$$\left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_0} = \frac{\partial f(\boldsymbol{x}_0)}{\partial \boldsymbol{x}} \equiv \begin{bmatrix} \left.\dfrac{\partial f}{\partial x_1}\right|_{\boldsymbol{x}=\boldsymbol{x}_0} \\ \vdots \\ \left.\dfrac{\partial f}{\partial x_m}\right|_{\boldsymbol{x}=\boldsymbol{x}_0} \end{bmatrix} \qquad (m \times 1)$$

and

$$\left.\frac{\partial f}{\partial \boldsymbol{x}'}\right|_{\boldsymbol{x}_0} = \left.\frac{\partial f}{\partial \boldsymbol{x}'}\right|_{\boldsymbol{x}=\boldsymbol{x}_0} = \frac{\partial f(\boldsymbol{x}_0)}{\partial \boldsymbol{x}'} \equiv \left[\frac{\partial f(\boldsymbol{x}_0)}{\partial \boldsymbol{x}}\right]' \qquad (1 \times m)$$

are vectors of **first order partial derivatives evaluated at** the $(m \times 1)$ vector x_0.

$$\frac{\partial^2 f}{\partial x \partial x'} = \frac{\partial^2 f(x)}{\partial x \partial x'} \equiv \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right] = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_m \partial x_m} \end{bmatrix} \qquad (m \times m)$$

is the **Hessian matrix** of second order partial derivatives of $f(x)$ and

$$\left.\frac{\partial^2 f}{\partial x \partial x'}\right|_{x=x_0} = \left.\frac{\partial^2 f}{\partial x \partial x'}\right|_{x_0} = \frac{\partial^2 f(x_0)}{\partial x \partial x'} \equiv \left[\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{x=x_0}\right] \qquad (m \times m)$$

is the **Hessian matrix** of second order partial derivatives of $f(x)$ **evaluated at** $x = x_0$.

Let $f(X)$ be a differentiable real valued function of the real $(m \times n)$ matrix $X = [x_{ij}]$.

$$\frac{\partial f}{\partial X} = \frac{\partial f(X)}{\partial X} \equiv \left[\frac{\partial f(X)}{\partial x_{ij}}\right] = \begin{bmatrix} \dfrac{\partial f}{\partial x_{11}} & \cdots & \dfrac{\partial f}{\partial x_{1n}} \\ \vdots & & \vdots \\ \dfrac{\partial f}{\partial x_{m1}} & \cdots & \dfrac{\partial f}{\partial x_{mn}} \end{bmatrix} \qquad (m \times n)$$

is the matrix of first order partial derivatives of $f(X)$ and

$$\left.\frac{\partial f}{\partial X}\right|_{X=X_0} = \left.\frac{\partial f}{\partial X}\right|_{X_0} = \frac{\partial f(X_0)}{\partial X} \equiv \left[\left.\frac{\partial f}{\partial x_{ij}}\right|_{X=X_0}\right] \qquad (m \times n)$$

is the matrix of first order partial derivatives of $f(X)$ evaluated at the $(m \times n)$ matrix X_0.

$$\frac{\partial^2 f}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = \frac{\partial^2 f(X)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} \qquad (mn \times mn)$$

is the **Hessian matrix** of second order partial derivatives of $f(X)$. Here it is important to note the order in which the partial derivatives are arranged. They have the same order as for the function $f(\mathrm{vec}(X))$. Accordingly,

$$\left.\frac{\partial^2 f}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'}\right|_{X=X_0} = \frac{\partial^2 f(X_0)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'}$$

is the Hessian matrix evaluated at $X = X_0$.

Let $y(x) = [y_1(x), \ldots, y_n(x)]'$ be a real $(n \times 1)$ vector of differentiable functions of the real $(m \times 1)$ vector $x = (x_1, \ldots, x_m)'$, that is, y is a function mapping a subset of $\mathbb{R}^m$ on a subset of $\mathbb{R}^n$.

$$\frac{\partial y}{\partial x'} = \frac{\partial y(x)}{\partial x'} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix} \qquad (n \times m)$$

and

$$\frac{\partial y'}{\partial x} = \frac{\partial y(x)'}{\partial x} \equiv \left(\frac{\partial y}{\partial x'}\right)' \qquad (m \times n)$$

are matrices of first order partial derivatives of $y(x)$. $\partial y/\partial x'$ is sometimes called the **Jacobian matrix** of y and $\partial y'/\partial x$ is sometimes called the **gradient** of y.

$$\left.\frac{\partial y}{\partial x'}\right|_{x=x_0} = \left.\frac{\partial y}{\partial x'}\right|_{x_0} = \frac{\partial y(x_0)}{\partial x'} \equiv \begin{bmatrix} \left.\frac{\partial y_1}{\partial x_1}\right|_{x=x_0} & \cdots & \left.\frac{\partial y_1}{\partial x_m}\right|_{x=x_0} \\ \vdots & & \vdots \\ \left.\frac{\partial y_n}{\partial x_1}\right|_{x=x_0} & \cdots & \left.\frac{\partial y_n}{\partial x_m}\right|_{x=x_0} \end{bmatrix}$$

is the Jacobian matrix evaluated at $x = x_0$ and its transpose is the gradient of y evaluated at x_0. For $m = n$,

$$\det\left(\frac{\partial y(x)}{\partial x'}\right) = \det\left(\frac{\partial y}{\partial x'}\right)$$

is the **Jacobian** or **Jacobian determinant** of $y(x)$. The **Hessian matrix** of $y(x)$ is

$$\frac{\partial \mathrm{vec}(\partial y/\partial x')}{\partial x'}.$$

For a matrix valued function of a single variable, $A : S \subset \mathbb{R} \to \mathbb{R}^{m\times n}$, $x \mapsto A(x) = [a_{ij}(x)]$, the matrix of first order derivatives is

$$\frac{\mathrm{d}A(x)}{\mathrm{d}x} = \frac{\mathrm{d}A}{\mathrm{d}x} \equiv \begin{bmatrix} \frac{\mathrm{d}a_{11}}{\mathrm{d}x} & \cdots & \frac{\mathrm{d}a_{1n}}{\mathrm{d}x} \\ \vdots & & \vdots \\ \frac{\mathrm{d}a_{m1}}{\mathrm{d}x} & \cdots & \frac{\mathrm{d}a_{mn}}{\mathrm{d}x} \end{bmatrix}$$

and the corresponding matrix of first order derivatives evaluated at x_0 is

$$\frac{\mathrm{d}A(x_0)}{\mathrm{d}x} = \left.\frac{\mathrm{d}A}{\mathrm{d}x}\right|_{x=x_0} \equiv \left[\left.\frac{\mathrm{d}a_{ij}}{\mathrm{d}x}\right|_{x=x_0}\right].$$

In the remainder of this chapter all functions are assumed to be differentiable.

10.2 Gradients and Hessian Matrices of Real Valued Functions with Vector Arguments

10.2.1 Gradients

(1) $x\ (m \times 1), c \in \mathbb{R}$ constant: $\dfrac{\partial c}{\partial x} = O_{m\times 1}$.

(2) (Linearity)
$x\ (m \times 1), f(x), g(x)$ real valued functions, $c_1, c_2 \in \mathbb{R}$:

$$\frac{\partial[c_1 f(x) + c_2 g(x)]}{\partial x} = c_1 \frac{\partial f(x)}{\partial x} + c_2 \frac{\partial g(x)}{\partial x}.$$

(3) (Product rule)
$x\ (m \times 1), f(x), g(x)$ real valued functions:

$$\frac{\partial f(x)g(x)}{\partial x} = f(x)\frac{\partial g(x)}{\partial x} + g(x)\frac{\partial f(x)}{\partial x}.$$

(4) $x\ (m \times 1), f(x), g(x), h(x)$ real valued functions:

$$\frac{\partial f(x)g(x)h(x)}{\partial x} = f(x)g(x)\frac{\partial h(x)}{\partial x} + f(x)h(x)\frac{\partial g(x)}{\partial x} + g(x)h(x)\frac{\partial f(x)}{\partial x}.$$

(5) (Ratio rule)
$x\ (m \times 1), f(x), g(x) \neq 0$ real valued functions:

$$\frac{\partial[f(x)/g(x)]}{\partial x} = \frac{1}{g(x)^2}\left[g(x)\frac{\partial f(x)}{\partial x} - f(x)\frac{\partial g(x)}{\partial x}\right].$$

(6) (Chain rule)
$x\ (m \times 1), y(x)\ (n \times 1), f(y)$ a real valued function:

$$\frac{\partial f(y(x))}{\partial x} = \frac{\partial y(x)'}{\partial x}\frac{\partial f(y)}{\partial y}.$$

(7) $x, a\ (m \times 1)$: $\dfrac{\partial a'x}{\partial x} = a$.

(8) $x\ (m \times 1),\ a, y(x)\ (n \times 1)$: $\dfrac{\partial a'y(x)}{\partial x} = \dfrac{\partial y(x)'}{\partial x}a$.

(9) $x\ (m \times 1),\ A\ (m \times m)$: $\dfrac{\partial x'Ax}{\partial x} = (A + A')x$.

(10) $x\ (m \times 1), a\ (n \times 1), A\ (n \times m), B\ (n \times n)$ symmetric:

$$\frac{(a - Ax)'B(a - Ax)}{\partial x} = -2A'B(a - Ax).$$

(11) $x\ (m \times 1), y(x)\ (n \times 1), A\ (n \times n)$:

$$\frac{\partial y(x)'Ay(x)}{\partial x} = \frac{\partial y(x)'}{\partial x}(A + A')y(x).$$

(12) $x\ (m \times 1), y(x)\ (n \times 1), z(x)\ (p \times 1), A\ (n \times p)$:

$$\frac{\partial y(x)'Az(x)}{\partial x} = \frac{\partial y(x)'}{\partial x}Az(x) + \frac{\partial z(x)'}{\partial x}A'y(x).$$

Note: (1) – (6) are standard rules of calculus for functions with vector arguments. For the chain rule see, e.g., Magnus & Neudecker (1988, Chapter 5, Sec. 12). (7) and (8) follow from (2). (9) is obtained from standard rules for derivatives by writing $x'Ax$ in summation notation and computing the derivative for each individual component of x. (10) and (11) follow from (9) and the chain rule (6). (12) is a consequence of the product and chain rules.

10.2.2 Hessian Matrices

(1) $x, a\ (m \times 1)$: $\quad \dfrac{\partial^2 a'x}{\partial x \partial x'} = O_{m \times m}.$

(2) $x\ (m \times 1), A\ (m \times m)$: $\quad \dfrac{\partial^2 x'Ax}{\partial x \partial x'} = A + A'.$

(3) $x\ (m \times 1), a\ (n \times 1), A\ (n \times m), B\ (n \times n)$ symmetric:

$$\frac{\partial^2 (a - Ax)'B(a - Ax)}{\partial x \partial x'} = 2A'BA.$$

(4) $x\ (m \times 1), y(x)\ (n \times 1), A\ (n \times n)$:

$$\begin{aligned}\frac{\partial^2 y(x)'Ay(x)}{\partial x \partial x'} &= \frac{\partial y(x)'}{\partial x}(A + A')\frac{\partial y(x)}{\partial x'} \\ &\quad + [y(x)'(A + A') \otimes I_m]\frac{\partial}{\partial x'}\left[\text{vec}\left(\frac{\partial y(x)'}{\partial x}\right)\right].\end{aligned}$$

(5) $x\ (m \times 1), y(x)\ (n \times 1), z(x)\ (p \times 1), A\ (n \times p)$:

$$\begin{aligned}\frac{\partial^2 y(x)'Az(x)}{\partial x \partial x'} &= \frac{\partial y(x)'}{\partial x}A\frac{\partial z(x)}{\partial x'} + \frac{\partial z(x)'}{\partial x}A'\frac{\partial y(x)}{\partial x'} \\ &\quad + [z(x)'A' \otimes I_m]\frac{\partial}{\partial x'}\left[\text{vec}\left(\frac{\partial y(x)'}{\partial x}\right)\right] \\ &\quad + [y(x)'A \otimes I_m]\frac{\partial}{\partial x'}\left[\text{vec}\left(\frac{\partial z(x)'}{\partial x}\right)\right].\end{aligned}$$

Note: All results are simple consequences of the results in the previous subsection and the rules for vector valued functions with vector arguments given in the following sections.

10.3 Derivatives of Real Valued Functions with Matrix Arguments

10.3.1 General and Miscellaneous Rules

(1) X $(m \times n)$, $c \in \mathbb{R}$ constant: $\dfrac{\partial c}{\partial X} = O_{m \times n}$.

(2) (Linearity)
X $(m \times n)$, $f(X)$, $g(X)$ real valued functions, $c_1, c_2 \in \mathbb{R}$:

$$\frac{\partial[c_1 f(X) + c_2 g(X)]}{\partial X} = c_1 \frac{\partial f(X)}{\partial X} + c_2 \frac{\partial g(X)}{\partial X}.$$

(3) (Product rule)
X $(m \times n)$, $f(X)$, $g(X)$ real valued functions:

$$\frac{\partial f(X) g(X)}{\partial X} = f(X) \frac{\partial g(X)}{\partial X} + g(X) \frac{\partial f(X)}{\partial X}.$$

(4) X $(m \times n)$, $f(X)$, $g(X)$, $h(X)$ real valued functions:

$$\begin{aligned} \frac{\partial f(X) g(X) h(X)}{\partial X} &= f(X) g(X) \frac{\partial h(X)}{\partial X} \\ &\quad + f(X) h(X) \frac{\partial g(X)}{\partial X} + g(X) h(X) \frac{\partial f(X)}{\partial X}. \end{aligned}$$

(5) (Ratio rule)
X $(m \times n)$, $f(X)$, $g(X) \neq 0$ real valued functions:

$$\frac{\partial [f(X)/g(X)]}{\partial X} = \frac{1}{g(X)^2} \left[g(X) \frac{\partial f(X)}{\partial X} - f(X) \frac{\partial g(X)}{\partial X} \right].$$

(6) (Chain rule)
X $(m \times n)$, $y = f(X)$, $g(y)$ real valued functions:

$$\frac{\partial g(f(X))}{\partial X} = \frac{\mathrm{d} g(y)}{\mathrm{d} y} \frac{\partial f(X)}{\partial X}.$$

(7) X $(m \times n)$, $f(X)$ a real valued function:

$$\mathrm{vec}\left(\frac{\partial f(X)}{\partial X} \right) = \frac{\partial f(X)}{\partial \mathrm{vec}(X)}.$$

(8) X $(m \times n), a$ $(m \times 1), b$ $(n \times 1)$:

$$\frac{\partial a'Xb}{\partial X} = ab'.$$

(9) X $(m \times m), a, b$ $(m \times 1)$:

$$\frac{\partial a'X^ib}{\partial X} = \sum_{j=0}^{i-1}(X^j)'ab'(X^{i-1-j})', \quad i = 1, 2, \ldots$$

(10) X $(m \times m)$ nonsingular, a, b $(m \times 1)$:

$$\frac{\partial a'X^{-1}b}{\partial X} = -(X^{-1})'ab'(X^{-1})'.$$

(11) X $(m \times n), a, b$ $(n \times 1)$:

$$\frac{\partial a'X'Xb}{\partial X} = X(ba' + ab').$$

(12) X $(m \times n), a, b$ $(m \times 1)$:

$$\frac{\partial a'XX'b}{\partial X} = (ba' + ab')X.$$

(13) X $(m \times m)$ symmetric with simple eigenvalue λ and corresponding eigenvector v satisfying $v'v = 1$:

$$\frac{\partial \lambda}{\partial X} = vv'.$$

Note: (1) – (7) are standard results from calculus. (9) and (10) are given in Magnus & Neudecker (1988, Chapter 9, Sec. 13) and (8) is a special case of (9). (11) and (12) follow from results given in Magnus & Neudecker (1988, Chapter 9, Sec. 9) by noting that $a'Ab = \text{tr}(Aba')$. Result (13) is from Magnus & Neudecker (1988, Chapter 9, Sec. 11).

10.3.2 Derivatives of the Trace

First Order Derivatives

(1) X $(m \times m)$: $\dfrac{\partial \text{tr}(X)}{\partial X} = \dfrac{\partial \text{tr}(X')}{\partial X} = I_m.$

(2) X $(m \times n)$, A $(n \times m)$: $\dfrac{\partial \text{tr}(AX)}{\partial X} = \dfrac{\partial \text{tr}(XA)}{\partial X} = A'.$

(3) X $(m \times n)$, A $(m \times n)$: $\dfrac{\partial \text{tr}(X'A)}{\partial X} = \dfrac{\partial \text{tr}(AX')}{\partial X} = A.$

(4) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times p):\quad \dfrac{\partial \mathrm{tr}(AXB)}{\partial X} = A'B'.$

(5) $X\ (m \times n),\ A\ (p \times n),\ B\ (m \times p):\quad \dfrac{\partial \mathrm{tr}(AX'B)}{\partial X} = BA.$

(6) $X\ (m \times m):\quad \dfrac{\partial \mathrm{tr}(X^2)}{\partial X} = 2X',$

$$\frac{\partial \mathrm{tr}(X^i)}{\partial X} = i(X')^{i-1}, \quad i = 1, 2, \ldots$$

(7) $X, A, B\ (m \times m):$

$$\frac{\partial \mathrm{tr}(AX^iB)}{\partial X} = \sum_{j=0}^{i-1} (X^j)'A'B'(X^{i-1-j})', \quad i = 1, 2, \ldots$$

(8) $X\ (m \times n):\quad \dfrac{\partial \mathrm{tr}(X'X)}{\partial X} = \dfrac{\partial \mathrm{tr}(XX')}{\partial X} = 2X.$

(9) $X\ (m \times n),\ A\ (m \times m):\quad \dfrac{\partial \mathrm{tr}(X'AX)}{\partial X} = (A + A')X.$

(10) $X\ (m \times n),\ A\ (m \times m)$ symmetric : $\dfrac{\partial \mathrm{tr}(X'AX)}{\partial X} = 2AX.$

(11) $X\ (m \times n),\ A\ (n \times n):\quad \dfrac{\partial \mathrm{tr}(XAX')}{\partial X} = X(A + A').$

(12) $X\ (m \times n),\ A\ (n \times n)$ symmetric : $\dfrac{\partial \mathrm{tr}(XAX')}{\partial X} = 2XA.$

(13) $X, A\ (m \times m):\quad \dfrac{\partial \mathrm{tr}(XAX)}{\partial X} = X'A' + A'X'.$

(14) $X\ (m \times n),\ A\ (p \times m):\quad \dfrac{\partial \mathrm{tr}(AXX'A')}{\partial X} = 2A'AX.$

(15) $X\ (m \times n),\ A\ (p \times n):\quad \dfrac{\partial \mathrm{tr}(AX'XA')}{\partial X} = 2XA'A.$

(16) $X\ (m \times n),\ A\ (p \times m),\ B\ (m \times p):\quad \dfrac{\partial \mathrm{tr}(AXX'B)}{\partial X} = (BA + A'B')X.$

(17) $X\ (m \times n),\ A\ (p \times n),\ B\ (n \times p):\quad \dfrac{\partial \mathrm{tr}(AX'XB)}{\partial X} = X(BA + A'B').$

(18) $X\ (m \times n),\ A\ (n \times n),\ B\ (m \times m):\quad \dfrac{\partial \mathrm{tr}(XAX'B)}{\partial X} = B'XA' + BXA.$

(19) $X\ (m \times n),\ A, B\ (n \times m):\quad \dfrac{\partial \mathrm{tr}(XAXB)}{\partial X} = B'X'A' + A'X'B'.$

(20) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times n),\ C\ (m \times p)$:

$$\frac{\partial \operatorname{tr}(AXBX'C)}{\partial X} = A'C'XB' + CAXB.$$

(21) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times m),\ C\ (n \times p)$:

$$\frac{\partial \operatorname{tr}(AXBXC)}{\partial X} = A'C'X'B' + B'X'A'C'.$$

(22) $X\ (m \times m)$ nonsingular : $\dfrac{\partial \operatorname{tr}(X^{-1})}{\partial X} = -(X^{-2})'.$

(23) $X\ (m \times m)$ nonsingular, $A, B\ (m \times m)$:

$$\frac{\partial \operatorname{tr}(AX^{-1}B)}{\partial X} = -(X^{-1}BAX^{-1})'.$$

(24) $X\ (m \times n),\ F(X)\ (p \times p)$: $\dfrac{\partial \operatorname{tr}[F(X)]}{\partial \operatorname{vec}(X)'} = \operatorname{vec}(I_p)' \dfrac{\partial \operatorname{vec} F(X)}{\partial \operatorname{vec}(X)'}.$

(25) $X\ (m \times n),\ F(X)\ (p \times q),\ G(X)\ (r \times s),\ A\ (q \times r),\ B\ (s \times p)$:

$$\frac{\partial \operatorname{tr}[F(X)AG(X)B]}{\partial \operatorname{vec}(X)'} =$$

$$\operatorname{vec}(I_p)' \left[[B'G(X)'A' \otimes I_p] \frac{\partial \operatorname{vec} F(X)}{\partial \operatorname{vec}(X)'} + [B' \otimes F(X)A] \frac{\partial \operatorname{vec} G(X)}{\partial \operatorname{vec}(X)'} \right].$$

Hessian Matrices

(1) $X\ (m \times m)$: $\dfrac{\partial^2 \operatorname{tr}(X)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = O_{m^2 \times m^2}.$

(2) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times p)$: $\dfrac{\partial^2 \operatorname{tr}(AXB)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = O_{mn \times mn}.$

(3) $X\ (m \times n),\ A\ (p \times n),\ B\ (m \times p)$: $\dfrac{\partial^2 \operatorname{tr}(AX'B)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = O_{mn \times mn}.$

(4) $X\ (m \times m)$: $\dfrac{\partial^2 \operatorname{tr}(X^2)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = 2K_{mm},$

$$\frac{\partial^2 \operatorname{tr}(X^i)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = iK_{mm} \left(\sum_{j=0}^{i-2} (X')^{i-2-j} \otimes X^j \right), \quad i = 2, 3, \ldots$$

(5) $X\ (m \times n)$: $\dfrac{\partial^2 \operatorname{tr}(X'X)}{\partial \operatorname{vec}(X) \partial \operatorname{vec}(X)'} = 2I_{mn}.$

(6) X $(m \times n)$, A $(m \times m)$: $\dfrac{\partial^2 \mathrm{tr}(X'AX)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = I_n \otimes (A + A')$.

(7) X $(m \times n)$, A $(m \times m)$ symmetric: $\dfrac{\partial^2 \mathrm{tr}(X'AX)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = 2(I_n \otimes A)$.

(8) X $(m \times n)$, A $(n \times n)$: $\dfrac{\partial^2 \mathrm{tr}(XAX')}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = (A' + A) \otimes I_m$.

(9) X $(m \times n)$, A $(n \times n)$ symmetric: $\dfrac{\partial^2 \mathrm{tr}(XAX')}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = 2(A \otimes I_m)$.

(10) X, A $(m \times m)$: $\dfrac{\partial^2 \mathrm{tr}(XAX)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = (A \otimes I_m + I_m \otimes A')K_{mm}$.

(11) X $(m \times n)$, A $(p \times m)$: $\dfrac{\partial^2 \mathrm{tr}(AXX'A')}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = 2(I_n \otimes A'A)$.

(12) X $(m \times n)$, A $(p \times n)$: $\dfrac{\partial^2 \mathrm{tr}(AX'XA')}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = 2(A'A \otimes I_m)$.

(13) X $(m \times n)$, A $(p \times m)$, B $(m \times p)$:

$$\frac{\partial^2 \mathrm{tr}(AXX'B)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = I_n \otimes (BA + A'B').$$

(14) X $(m \times n)$, A $(p \times n)$, B $(n \times p)$:

$$\frac{\partial^2 \mathrm{tr}(AX'XB)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = (BA + A'B') \otimes I_m.$$

(15) X $(m \times n)$, A $(n \times n)$, B $(m \times m)$:

$$\frac{\partial^2 \mathrm{tr}(XAX'B)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = A \otimes B' + A' \otimes B.$$

(16) X $(m \times n)$, A, B $(n \times m)$:

$$\frac{\partial^2 \mathrm{tr}(XAXB)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = (A \otimes B' + B \otimes A')K_{mn}.$$

(17) X $(m \times n)$, A $(p \times m)$, B $(n \times n)$, C $(m \times p)$:

$$\frac{\partial^2 \mathrm{tr}(AXBX'C)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = B \otimes A'C' + B' \otimes CA.$$

(18) X $(m \times n)$, A $(p \times m)$, B $(n \times m)$, C $(n \times p)$:

$$\frac{\partial^2 \mathrm{tr}(AXBXC)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = (B \otimes A'C' + CA \otimes B')K_{mn}.$$

(19) X $(m \times m)$ nonsingular :

$$\frac{\partial^2 \operatorname{tr}(X^{-1})}{\partial \operatorname{vec}(X)\partial \operatorname{vec}(X)'} = K_{mm}(X'^{-2} \otimes X^{-1} + X'^{-1} \otimes X^{-2}).$$

(20) X $(m \times m)$ nonsingular, A, B $(m \times m)$:

$$\frac{\partial^2 \operatorname{tr}(AX^{-1}B)}{\partial \operatorname{vec}(X)\partial \operatorname{vec}(X)'}$$
$$= K_{mm}(X'^{-1}A'B'X'^{-1} \otimes X^{-1} + X'^{-1} \otimes X^{-1}BAX^{-1}).$$

Note: The results of this subsection are partly given in Magnus & Neudecker (1988, Chapter 9, Sec. 9 and 13) and partly follow from their results by straightforward application of the rules for the trace and matrix derivatives, notably the chain rule and the product rules of Section 10.5.

10.3.3 Derivatives of Determinants

(1) X $(m \times m)$ nonsingular : $\dfrac{\partial \det(X)}{\partial X} = \det(X)(X')^{-1} = (X^{adj})'.$

(2) X $(m \times n)$, A $(p \times m)$, B $(n \times p)$, $\det(AXB) \neq 0$:

$$\frac{\partial \det(AXB)}{\partial X} = \det(AXB)A'(B'X'A')^{-1}B'.$$

(3) X $(m \times n)$, $\operatorname{rk}(X) = m$: $\dfrac{\partial \det(XX')}{\partial X} = 2\det(XX')\,(XX')^{-1}X.$

(4) X $(m \times n)$, $\operatorname{rk}(X) = n$: $\dfrac{\partial \det(X'X)}{\partial X} = 2\det(X'X)\,X(X'X)^{-1}.$

(5) X $(m \times m)$ nonsingular:

(a) $\dfrac{\partial \det(X^2)}{\partial X} = 2(\det X)^2(X')^{-1}.$

(b) $\dfrac{\partial \det(X^{-1})}{\partial X} = -(\det X)^{-1}(X')^{-1}.$

(c) $\dfrac{\partial \det(X^i)}{\partial X} = i\,(\det X)^i(X')^{-1}, \quad i = \pm 1, \pm 2, \ldots$

(6) X, A, B $(m \times m)$, $\det(AX^iB) \neq 0$:

$$\frac{\partial \det(AX^iB)}{\partial X} = \det(AX^iB)\sum_{j=1}^{i-1}[(X^{i-1-j})B(AX^iB)^{-1}AX^j]'.$$

(7) $X\ (m \times n)$, $A\ (p \times m)$, $B\ (n \times n)$, $C\ (m \times p)$, $\det(AXBX'C) \neq 0$:

$$\frac{\partial \det(AXBX'C)}{\partial X} =$$
$$\det(AXBX'C)\{C[AXBX'C]^{-1}AXB + A'[C'XB'X'A']^{-1}C'XB'\}.$$

(8) $X\ (m \times n)$, $A\ (p \times n)$, $B\ (m \times m)$, $C\ (n \times p)$, $\det(AX'BXC) \neq 0$:

$$\frac{\partial \det(AX'BXC)}{\partial X} =$$
$$\det(AX'BXC)\{BXC[AX'BXC]^{-1}A + B'XA'[C'X'B'XA']^{-1}C'\}.$$

(9) $X\ (m \times n)$, $A\ (p \times m)$, $B\ (n \times m)$, $C\ (n \times p)$, $\det(AXBXC) \neq 0$:

$$\frac{\partial \det(AXBXC)}{\partial X} =$$
$$\det(AXBXC)\{C[AXBXC]^{-1}AXB + BXC[AXBXC]^{-1}A\}'.$$

(10) $X\ (m \times m)$, $\det(X) > 0$: $\quad \dfrac{\partial \ln \det(X)}{\partial X} = (X')^{-1}$.

(11) $X\ (m \times n)$, $A\ (m \times m)$ positive definite, $\det(X'AX) > 0$:

$$\frac{\partial \ln \det(X'AX)}{\partial X} = 2AX(X'AX)^{-1}.$$

(12) $X\ (m \times n)$, $A\ (n \times n)$ positive definite, $\det(XAX') > 0$:

$$\frac{\partial \ln \det(XAX')}{\partial X} = 2(XAX')^{-1}XA.$$

(13) $X\ (m \times n)$, $A\ (n \times p)$, $B\ (m \times m)$ positive definite, $\det(A'X'BXA) > 0$:

$$\frac{\partial \ln \det(A'X'BXA)}{\partial X} = 2BXA(A'X'BXA)^{-1}A'.$$

(14) $X\ (m \times n)$, $A\ (p \times m)$, $B\ (n \times n)$ positive definite, $\det(AXBX'A') > 0$:

$$\frac{\partial \ln \det(AXBX'A')}{\partial X} = 2A'(AXBX'A')^{-1}AXB.$$

(15) $X\ (m \times m)$, $\det(X) > 0$:

$$\frac{\partial^2 \ln \det(X)}{\partial \mathrm{vec}(X) \partial \mathrm{vec}(X)'} = -K_{mm}\left[(X')^{-1} \otimes X^{-1}\right].$$

Note: (1) – (5) and (7) – (9) are either given in or follow straightforwardly from Magnus & Neudecker (1988, Chapter 9, Sec. 10). Rules (10) – (15) follow from the chain rule and the derivative of the logarithm. In (15), the derivative of $\mathrm{vec}(X^{-1})$ from Sec. 10.6 is used in addition. (6) results from (2), a chain rule and a product rule.

10.4 Jacobian Matrices of Linear Functions

10.4.1 Linear Functions with General Matrix Arguments

(1) $X\ (m \times n):$

$$\frac{\partial \mathrm{vec}(X)}{\partial \mathrm{vec}(X)'} = I_{mn},$$

$$\frac{\partial \mathrm{vec}(X')}{\partial \mathrm{vec}(X)'} = K_{mn},$$

$$\frac{\partial \mathrm{vec}(X)}{\partial \mathrm{vec}(X')'} = K_{nm}.$$

(2) $X\ (m \times n), c \in \mathbb{R}:\quad \dfrac{\partial \mathrm{vec}(cX)}{\partial \mathrm{vec}(X)'} = cI_{mn}.$

(3) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times q):\quad \dfrac{\partial \mathrm{vec}(AXB)}{\partial \mathrm{vec}(X)'} = B' \otimes A.$

(4) $X\ (m \times n),\ A\ (p \times n),\ B\ (m \times q):\quad \dfrac{\partial \mathrm{vec}(AX'B)}{\partial \mathrm{vec}(X)'} = (B' \otimes A)K_{mn}.$

(5) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times q),\ C\ (p \times q):$

$$\frac{\partial \mathrm{vec}(AXB + C)}{\partial \mathrm{vec}(X)'} = B' \otimes A.$$

(6) $X\ (m \times n),\ A, C\ (p \times m),\ B, D\ (n \times q):$

$$\frac{\partial \mathrm{vec}(AXB \pm CXD)}{\partial \mathrm{vec}(X)'} = B' \otimes A \pm D' \otimes C.$$

(7) $X\ (m \times n),\ A, C\ (p \times n),\ B, D\ (m \times q):$

$$\frac{\partial \mathrm{vec}(AX'B \pm CX'D)}{\partial \mathrm{vec}(X)'} = (B' \otimes A \pm D' \otimes C)K_{mn}.$$

(8) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times q),\ C\ (p \times n),\ D\ (m \times q):$

$$\begin{aligned}\frac{\partial \mathrm{vec}(AXB + CX'D)}{\partial \mathrm{vec}(X)'} &= \frac{\partial \mathrm{vec}(CX'D + AXB)}{\partial \mathrm{vec}(X)'} \\ &= B' \otimes A + (D' \otimes C)K_{mn}.\end{aligned}$$

(9) $X\ (m \times n),\ A_i\ (p \times m),\ B_i\ (n \times q),\ i = 1, \ldots, r:$

$$\frac{\partial \mathrm{vec}(\sum_{i=1}^{r} A_i X B_i)}{\partial \mathrm{vec}(X)'} = \sum_{i=1}^{r} B_i' \otimes A_i.$$

(10) $X\ (m \times n),\ A_i\ (p \times n),\ B_i\ (m \times q),\ i = 1, \ldots, r$:

$$\frac{\partial \mathrm{vec}(\sum_{i=1}^{r} A_i X' B_i)}{\partial \mathrm{vec}(X)'} = \left(\sum_{i=1}^{r} B_i' \otimes A_i\right) K_{mn}.$$

Kronecker and Hadamard Products

(11) $X\ (m \times n),\ A\ (p \times q)$:

(a) $$\frac{\partial \mathrm{vec}(A) \otimes \mathrm{vec}(X)}{\partial \mathrm{vec}(X)'} = \mathrm{vec}(A) \otimes I_{mn}.$$

(b) $$\frac{\partial \mathrm{vec}(X) \otimes \mathrm{vec}(A)}{\partial \mathrm{vec}(X)'} = I_{mn} \otimes \mathrm{vec}(A).$$

(c) $$\frac{\partial \mathrm{vec}(A \otimes X)}{\partial \mathrm{vec}(X)'} = (I_q \otimes K_{np} \otimes I_m)[\mathrm{vec}(A) \otimes I_{mn}].$$

(d) $$\frac{\partial \mathrm{vec}(X \otimes A)}{\partial \mathrm{vec}(X)'} = (I_n \otimes K_{qm} \otimes I_p)[I_{mn} \otimes \mathrm{vec}(A)].$$

(12) $X\ (m \times n),\ A\ (p \times q),\ B\ (r \times m),\ C\ (n \times s)$:

$$\frac{\partial \mathrm{vec}(A \otimes BXC)}{\partial \mathrm{vec}(X)'} = (I_q \otimes K_{sp} \otimes I_r)[\mathrm{vec}(A) \otimes C' \otimes B],$$

$$\frac{\partial \mathrm{vec}(BXC \otimes A)}{\partial \mathrm{vec}(X)'} = (I_s \otimes K_{qr} \otimes I_p)[C' \otimes B \otimes \mathrm{vec}(A)].$$

(13) $X\ (m \times n),\ A\ (p \times q), B\ (r \times s)$:

$$\frac{\partial \mathrm{vec}(A \otimes X \otimes B)}{\partial \mathrm{vec}(X)'}$$
$$= (I_q \otimes K_{n\ s,p} \otimes I_{mr})[\mathrm{vec}(A) \otimes (I_n \otimes K_{sm} \otimes I_r)(I_{mn} \otimes \mathrm{vec}(B))].$$

(14) $X\ (m \times n),\ A\ (p \times q),\ B\ (r \times s),\ C\ (k \times m),\ D\ (n \times l)$:

$$\frac{\partial \mathrm{vec}(A \otimes CXD \otimes B)}{\partial \mathrm{vec}(X)'}$$
$$= (I_q \otimes K_{s\ l,p} \otimes I_{kr})[\mathrm{vec}(A) \otimes (I_l \otimes K_{sk} \otimes I_r)(D' \otimes C \otimes \mathrm{vec}(B))].$$

(15) $X\ (m \times n),\ A\ (p \times q),\ B\ (mp \times nq)$:

$$\frac{\partial \mathrm{vec}[(A \otimes X) \odot B]}{\partial \mathrm{vec}(X)'} = \mathrm{diag}(\mathrm{vec}\, B)(I_q \otimes K_{np} \otimes I_m)[\mathrm{vec}(A) \otimes I_{mn}],$$

$$\frac{\partial \mathrm{vec}[(X \otimes A) \odot B]}{\partial \mathrm{vec}(X)'} = \mathrm{diag}(\mathrm{vec}\, B)(I_n \otimes K_{qm} \otimes I_p)[I_{mn} \otimes \mathrm{vec}(A)].$$

(16) $X, A\ (m \times n)$:

$$\frac{\partial \text{vec}(A \odot X)}{\partial \text{vec}(X)'} = \frac{\partial \text{vec}(X \odot A)}{\partial \text{vec}(X)'} = \text{diag}(\text{vec}\, A).$$

(17) $X\ (m \times n),\ A\ (p \times q),\ B\ (p \times m),\ C\ (n \times q)$:

$$\frac{\partial \text{vec}(A \odot BXC)}{\partial \text{vec}(X)'} = \frac{\partial \text{vec}(BXC \odot A)}{\partial \text{vec}(X)'} = \text{diag}(\text{vec}\, A)(C' \otimes B).$$

(18) $X\ (m \times n),\ A\ (p \times q),\ B\ (p \times n),\ C\ (m \times q)$:

$$\frac{\partial \text{vec}(A \odot BX'C)}{\partial \text{vec}(X)'} = \frac{\partial \text{vec}(BX'C \odot A)}{\partial \text{vec}(X)'} = \text{diag}(\text{vec}\, A)(C' \otimes B)K_{mn}.$$

(19) $X, A, B\ (m \times n)$:

$$\frac{\partial \text{vec}(A \odot X \odot B)}{\partial \text{vec}(X)'} = \text{diag}[\text{vec}(A \odot B)].$$

(20) $X\ (m \times n),\ A, B\ (n \times m)$:

$$\frac{\partial \text{vec}(A \odot X' \odot B)}{\partial \text{vec}(X)'} = \text{diag}[\text{vec}(A \odot B)]K_{mn}.$$

(21) $X, A\ (m \times n),\ B\ (p \times q)$:

$$\frac{\partial \text{vec}[(A \odot X) \otimes B]}{\partial \text{vec}(X)'} = (I_n \otimes K_{qm} \otimes I_p)[\text{diag}(\text{vec}\, A) \otimes \text{vec}(B)],$$

$$\frac{\partial \text{vec}[B \otimes (A \odot X)]}{\partial \text{vec}(X)'} = (I_q \otimes K_{np} \otimes I_m)[\text{vec}(B) \otimes \text{diag}(\text{vec}\, A)].$$

Note: (1) – (10) follow from basic properties of the derivative and the rule $\text{vec}(AXB) = (B' \otimes A)\text{vec}(X)$ for the vec operator (see Section 7.2). (11) – (21) may be derived using $\text{vec}(A \otimes B) = (I \otimes K \otimes I)[\text{vec}(A) \otimes \text{vec}(B)]$ and $\text{vec}(A \odot B) = \text{diag}(\text{vec}\, A)\text{vec}(B)$ and other standard rules for Kronecker and Hadamard products (see Chapter 2 and Magnus & Neudecker (1988, Chapter 9, Sec. 14)).

10.4.2 Linear Functions with Symmetric Matrix Arguments

Reminder: D_m denotes a duplication matrix and D_m^+ its Moore–Penrose inverse (see Section 9.5). L_m is an elimination matrix (see Section 9.6).

(1) X $(m \times m)$ symmetric: $\dfrac{\partial \text{vech}(X)}{\partial \text{vech}(X)'} = I_{m(m+1)/2}$,

$$\frac{\partial \text{vec}(X)}{\partial \text{vech}(X)'} = D_m.$$

(2) X $(m \times m)$ symmetric, $c \in \mathbb{R}$: $\dfrac{\partial \text{vec}(cX)}{\partial \text{vech}(X)'} = cD_m$.

(3) X $(m \times m)$ symmetric, A $(m \times n)$:

$$\frac{\partial \text{vech}(A'XA)}{\partial \text{vech}(X)'} = L_n(A' \otimes A')D_m = D_n^+(A' \otimes A')D_m.$$

(4) X $(m \times m)$ symmetric, A $(n \times m), B$ $(m \times p)$:

$$\frac{\partial \text{vec}(AXB)}{\partial \text{vech}(X)'} = (B' \otimes A)D_m.$$

(5) X $(m \times m)$ symmetric, A, B $(m \times m)$:

$$\frac{\partial \text{vech}(AXA' \pm BXB')}{\partial \text{vech}(X)'} = D_m^+(A \otimes A \pm B \otimes B)D_m.$$

(6) X $(m \times m)$ symmetric, A, B $(m \times m)$:

$$\frac{\partial \text{vech}(AXB' + BXA')}{\partial \text{vech}(X)'} = D_m^+(B \otimes A + A \otimes B)D_m.$$

(7) X $(m \times m)$ symmetric, A, C $(n \times m)$, B, D $(m \times p)$:

$$\frac{\partial \text{vec}(AXB \pm CXD)}{\partial \text{vech}(X)'} = (B' \otimes A \pm D' \otimes C)D_m.$$

(8) X $(m \times m)$ symmetric, A_i $(n \times m), B_i$ $(m \times p)$, $i = 1, \ldots, r$:

$$\frac{\partial \text{vec}(\sum_{i=1}^r A_i X B_i)}{\partial \text{vech}(X)'} = \sum_{i=1}^r (B_i' \otimes A_i)D_m.$$

(9) X $(m \times m)$ symmetric, A_i $(m \times n)$, $i = 1, \ldots, r$:

$$\frac{\partial \text{vech}(\sum_{i=1}^r A_i' X A_i)}{\partial \text{vech}(X)'} = \sum_{i=1}^r L_n(A_i' \otimes A_i')D_m = \sum_{i=1}^r D_n^+(A_i' \otimes A_i')D_m.$$

Note: These results follow from basic properties of the derivative and the vec and vech operators. Notably, the rule $\text{vec}(AXB) = (B' \otimes A)\text{vec}(X)$ is useful (see Chapter 7). For proofs see also Magnus (1988, Capter 8, Sec. 8.2).

10.4.3 Linear Functions with Triangular Matrix Arguments

Reminder: L_m denotes an elimination matrix (see Section 9.6).

(1) X $(m \times m)$ lower triangular: $\dfrac{\partial \mathrm{vech}(X)}{\partial \mathrm{vech}(X)'} = I_{m(m+1)/2}$,

$$\frac{\partial \mathrm{vec}(X)}{\partial \mathrm{vech}(X)'} = L_m'.$$

(2) X $(m \times m)$ lower triangular, $c \in \mathbb{R}$: $\dfrac{\partial \mathrm{vec}(cX)}{\partial \mathrm{vech}(X)'} = cL_m'$.

(3) X $(m \times m)$ lower triangular, A $(n \times m)$, B $(m \times p)$:

$$\frac{\partial \mathrm{vec}(AXB)}{\partial \mathrm{vech}(X)'} = (B' \otimes A)L_m'.$$

(4) X $(m \times m)$ lower triangular, A, C $(n \times m)$, B, D $(m \times p)$:

$$\frac{\partial \mathrm{vec}(AXB \pm CXD)}{\partial \mathrm{vech}(X)'} = (B' \otimes A \pm D' \otimes C)L_m'.$$

(5) X $(m \times m)$ lower triangular, A_i $(n \times m)$, B_i $(m \times p)$, $i = 1, \ldots, r$:

$$\frac{\partial \mathrm{vec}(\sum_{i=1}^r A_i X B_i)}{\partial \mathrm{vech}(X)'} = \sum_{i=1}^r (B_i' \otimes A_i)L_m'.$$

(6) X, A, B $(m \times m)$ lower triangular: $\dfrac{\partial \mathrm{vech}(AXB)}{\partial \mathrm{vech}(X)'} = L_m(B' \otimes A)L_m'$.

(7) X, A, B, C, D $(m \times m)$ lower triangular:

$$\frac{\partial \mathrm{vech}(AXB \pm CXD)}{\partial \mathrm{vech}(X)'} = L_m(B' \otimes A \pm D' \otimes C)L_m'.$$

(8) X, A_i, B_i $(m \times m)$ lower triangular, $i = 1, \ldots, r$:

$$\frac{\partial \mathrm{vech}(\sum_{i=1}^r A_i X B_i)}{\partial \mathrm{vech}(X)'} = L_m \left(\sum_{i=1}^r B_i' \otimes A_i \right) L_m'.$$

Note: The results of this subsection follow from basic properties of derivatives and the vec and vech operators. In particular, the rules $\mathrm{vec}(A) = L_m'\mathrm{vech}(A)$ for lower triangular matrices A (see Section 9.14) and $\mathrm{vec}(AXB) = (B' \otimes A)\mathrm{vec}(X)$ (see Section 7.2) are useful. For proofs see also Magnus (1988, Chapter 8, Sec. 8.3).

10.4.4 Linear Functions of Vector and Matrix Valued Functions with Vector Arguments

(1) $x\ (m \times 1)$, $c\ (n \times 1)$ constant: $\quad \dfrac{\partial c}{\partial x'} = O_{n \times m}$.

(2) $x\ (m \times 1)$, $y(x)$, $z(x)\ (n \times 1)$, $c_1, c_2 \in \mathbb{R}$:

$$\frac{\partial[c_1 y(x) \pm c_2 z(x)]}{\partial x'} = c_1 \frac{\partial y(x)}{\partial x'} \pm c_2 \frac{\partial z(x)}{\partial x'}.$$

(3) $x\ (m \times 1)$, $y_i(x)\ (n \times 1)$, $c_i \in \mathbb{R}$, $i = 1, \dots, r$:

$$\frac{\partial[\sum_{i=1}^{r} c_i y_i(x)]}{\partial x'} = \sum_{i=1}^{r} c_i \frac{\partial y_i(x)}{\partial x'}.$$

(4) $x\ (m \times 1)$, $y(x)\ (n \times 1)$, $z(x)\ (p \times 1)$, $A\ (q \times n)$, $B\ (q \times p)$:

$$\frac{\partial[Ay(x) \pm Bz(x)]}{\partial x'} = A\frac{\partial y(x)}{\partial x'} \pm B\frac{\partial z(x)}{\partial x'}.$$

(5) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $c \in \mathbb{R}$: $\quad \dfrac{\partial \operatorname{vec}(cY)}{\partial x'} = c\dfrac{\partial \operatorname{vec}(Y)}{\partial x'}$.

(6) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (q \times n)$, $B\ (p \times r)$:

$$\frac{\partial \operatorname{vec}(AYB)}{\partial x'} = (B' \otimes A)\frac{\partial \operatorname{vec}(Y)}{\partial x'}.$$

(7) $x\ (m \times 1)$, $Y_i(x)\ (n_i \times p_i)$, $A_i\ (q \times n_i)$, $B_i\ (p_i \times r)$, $i = 1, \dots, s$:

$$\frac{\partial \operatorname{vec}(\sum_{i=1}^{s} A_i Y_i B_i)}{\partial x'} = \sum_{i=1}^{s} (B_i' \otimes A_i)\frac{\partial \operatorname{vec}(Y_i)}{\partial x'}.$$

(8) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (q \times n)$, $B\ (p \times r)$, $C\ (q \times r)$:

$$\frac{\partial \operatorname{vec}(AYB + C)}{\partial x'} = (B' \otimes A)\frac{\partial \operatorname{vec}(Y)}{\partial x'}.$$

(9) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (r \times s)$:

$$\frac{\partial \operatorname{vec}(A \otimes Y)}{\partial x'} = (I_s \otimes K_{pr} \otimes I_n)\left(\operatorname{vec}(A) \otimes \frac{\partial \operatorname{vec}(Y)}{\partial x'}\right),$$

$$\frac{\partial \operatorname{vec}(Y \otimes A)}{\partial x'} = (I_p \otimes K_{sn} \otimes I_r)\left(\frac{\partial \operatorname{vec}(Y)}{\partial x'} \otimes \operatorname{vec}(A)\right).$$

(10) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (r \times s)$, $B\ (k \times l)$:

$$\begin{aligned}&\frac{\partial \mathrm{vec}(A \otimes Y \otimes B)}{\partial x'}\\ &= (I_s \otimes K_{p\cdot l,r} \otimes I_{nk})\\ &\quad \times \left[\mathrm{vec}(A) \otimes (I_p \otimes K_{ln} \otimes I_k)\left(\frac{\partial \mathrm{vec}(Y)}{\partial x'} \otimes \mathrm{vec}(B)\right)\right].\end{aligned}$$

(11) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (r \times s)$, $B\ (k \times n)$, $C\ (p \times l)$:

$$\frac{\partial \mathrm{vec}(BYC \otimes A)}{\partial x'} = (I_l \otimes K_{sk} \otimes I_r)\left(\left[(C' \otimes B)\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right] \otimes \mathrm{vec}(A)\right),$$

$$\frac{\partial \mathrm{vec}(A \otimes BYC)}{\partial x'} = (I_s \otimes K_{lr} \otimes I_k)\left(\mathrm{vec}(A) \otimes \left[(C' \otimes B)\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right]\right).$$

(12) $x\ (m \times 1)$, $Y(x)\ (n \times p)$, $A\ (n \times p)$:

$$\frac{\partial \mathrm{vec}(Y \odot A)}{\partial x'} = \frac{\partial \mathrm{vec}(A \odot Y)}{\partial x'} = \mathrm{diag}(\mathrm{vec}\, A)\frac{\partial \mathrm{vec}(Y)}{\partial x'}.$$

Note: The results of this subsection follow from the linearity of the functions considered, the rules for the vec operator and basic matrix operations.

10.5 Product Rules

10.5.1 Matrix Products

(1) $X\ (m \times m)$:

(a) $$\frac{\partial \mathrm{vec}(X^2)}{\partial \mathrm{vec}(X)'} = X' \otimes I_m + I_m \otimes X.$$

(b) $$\frac{\partial \mathrm{vec}(X^i)}{\partial \mathrm{vec}(X)'} = \sum_{j=0}^{i-1}(X')^{i-1-j} \otimes X^j, \quad i = 1, 2, \ldots$$

(c) $$\frac{\partial \mathrm{vec}(X'^2)}{\partial \mathrm{vec}(X)'} = (X \otimes I_m + I_m \otimes X')K_{mm}.$$

(2) $X\ (m \times n)$:

(a) $$\frac{\partial \mathrm{vec}(X'X)}{\partial \mathrm{vec}(X)'} = (I_{n^2} + K_{nn})(I_n \otimes X').$$

(b) $$\frac{\partial \mathrm{vec}(XX')}{\partial \mathrm{vec}(X)'} = (I_{m^2} + K_{mm})(X \otimes I_m).$$

(3) X $(m \times n)$, A $(n \times m)$:

$$\frac{\partial \text{vec}(XAX)}{\partial \text{vec}(X)'} = X'A' \otimes I_m + I_n \otimes XA.$$

(4) X $(m \times n)$, A $(m \times m)$:

$$\frac{\partial \text{vec}(X'AX)}{\partial \text{vec}(X)'} = (X'A' \otimes I_n)K_{mn} + (I_n \otimes X'A).$$

(5) X $(m \times n)$, A $(m \times m)$ symmetric:

$$\frac{\partial \text{vec}(X'AX)}{\partial \text{vec}(X)'} = (I_{n^2} + K_{nn})(I_n \otimes X'A).$$

(6) X $(m \times n)$, A $(n \times n)$:

$$\frac{\partial \text{vec}(XAX')}{\partial \text{vec}(X)'} = (XA' \otimes I_m) + (I_m \otimes XA)K_{mn}.$$

(7) X $(m \times n)$, A $(n \times n)$ symmetric:

$$\frac{\partial \text{vec}(XAX')}{\partial \text{vec}(X)'} = (I_{m^2} + K_{mm})(XA \otimes I_m).$$

(8) X $(m \times n)$, A $(p \times m)$, B $(n \times m)$, C $(n \times q)$:

$$\frac{\partial \text{vec}(AXBXC)}{\partial \text{vec}(X)'} = C'X'B' \otimes A + C' \otimes AXB.$$

(9) X $(m \times n)$, A $(p \times n)$, B $(m \times m)$, C $(n \times q)$:

$$\frac{\partial \text{vec}(AX'BXC)}{\partial \text{vec}(X)'} = (C'X'B' \otimes A)K_{mn} + (C' \otimes AX'B).$$

(10) X $(m \times n)$, A $(p \times n)$, B $(m \times m)$ symmetric:

$$\frac{\partial \text{vec}(AX'BXA')}{\partial \text{vec}(X)'} = (I_{p^2} + K_{pp})(A \otimes AX'B).$$

(11) X $(m \times n)$, A $(p \times m)$, B $(n \times n)$, C $(m \times q)$:

$$\frac{\partial \text{vec}(AXBX'C)}{\partial \text{vec}(X)'} = (C'XB' \otimes A) + (C' \otimes AXB)K_{mn}.$$

(12) X $(m \times n)$, A $(p \times m)$, B $(n \times n)$ symmetric:

$$\frac{\partial \text{vec}(AXBX'A')}{\partial \text{vec}(X)'} = (I_{p^2} + K_{pp})(AXB \otimes A).$$

(13) $X\ (m \times m),\ A\ (n \times m),\ B\ (m \times p)$:

$$\frac{\partial \mathrm{vec}(AX^iB)}{\partial \mathrm{vec}(X)'} = \sum_{j=0}^{i-1} B'(X')^{i-1-j} \otimes AX^j, \quad i = 1, 2, \ldots$$

(14) $X\ (m \times n),\ A\ (p \times m),\ B\ (n \times p)$:

$$\frac{\partial \mathrm{vec}(AXB)^i}{\partial \mathrm{vec}(X)'} = \sum_{j=0}^{i-1} (B'X'A')^{i-1-j}B' \otimes (AXB)^j A, \quad i = 1, 2, \ldots$$

(15) $X, A\ (m \times m)$:

$$\frac{\partial \mathrm{vec}(A+X)^2}{\partial \mathrm{vec}(X)'} = (A' + X') \otimes I_m + I_m \otimes (A + X),$$

$$\frac{\partial \mathrm{vec}(A+X)^i}{\partial \mathrm{vec}(X)'} = \sum_{j=0}^{i-1} (A' + X')^{i-1-j} \otimes (A + X)^j, \quad i = 1, 2, \ldots$$

(16) $X, A\ (m \times n)$:

$$\frac{\partial \mathrm{vec}[(A+X)'(A+X)]}{\partial \mathrm{vec}(X)'} = (I_{n^2} + K_{nn})[I_n \otimes (A' + X')],$$

$$\frac{\partial \mathrm{vec}[(A+X)(A+X)']}{\partial \mathrm{vec}(X)'} = (I_{m^2} + K_{mm})[(A + X) \otimes I_m].$$

(17) $X\ (m \times n),\ A\ (p \times p),\ B\ (p \times m),\ C\ (n \times p)$:

$$\frac{\partial \mathrm{vec}(A+BXC)^2}{\partial \mathrm{vec}(X)'} = (A' + C'X'B')C' \otimes B + C' \otimes (A + BXC)B,$$

$$\frac{\partial \mathrm{vec}(A+BXC)^i}{\partial \mathrm{vec}(X)'} = \sum_{j=0}^{i-1} (A' + C'X'B')^{i-1-j}C' \otimes (A + BXC)^j B,$$

$i = 1, 2, \ldots$

(18) $X\ (m \times n),\ A\ (p \times q),\ B\ (p \times m),\ C\ (n \times q)$:

$$\frac{\partial \mathrm{vec}[(A+BXC)'(A+BXC)]}{\partial \mathrm{vec}(X)'} = (I_{q^2} + K_{qq})[C' \otimes (A' + C'X'B')B],$$

$$\frac{\partial \mathrm{vec}[(A+BXC)(A+BXC)']}{\partial \mathrm{vec}(X)'} = (I_{p^2} + K_{pp})[(A + BXC)C' \otimes B].$$

Note: (1) – (2) are given in Magnus & Neudecker (1988, Chapter 9, Sec. 13). The remaining results follow via the product and chain rules of differential calculus.

10.5.2 Kronecker and Hadamard Products

(1) X $(m \times n)$:

(a) $$\frac{\partial \text{vec}(X) \otimes \text{vec}(X)}{\partial \text{vec}(X)'} = I_{mn} \otimes \text{vec}(X) + \text{vec}(X) \otimes I_{mn}.$$

(b) $$\frac{\partial \text{vec}(X \otimes X)}{\partial \text{vec}(X)'} = (I_n \otimes K_{nm} \otimes I_m)[I_{mn} \otimes \text{vec}(X) + \text{vec}(X) \otimes I_{mn}].$$

(c) $$\frac{\partial \text{vec}(X \otimes X')}{\partial \text{vec}(X)'}$$
$$= (I_n \otimes K_{mm} \otimes I_n)[I_{mn} \otimes \text{vec}(X') + \text{vec}(X) \otimes K_{mn}].$$

(d) $$\frac{\partial \text{vec}(X' \otimes X)}{\partial \text{vec}(X)'}$$
$$= (I_m \otimes K_{nn} \otimes I_m)[K_{mn} \otimes \text{vec}(X) + \text{vec}(X') \otimes I_{mn}].$$

(e) $$\frac{\partial \text{vec}(X' \otimes X')}{\partial \text{vec}(X)'}$$
$$= (I_m \otimes K_{mn} \otimes I_n)[K_{mn} \otimes \text{vec}(X') + \text{vec}(X') \otimes K_{mn}].$$

(2) X $(m \times n)$, A $(p \times m)$, B $(n \times q)$, C $(r \times m)$, D $(n \times s)$:

$$\frac{\partial \text{vec}(AXB \otimes CXD)}{\partial \text{vec}(X)'}$$
$$= (I_q \otimes K_{sp} \otimes I_r)[B' \otimes A \otimes \text{vec}(CXD) + \text{vec}(AXB) \otimes D' \otimes C].$$

(3) X $(m \times n)$, A $(p \times m)$, B $(n \times q)$, C $(r \times n)$, D $(m \times s)$:

$$\frac{\partial \text{vec}(AXB \otimes CX'D)}{\partial \text{vec}(X)'}$$
$$= (I_q \otimes K_{sp} \otimes I_r)$$
$$\times [B' \otimes A \otimes \text{vec}(CX'D) + \text{vec}(AXB) \otimes (D' \otimes C)K_{mn}].$$

(4) X $(m \times n)$, A $(p \times n)$, B $(m \times q)$, C $(r \times m)$, D $(n \times s)$:

$$\frac{\partial \text{vec}(AX'B \otimes CXD)}{\partial \text{vec}(X)'}$$
$$= (I_q \otimes K_{sp} \otimes I_r)$$
$$\times [(B' \otimes A)K_{mn} \otimes \text{vec}(CXD) + \text{vec}(AX'B) \otimes D' \otimes C].$$

(5) X $(m \times n)$, A $(p \times n)$, B $(m \times q)$, C $(r \times n)$, D $(m \times s)$:

$$\frac{\partial \text{vec}(AX'B \otimes CX'D)}{\partial \text{vec}(X)'}$$

$$= (I_q \otimes K_{sp} \otimes I_r)[(B' \otimes A)K_{mn} \otimes \text{vec}(CX'D)+$$

$$+ \ \mathrm{vec}(AX'B) \otimes (D' \otimes C)K_{mn}].$$

(6) $X\ (m \times n)$:

(a) $$\frac{\partial \mathrm{vec}(X \odot X)}{\partial \mathrm{vec}(X)'} = 2\ \mathrm{diag}(\mathrm{vec}\, X).$$

(b) $$\frac{\partial \mathrm{vec}(X \odot X \odot X)}{\partial \mathrm{vec}(X)'} = 3\ \mathrm{diag}[\mathrm{vec}(X \odot X)].$$

(c) $$\frac{\partial \mathrm{vec}(X' \odot X')}{\partial \mathrm{vec}(X)'} = 2\ \mathrm{diag}(\mathrm{vec}\, X')K_{mn}.$$

(7) $X\ (m \times m)$:

$$\frac{\partial \mathrm{vec}(X \odot X')}{\partial \mathrm{vec}(X)'} = \frac{\partial \mathrm{vec}(X' \odot X)}{\partial \mathrm{vec}(X)'} = \mathrm{diag}(\mathrm{vec}\, X)K_{mn} + \mathrm{diag}(\mathrm{vec}\, X').$$

(8) $X\ (m \times n),\ A, C\ (p \times m),\ B, D\ (n \times q)$:

$$\frac{\partial \mathrm{vec}(AXB \odot CXD)}{\partial \mathrm{vec}(X)'}$$
$$= \mathrm{diag}[\mathrm{vec}(AXB)](D' \otimes C) + \mathrm{diag}[\mathrm{vec}(CXD)](B' \otimes A).$$

(9) $X\ (m \times n),\ A\ (p \times n),\ B\ (m \times q),\ C\ (p \times m),\ D\ (n \times q)$:

$$\begin{aligned} \frac{\partial \mathrm{vec}(AX'B \odot CXD)}{\partial \mathrm{vec}(X)'} &= \frac{\partial \mathrm{vec}(CXD \odot AX'B)}{\partial \mathrm{vec}(X)'} \\ &= \mathrm{diag}[\mathrm{vec}(AX'B)](D' \otimes C) \\ &\quad + \mathrm{diag}[\mathrm{vec}(CXD)](B' \otimes A)K_{mn}. \end{aligned}$$

(10) $X\ (m \times n),\ A, C\ (p \times n),\ B, D\ (m \times q)$:

$$\begin{aligned} \frac{\partial \mathrm{vec}(AX'B \odot CX'D)}{\partial \mathrm{vec}(X)'} &= \{\mathrm{diag}[\mathrm{vec}(AX'B)](D' \otimes C) \\ &\quad + \mathrm{diag}[\mathrm{vec}(CX'D)](B' \otimes A)\}K_{mn}. \end{aligned}$$

Note: The results of this subsection follow from the basic product rule for differentiation and the rules for Kronecker and Hadamard products. See also Magnus & Neudecker (1988, Chapter 9, Sec. 14).

10.5.3 Functions with Symmetric Matrix Arguments

Reminder: D_m denotes a duplication matrix and D_m^+ its Moore–Penrose inverse (see Section 9.5). L_m is an elimination matrix (see Section 9.6).

(1) $X\ (m \times m)$ symmetric:

(a) $\dfrac{\partial \text{vec}(X^2)}{\partial \text{vech}(X)'} = (X \otimes I_m + I_m \otimes X)D_m.$

(b) $\dfrac{\partial \text{vec}(X^i)}{\partial \text{vech}(X)'} = \sum_{j=0}^{i-1}(X^{i-1-j} \otimes X^j)D_m, \quad i = 1, 2, \ldots$

(c) $\dfrac{\partial \text{vech}(X^i)}{\partial \text{vech}(X)'} = D_m^+ \left(\sum_{j=0}^{i-1} X^{i-1-j} \otimes X^j\right) D_m, \quad i = 1, 2, \ldots$

(2) X, A $(m \times m)$ symmetric:

$$\begin{aligned}\frac{\partial \text{vech}(XAX)}{\partial \text{vech}(X)'} &= D_m^+(XA \otimes I_m + I_m \otimes XA)D_m \\ &= L_m(XA \otimes I_m + I_m \otimes XA)D_m.\end{aligned}$$

(3) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times m)$, C $(m \times q)$:

$$\frac{\partial \text{vec}(AXBXC)}{\partial \text{vech}(X)'} = (C'XB' \otimes A + C' \otimes AXB)D_m.$$

(4) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times q)$:

$$\frac{\partial \text{vec}(AX^iB)}{\partial \text{vech}(X)'} = \sum_{j=0}^{i-1}(B'X^{i-1-j} \otimes AX^j)D_m, \quad i = 1, 2, \ldots$$

(5) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times p)$:

$$\frac{\partial \text{vec}(AXB)^i}{\partial \text{vech}(X)'} = \left(\sum_{j=0}^{i-1}(B'XA')^{i-1-j}B' \otimes (AXB)^jA\right) D_m,$$

$i = 1, 2, \ldots$

(6) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times p)$, C $(p \times p)$:

$$\begin{aligned}&\frac{\partial \text{vec}(AXB + C)^i}{\partial \text{vech}(X)'} \\ &\qquad = \left(\sum_{j=0}^{i-1}(B'XA' + C')^{i-1-j}B' \otimes (AXB + C)^jA\right) D_m,\end{aligned}$$

$i = 1, 2, \ldots$

(7) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times q)$, C $(p \times q)$:

$$\frac{\partial \text{vec}[(AXB + C)'(AXB + C)]}{\partial \text{vech}(X)'} = (I_{q^2}+K_{qq})[B'\otimes(B'XA'+C')A]D_m.$$

(8) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times q)$, C $(p \times q)$:

$$\frac{\partial \text{vec}[(AXB+C)(AXB+C)']}{\partial \text{vech}(X)'} = (I_{p^2} + K_{pp})[(AXB+C)B' \otimes A]D_m.$$

(9) X $(m \times m)$ symmetric:

$$\frac{\partial \text{vec}(X \otimes X)}{\partial \text{vech}(X)'} = (I_m \otimes K_{mm} \otimes I_m)[I_{m^2} \otimes \text{vec}(X) + \text{vec}(X) \otimes I_{m^2}]D_m.$$

(10) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times q)$, C $(r \times m)$, D $(m \times s)$:

$$\frac{\partial \text{vec}(AXB \otimes CXD)}{\partial \text{vech}(X)'}$$
$$= (I_q \otimes K_{sp} \otimes I_r)[B' \otimes A \otimes \text{vec}(CXD) + \text{vec}(AXB) \otimes D' \otimes C]D_m.$$

(11) X $(m \times m)$ symmetric:

(a) $$\frac{\partial \text{vec}(X \odot X)}{\partial \text{vech}(X)'} = 2 \text{ diag}(\text{vec}\, X)D_m.$$

(b) $$\frac{\partial \text{vec}(X \odot X \odot X)}{\partial \text{vech}(X)'} = 3 \text{ diag}[\text{vec}(X \odot X)]D_m.$$

(12) X $(m \times m)$ symmetric, A $(p \times m)$, B $(m \times q)$, C $(p \times m)$, D $(m \times q)$:

$$\frac{\partial \text{vec}(AXB \odot CXD)}{\partial \text{vech}(X)'} = \{\text{diag}[\text{vec}(AXB)](D' \otimes C) + \text{diag}[\text{vec}(CXD)](B' \otimes A)\}D_m.$$

Note: These results follow from those of the previous subsection and the chain rule for matrix differentiation. See also Magnus (1988, Chapter 8, Sec. 8.2).

10.5.4 Functions with Lower Triangular Matrix Arguments

Reminder: D_m^+ is the Moore–Penrose inverse of the duplication matrix D_m (see Section 9.5) and L_m denotes an elimination matrix (see Section 9.6).

(1) X $(m \times m)$ lower triangular:

$$\frac{\partial \text{vech}(X^i)}{\partial \text{vech}(X)'} = L_m \left(\sum_{j=0}^{i-1} (X')^{i-1-j} \otimes X^j \right) L_m', \quad i = 1, 2, \ldots$$

(2) X, A $(m \times m)$ lower triangular:

$$\frac{\partial \text{vech}(XAX)}{\partial \text{vech}(X)'} = L_m(X'A' \otimes I_m + I_m \otimes XA)L_m'.$$

(3) X $(m \times m)$ lower triangular:

$$\frac{\partial \text{vech}(X'X)}{\partial \text{vech}(X)'} = 2D_m^+(I_m \otimes X')L_m',$$

$$\frac{\partial \text{vech}(XX')}{\partial \text{vech}(X)'} = 2D_m^+(X \otimes I_m)L_m'.$$

Note: For proofs see Magnus (1988, Chapter 8, Sec. 8.3 - 8.4).

10.5.5 Products of Matrix Valued Functions with Vector Arguments

(1) x $(m \times 1)$, $Y(x)$ $(n \times n)$:

$$\frac{\partial \text{vec}(Y^i)}{\partial x'} = \sum_{j=0}^{i-1}[(Y')^{i-1-j} \otimes Y^j]\frac{\partial \text{vec}(Y)}{\partial x'}.$$

(2) x $(m \times 1)$, $Y(x)$ $(n \times p)$:

$$\frac{\partial \text{vec}(Y'Y)}{\partial x'} = (I_{p^2} + K_{pp})(I_p \otimes Y')\frac{\partial \text{vec}(Y)}{\partial x'},$$

$$\frac{\partial \text{vec}(YY')}{\partial x'} = (I_{n^2} + K_{nn})(Y \otimes I_n)\frac{\partial \text{vec}(Y)}{\partial x'}.$$

(3) x $(m \times 1)$, $Y(x)$ $(n \times p)$, $Z(x)$ $(p \times q)$:

$$\frac{\partial \text{vec}(YZ)}{\partial x'} = (I_q \otimes Y)\frac{\partial \text{vec}(Z)}{\partial x'} + (Z' \otimes I_n)\frac{\partial \text{vec}(Y)}{\partial x'}.$$

(4) x $(m \times 1)$, $Y(x)$ $(n \times p)$, $Z(x)$ $(q \times r)$, A $(s \times n)$, B $(p \times q)$, C $(r \times k)$:

$$\frac{\partial \text{vec}(AYBZC)}{\partial x'} = (C' \otimes AYB)\frac{\partial \text{vec}(Z)}{\partial x'} + (C'Z'B' \otimes A)\frac{\partial \text{vec}(Y)}{\partial x'}.$$

(5) x $(m \times 1)$, $Y(x)$ $(n \times p)$, A $(q \times n)$, B $(p \times q)$, C $(q \times q)$:

$$\frac{\partial \text{vec}(AYB + C)^i}{\partial x'}$$
$$= \left(\sum_{j=0}^{i-1}(B'Y'A' + C')^{i-1-j}B' \otimes (AYB + C)^j A\right)\frac{\partial \text{vec}(Y)}{\partial x'},$$

$i = 1, 2, \ldots$

(6) x $(m \times 1)$, $Y(x)$ $(n \times p)$, A $(q \times n)$, B $(p \times r)$, C $(q \times r)$:

(a) $$\frac{\partial \mathrm{vec}[(AYB+C)'(AYB+C)]}{\partial \boldsymbol{x}'} = (I_{r^2}+K_{rr})[B'\otimes(B'Y'A'+C')A]\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}.$$

(b) $$\frac{\partial \mathrm{vec}[(AYB+C)(AYB+C)']}{\partial \boldsymbol{x}'} = (I_{q^2}+K_{qq})[(AYB+C)B'\otimes A]\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}.$$

(7) $\boldsymbol{x}$ $(m\times 1)$, $Y(\boldsymbol{x})$ $(n\times p)$, $Z(\boldsymbol{x})$ $(q\times r)$:

(a) $$\frac{\partial[\mathrm{vec}(Y)\otimes \mathrm{vec}(Z)]}{\partial \boldsymbol{x}'} = \frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}\otimes \mathrm{vec}(Z)+\mathrm{vec}(Y)\otimes\frac{\partial \mathrm{vec}(Z)}{\partial \boldsymbol{x}'}.$$

(b) $$\frac{\partial \mathrm{vec}(Y\otimes Z)}{\partial \boldsymbol{x}'} = (I_p\otimes K_{rn}\otimes I_q)\left[\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}\otimes \mathrm{vec}(Z)+\mathrm{vec}(Y)\otimes\frac{\partial \mathrm{vec}(Z)}{\partial \boldsymbol{x}'}\right].$$

(8) $\boldsymbol{x}$ $(m\times 1)$, $Y(\boldsymbol{x})$ $(n\times p)$, $Z(\boldsymbol{x})$ $(q\times r)$, A $(s\times n)$, B $(p\times k)$, C $(l\times q)$, D $(r\times h)$:

$$\begin{aligned}\frac{\partial \mathrm{vec}(AYB\otimes CZD)}{\partial \boldsymbol{x}'} &= (I_k\otimes K_{hs}\otimes I_l)\\ &\quad\times\left[(B'\otimes A)\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}\otimes \mathrm{vec}(CZD)\right.\\ &\quad\left.+\ \mathrm{vec}(AYB)\otimes(D'\otimes C)\frac{\partial \mathrm{vec}(Z)}{\partial \boldsymbol{x}'}\right].\end{aligned}$$

(9) $\boldsymbol{x}$ $(m\times 1)$, $Y(\boldsymbol{x}), Z(\boldsymbol{x})$ $(n\times p)$:

$$\begin{aligned}\frac{\partial \mathrm{vec}(Y\odot Z)}{\partial \boldsymbol{x}'} &= \frac{\partial \mathrm{vec}(Z\odot Y)}{\partial \boldsymbol{x}'}\\ &= \mathrm{diag}(\mathrm{vec}\, Z)\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}+\mathrm{diag}(\mathrm{vec}\, Y)\frac{\partial \mathrm{vec}(Z)}{\partial \boldsymbol{x}'}.\end{aligned}$$

(10) $\boldsymbol{x}$ $(m\times 1)$, $Y(\boldsymbol{x})$ $(n\times p)$, $Z(\boldsymbol{x})$ $(q\times r)$, A $(s\times n)$, B $(p\times k)$, C $(s\times q)$, D $(r\times k)$:

$$\begin{aligned}\frac{\partial \mathrm{vec}(AYB\odot CZD)}{\partial \boldsymbol{x}'} &= \frac{\partial \mathrm{vec}(CZD\odot AYB)}{\partial \boldsymbol{x}'}\\ &= \mathrm{diag}[\mathrm{vec}(CZD)](B'\otimes A)\frac{\partial \mathrm{vec}(Y)}{\partial \boldsymbol{x}'}\\ &\quad+\mathrm{diag}[\mathrm{vec}(AYB)](D'\otimes C)\frac{\partial \mathrm{vec}(Z)}{\partial \boldsymbol{x}'}.\end{aligned}$$

Note: The results of this subsection follow from the chain rule and the rules for basic matrix operations. See also the results of the previous subsections, in particular Sections 10.5.1 and 10.5.2.

10.6 Jacobian Matrices of Functions Involving Inverse Matrices

(1) X $(m \times m)$ nonsingular:

$$\frac{\partial \mathrm{vec}(X^{-1})}{\partial \mathrm{vec}(X)'} = -X'^{-1} \otimes X^{-1},$$

$$\frac{\partial \mathrm{vec}(X'^{-1})}{\partial \mathrm{vec}(X)'} = -(X^{-1} \otimes X'^{-1})K_{mn}.$$

(2) X $(m \times m)$ symmetric nonsingular:

$$\frac{\partial \mathrm{vech}(X^{-1})}{\partial \mathrm{vech}(X)'} = -D_m^+(X^{-1} \otimes X^{-1})D_m,$$

where D_m denotes a duplication matrix and D_m^+ its Moore–Penrose inverse (see Section 9.5).

Note: For proofs of these results see Magnus & Neudecker (1988, Chapter 9, Sec. 13) and Magnus (1988, Chapter 8, Sec. 8.2).

10.6.1 Matrix Products

(1) X $(m \times m)$ nonsingular, A $(n \times m)$, B $(m \times p)$:

$$\frac{\partial \mathrm{vec}(AX^{-1}B)}{\partial \mathrm{vec}(X)'} = -B'X'^{-1} \otimes AX^{-1}.$$

(2) X $(m \times m)$ nonsingular :

$$\frac{\partial \mathrm{vec}[(X^{-1})^i]}{\partial \mathrm{vec}(X)'} = -\sum_{j=0}^{i-1}(X')^{j-i} \otimes X^{-j-1}, \quad i = 1, 2, \ldots$$

(3) X $(m \times n)$, $\mathrm{rk}(X) = n$:

$$\frac{\partial \mathrm{vec}[(X'X)^{-1}]}{\partial \mathrm{vec}(X)'} = -(I_{n^2} + K_{nn})[(X'X)^{-1} \otimes (X'X)^{-1}X'].$$

(4) X $(m \times n)$, $\mathrm{rk}(X) = m$:

$$\frac{\partial \mathrm{vec}[(XX')^{-1}]}{\partial \mathrm{vec}(X)'} = -(I_{m^2} + K_{mm})[(XX')^{-1}X \otimes (XX')^{-1}].$$

(5) X $(m \times m)$ nonsingular, A $(n \times m)$, B $(m \times m)$, C $(m \times p)$:

(a) $\dfrac{\partial \text{vec}(AX^{-1}BXC)}{\partial \text{vec}(X)'} = C' \otimes AX^{-1}B - C'X'B'X'^{-1} \otimes AX^{-1}.$

(b) $\dfrac{\partial \text{vec}(AXBX^{-1}C)}{\partial \text{vec}(X)'} = C'X'^{-1}B' \otimes A - C'X'^{-1} \otimes AXBX^{-1}.$

(c) $\dfrac{\partial \text{vec}(AX^{-1}BX^{-1}C)}{\partial \text{vec}(X)'}$

$$= -C'X'^{-1}B'X'^{-1} \otimes AX^{-1} - C'X'^{-1} \otimes AX^{-1}BX^{-1}.$$

Note: The results of this subsection follow from the chain rule and the results of the previous sections.

10.6.2 Kronecker and Hadamard Products

(1) X $(m \times m)$ nonsingular, A $(p \times q)$:

$$\frac{\partial \text{vec}(A \otimes X^{-1})}{\partial \text{vec}(X)'} = -(I_q \otimes K_{mp} \otimes I_m)[\text{vec}(A) \otimes X'^{-1} \otimes X^{-1}],$$

$$\frac{\partial \text{vec}(X^{-1} \otimes A)}{\partial \text{vec}(X)'} = -(I_m \otimes K_{qm} \otimes I_p)[X'^{-1} \otimes X^{-1} \otimes \text{vec}(A)].$$

(2) X $(m \times m)$ nonsingular:

(a) $\dfrac{\partial \text{vec}(X^{-1}) \otimes \text{vec}(X)}{\partial \text{vec}(X)'} = \text{vec}(X^{-1}) \otimes I_{m^2} - X'^{-1} \otimes X^{-1} \otimes \text{vec}(X).$

(b) $\dfrac{\partial \text{vec}(X) \otimes \text{vec}(X^{-1})}{\partial \text{vec}(X)'} = I_{m^2} \otimes \text{vec}(X^{-1}) - \text{vec}(X) \otimes X'^{-1} \otimes X^{-1}.$

(c) $\dfrac{\partial \text{vec}(X^{-1}) \otimes \text{vec}(X^{-1})}{\partial \text{vec}(X)'}$

$$= -[X'^{-1} \otimes X^{-1} \otimes \text{vec}(X^{-1}) + \text{vec}(X^{-1}) \otimes X'^{-1} \otimes X^{-1}].$$

(d) $\dfrac{\partial \text{vec}(X^{-1} \otimes X)}{\partial \text{vec}(X)'}$

$$= (I_m \otimes K_{mm} \otimes I_m)[\text{vec}(X^{-1}) \otimes I_{m^2} - X'^{-1} \otimes X^{-1} \otimes \text{vec}(X)].$$

(e) $\dfrac{\partial \text{vec}(X \otimes X^{-1})}{\partial \text{vec}(X)'}$

$$= (I_m \otimes K_{mm} \otimes I_m)$$
$$\times [I_{m^2} \otimes \text{vec}(X^{-1}) - \text{vec}(X) \otimes X'^{-1} \otimes X^{-1}].$$

(f) $\dfrac{\partial \text{vec}(X^{-1} \otimes X^{-1})}{\partial \text{vec}(X)'}$

$= -(I_m \otimes K_{mm} \otimes I_m)$

$\times [X'^{-1} \otimes X^{-1} \otimes \text{vec}(X^{-1}) + \text{vec}(X^{-1}) \otimes X'^{-1} \otimes X^{-1}].$

(3) X $(m \times m)$ nonsingular, A $(m \times m)$:

$$\frac{\partial \text{vec}(A \bullet X^{-1})}{\partial \text{vec}(X)'} = \frac{\partial \text{vec}(X^{-1} \bullet A)}{\partial \text{vec}(X)'} = -\text{diag}(\text{vec}\, A)(X'^{-1} \otimes X^{-1}).$$

(4) X $(m \times m)$ nonsingular:

(a) $\dfrac{\partial \text{vec}(X^{-1} \odot X)}{\partial \text{vec}(X)'} = \dfrac{\partial \text{vec}(X \odot X^{-1})}{\partial \text{vec}(X)'}$

$= \text{diag}(\text{vec}\, X^{-1}) - \text{diag}(\text{vec}\, X)(X'^{-1} \otimes X^{-1}).$

(b) $\dfrac{\partial \text{vec}(X^{-1} \bullet X^{-1})}{\partial \text{vec}(X)'} = -2\ \text{diag}(\text{vec}\, X^{-1})(X'^{-1} \otimes X^{-1}).$

Note: The results of this subsection follow from the chain rule for derivatives and the results of the previous sections.

10.6.3 Matrix Valued Functions with Vector Arguments

(1) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular:

$$\frac{\partial \text{vec}(Y^{-1})}{\partial x'} = -(Y'^{-1} \otimes Y^{-1}) \frac{\partial \text{vec}(Y)}{\partial x'}.$$

(2) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular, $Z(x)$ $(n \times p)$:

$$\frac{\partial \text{vec}(Y^{-1}Z)}{\partial x'} = (I_p \otimes Y^{-1}) \frac{\partial \text{vec}(Z)}{\partial x'} - (Z'Y'^{-1} \otimes Y^{-1}) \frac{\partial \text{vec}(Y)}{\partial x'}.$$

(3) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular, $Z(x)$ $(p \times n)$:

$$\frac{\partial \text{vec}(ZY^{-1})}{\partial x'} = (Y'^{-1} \otimes I_p) \frac{\partial \text{vec}(Z)}{\partial x'} - (Y'^{-1} \otimes ZY^{-1}) \frac{\partial \text{vec}(Y)}{\partial x'}.$$

(4) x $(m \times 1)$, $Y(x)$, $Z(x)$ $(n \times n)$ nonsingular:

$\dfrac{\partial \text{vec}(Y^{-1}Z^{-1})}{\partial x'}$

$= -(Z'^{-1}Y'^{-1} \otimes Y^{-1}) \dfrac{\partial \text{vec}(Y)}{\partial x'} - (Z'^{-1} \otimes Y^{-1}Z^{-1}) \dfrac{\partial \text{vec}(Z)}{\partial x'}.$

(5) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular, $Z(x)$ $(p \times p)$ nonsingular, $U(x)$ $(q \times n)$, $V(x)$ $(n \times p)$, $W(x)$ $(p \times r)$:

$$\begin{aligned}
\frac{\partial \mathrm{vec}(UY^{-1}VZ^{-1}W)}{\partial x'} &= (W'Z'^{-1}V'Y'^{-1} \otimes I_q)\frac{\partial \mathrm{vec}(U)}{\partial x'} \\
&\quad - (W'Z'^{-1}V'Y'^{-1} \otimes UY^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'} \\
&\quad + (W'Z'^{-1} \otimes UY^{-1})\frac{\partial \mathrm{vec}(V)}{\partial x'} \\
&\quad - (W'Z'^{-1} \otimes UY^{-1}VZ^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'} \\
&\quad + (I_r \otimes UY^{-1}VZ^{-1})\frac{\partial \mathrm{vec}(W)}{\partial x'}.
\end{aligned}$$

(6) x $(m \times 1)$, $Y(x)$ $(n \times q)$, $Z(x)$ $(p \times p)$ nonsingular:

$$\begin{aligned}
\frac{\partial[\mathrm{vec}(Y) \otimes \mathrm{vec}(Z^{-1})]}{\partial x'} &= \frac{\partial \mathrm{vec}(Y)}{\partial x'} \otimes \mathrm{vec}(Z^{-1}) \\
&\quad - \mathrm{vec}(Y) \otimes (Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}.
\end{aligned}$$

(7) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular, $Z(x)$ $(p \times q)$:

$$\begin{aligned}
\frac{\partial[\mathrm{vec}(Y^{-1}) \otimes \mathrm{vec}(Z)]}{\partial x'} &= \mathrm{vec}(Y^{-1}) \otimes \frac{\partial \mathrm{vec}(Z)}{\partial x'} \\
&\quad - \left[(Y'^{-1} \otimes Y^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right] \otimes \mathrm{vec}(Z).
\end{aligned}$$

(8) x $(m \times 1)$, $Y(x)$ $(n \times n)$ nonsingular, $Z(x)$ $(p \times p)$ nonsingular:

$$\begin{aligned}
\frac{\partial[\mathrm{vec}(Y^{-1}) \otimes \mathrm{vec}(Z^{-1})]}{\partial x'} &= -\,\mathrm{vec}(Y^{-1}) \otimes \left[(Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}\right] \\
&\quad - \left[(Y'^{-1} \otimes Y^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right] \otimes \mathrm{vec}(Z^{-1}).
\end{aligned}$$

(9) x $(m \times 1)$, $Y(x)$ $(n \times q)$, $Z(x)$ $(p \times p)$ nonsingular:

$$\begin{aligned}
\frac{\partial \mathrm{vec}(Y \otimes Z^{-1})}{\partial x'} &= (I_q \otimes K_{pn} \otimes I_p) \\
&\quad \times \left[\frac{\partial \mathrm{vec}(Y)}{\partial x'} \otimes \mathrm{vec}(Z^{-1})\right. \\
&\quad \left. - \mathrm{vec}(Y) \otimes \left((Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}\right)\right].
\end{aligned}$$

(10) $x\,(m \times 1)$, $Y(x)\,(n \times n)$ nonsingular, $Z(x)\,(p \times q)$:

$$\begin{aligned}\frac{\partial \mathrm{vec}(Y^{-1} \otimes Z)}{\partial x'} &= (I_n \otimes K_{qn} \otimes I_p) \\ &\quad \times \left[\mathrm{vec}(Y^{-1}) \otimes \frac{\partial \mathrm{vec}(Z)}{\partial x'} \right. \\ &\quad \left. - \left((Y'^{-1} \otimes Y^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right) \otimes \mathrm{vec}(Z)\right].\end{aligned}$$

(11) $x\,(m \times 1)$, $Y(x)\,(n \times n)$ nonsingular, $Z(x)\,(p \times p)$ nonsingular:

$$\begin{aligned}\frac{\partial \mathrm{vec}(Y^{-1} \otimes Z^{-1})}{\partial x'} &= -(I_n \otimes K_{pn} \otimes I_p) \\ &\quad \times \left[\mathrm{vec}(Y^{-1}) \otimes \left((Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}\right) \right. \\ &\quad \left. + \left((Y'^{-1} \otimes Y^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'}\right) \otimes \mathrm{vec}(Z^{-1})\right].\end{aligned}$$

(12) $x\,(m \times 1)$, $Y(x)\,(n \times n)$, $Z(x)\,(n \times n)$ nonsingular:

$$\begin{aligned}\frac{\partial \mathrm{vec}(Y \odot Z^{-1})}{\partial x'} &= \frac{\partial \mathrm{vec}(Z^{-1} \odot Y)}{\partial x'} \\ &= \mathrm{diag}(\mathrm{vec}\, Z^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'} \\ &\quad - \mathrm{diag}(\mathrm{vec}\, Y)(Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}.\end{aligned}$$

(13) $x\,(m \times 1)$, $Y(x)\,(n \times n)$ nonsingular, $Z(x)\,(n \times n)$ nonsingular:

$$\begin{aligned}\frac{\partial \mathrm{vec}(Y^{-1} \odot Z^{-1})}{\partial x'} &= -\left[\mathrm{diag}(\mathrm{vec}\, Z^{-1})(Y'^{-1} \otimes Y^{-1})\frac{\partial \mathrm{vec}(Y)}{\partial x'} \right. \\ &\quad \left. + \mathrm{diag}(\mathrm{vec}\, Y^{-1})(Z'^{-1} \otimes Z^{-1})\frac{\partial \mathrm{vec}(Z)}{\partial x'}\right].\end{aligned}$$

Note: The results of this subsection follow from the chain rule for derivatives and the rules of the previous sections.

10.7 Chain Rules and Miscellaneous Jacobian Matrices

Reminder: D_m denotes a duplication matrix and D_m^+ its Moore-Penrose inverse (see Section 9.5). L_m is an elimination matrix (see Section 9.6).

(1) $x\ (m \times 1),\ y(x)\ (n \times 1),\ z(y)\ (p \times 1):$

$$\frac{\partial z(y(x))}{\partial x'} = \frac{\partial z(y)}{\partial y'}\frac{\partial y(x)}{\partial x'}.$$

(2) $X\ (m \times n),\ Y(X)\ (p \times q),\ Z(Y)\ (r \times s):$

$$\frac{\partial \text{vec}\, Z(Y(X))}{\partial \text{vec}(X)'} = \frac{\partial \text{vec}\, Z(Y)}{\partial \text{vec}(Y)'}\frac{\partial \text{vec}\, Y(X)}{\partial \text{vec}(X)'}.$$

(3) $X\ (m \times n)$, $\mathcal{S}$ open subset of $\mathbb{R}^{m\times n}$: $\quad \text{rk}(X) = r$ for all $X \in \mathcal{S}$
$\Rightarrow \quad X^+$ is a differentiable function of X on $\mathcal{S}$ and

$$\frac{\partial \text{vec}(X^+)}{\partial \text{vec}(X)'} = -X^{+\prime} \otimes X^+ + [(I_m - XX^+) \otimes X^+X^{+\prime} + X^{+\prime}X^+ \otimes (I_n - X^+X)]K_{mn}.$$

(4) $X\ (m \times n),\ Y(X)\ (p \times q)$, $\mathcal{S}$ open subset of $\mathbb{R}^{m\times n}$:
$\text{rk}[Y(X)] = r$ for all $X \in \mathcal{S} \quad \Rightarrow \quad Y(X)^+$ is differentiable on $\mathcal{S}$ and

$$\frac{\partial \text{vec}\, Y(X)^+}{\partial \text{vec}(X)'} = \Big(-Y^{+\prime} \otimes Y^+ + [(I_p - YY^+) \otimes Y^+Y^{+\prime} + Y^{+\prime}Y^+ \otimes (I_q - Y^+Y)]K_{pq}\Big)\frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'}.$$

(5) $X\ (m \times m)$ nonsingular:

$$\frac{\partial \text{vec}(X^{adj})}{\partial \text{vec}(X)'} = \det(X)[(\text{vec}\, X^{-1})(\text{vec}\, X'^{-1})' - X'^{-1} \otimes X^{-1}].$$

(6) $X\ (m \times m)$ symmetric nonsingular:

$$\frac{\partial \text{vech}(X^{adj})}{\partial \text{vech}(X)'} = \det(X)D_m^+[(\text{vec}\, X^{-1})(\text{vec}\, X^{-1})' - X^{-1} \otimes X^{-1}]D_m.$$

(7) $X\ (m \times m)$ lower triangular, nonsingular:

$$\frac{\partial \text{vech}(X^{adj})}{\partial \text{vech}(X)'} = \det(X)L_m[(\text{vec}\, X^{-1})(\text{vec}\, X'^{-1})' - X'^{-1} \otimes X^{-1}]L_m'.$$

(8) $X\ (m \times n),\ Y(X)\ (p \times p)$ nonsingular:

$$\frac{\partial \text{vec}[Y(X)^{adj}]}{\partial \text{vec}\,(X)'} = \det(Y)[(\text{vec}\, Y^{-1})(\text{vec}\, Y'^{-1})' - Y'^{-1} \otimes Y^{-1}]\frac{\partial \text{vec}(Y)}{\partial \text{vec}(X)'}.$$

(9) X $(m \times m)$ positive definite, $Y(X) = [y_{ij}]$ $(m \times m)$ lower triangular with $y_{ii} \geq 0$, $i = 1, \ldots, m$, such that $X = YY'$, that is, Y is obtained by a Choleski decomposition of X :

$$\begin{aligned}\frac{\partial \text{vech}(Y)}{\partial \text{vech}(X)'} &= \{L_m[(I_m \otimes Y)K_{mm} + (Y \otimes I_m)]L_m'\}^{-1} \\ &= \{L_m(I_{m^2} + K_{mm})(Y \otimes I_m)L_m'\}^{-1} \\ &= \tfrac{1}{2}[D_m^+(Y \otimes I_m)L_m']^{-1}.\end{aligned}$$

Note: (1) and (2) are just chain rules. (3) is given in Magnus & Neudecker (1988, Chapter 8, Sec. 5). (4) follows from (3) and a chain rule. (5) is established in Magnus & Neudecker (1988, Chapter 8, Sec. 6). (6), (7) and (8) follow from (5) and a chain rule (see also Magnus (1988, Chapter 8, Sec. 8.2, 8.3)). (9) is, e.g., given in Lütkepohl (1991, Appendix A, Sec. A.13).

10.8 Jacobian Determinants

10.8.1 Linear Transformations

(1) X $(m \times n)$:

(a) $\det\left(\dfrac{\partial \text{vec}(X)}{\partial \text{vec}(X)'}\right) = 1.$

(b) $\det\left(\dfrac{\partial \text{vec}(X')}{\partial \text{vec}(X)'}\right) = (-1)^{mn(m-1)(n-1)/4}.$

(c) $\det\left(\dfrac{\partial \text{vec}(X)}{\partial \text{vec}(X')'}\right) = (-1)^{mn(m-1)(n-1)/4}.$

(2) X $(m \times m)$: $\det\left(\dfrac{\partial \text{vech}(X)}{\partial \text{vech}(X)'}\right) = 1.$

(3) X $(m \times n)$, $c \in \mathbb{R}$: $\det\left(\dfrac{\partial \text{vec}(cX)}{\partial \text{vec}(X)'}\right) = c^{mn}.$

(4) X $(m \times m)$, $c \in \mathbb{R}$: $\det\left(\dfrac{\partial \text{vech}(cX)}{\partial \text{vech}(X)'}\right) = c^{m(m+1)/2}.$

(5) X $(m \times n)$, A $(m \times m)$, B $(n \times n)$:

$$\det\left(\frac{\partial \text{vec}(AXB)}{\partial \text{vec}(X)'}\right) = (\det A)^n(\det B)^m.$$

(6) X $(m \times n)$, A $(n \times n)$, B $(m \times m)$:

$$\det\left(\frac{\partial \text{vec}(AX'B)}{\partial \text{vec}(X)'}\right) = (-1)^{mn(m-1)(n-1)/4}(\det A)^m(\det B)^n.$$

(7) X, $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(AXB)}{\partial \mathrm{vech}(X)'}\right) = \prod_{i=1}^{m} a_{ii}^{i} b_{ii}^{m-i+1}.$$

(8) X $(m \times m)$ symmetric, A $(m \times m)$:

$$\det\left(\frac{\partial \mathrm{vech}(A'XA)}{\partial \mathrm{vech}(X)'}\right) = (\det A)^{m+1}.$$

(9) X $(m \times m)$ symmetric, A, B $(m \times m)$, $\det(A) \neq 0$, $\lambda_1, \ldots, \lambda_m$ eigenvalues of BA^{-1} :

$$\det\left(\frac{\partial \mathrm{vech}(AXA' \pm BXB')}{\partial \mathrm{vech}(X)'}\right) = (\det A)^{m+1} \prod_{i \geq j} (1 \pm \lambda_i \lambda_j).$$

(10) X $(m \times m)$ symmetric, $A = [a_{ij}], B = [b_{ij}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(AXA' \pm BXB')}{\partial \mathrm{vech}(X)'}\right) = \prod_{i \geq j} (a_{ii}a_{jj} \pm b_{ii}b_{jj}).$$

(11) X $(m \times m)$ symmetric, A, B $(m \times m)$, $\det(B) \neq 0$, $\lambda_1, \ldots, \lambda_m$ eigenvalues of AB^{-1} :

$$\det\left(\frac{\partial \mathrm{vech}(AXB' + BXA')}{\partial \mathrm{vech}(X)'}\right) = 2^m \det(A)(\det B)^m \prod_{i > j} (\lambda_i + \lambda_j).$$

(12) $X, A = [a_{ij}], B = [b_{ij}], C = [c_{ij}], D = [d_{ij}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(AXB + CXD)}{\partial \mathrm{vech}(X)'}\right) = \prod_{i \geq j} (a_{ii}b_{jj} + c_{ii}d_{jj}).$$

(13) X $(m \times m)$ lower triangular, A, B $(m \times m)$ nonsingular:

$$\det\left(\frac{\partial \mathrm{vech}(B'XA + A'X'B)}{\partial \mathrm{vech}(X)'}\right) = 2^m \det(A)(\det B)^m \prod_{i=1}^{m-1} \det(C_{(i)}),$$

where

$$C_{(i)} = \begin{bmatrix} c_{11} & \ldots & c_{1i} \\ \vdots & \ddots & \vdots \\ c_{i1} & \ldots & c_{ii} \end{bmatrix}, \quad i = 1, \ldots, m-1,$$

are the principal submatrices of $C = [c_{ij}] = AB^{-1}$.

(14) $X, A = [a_{ij}]$ $(m \times n)$, B $(m \times m)$, C $(n \times n)$:

$$\det\left(\frac{\partial \operatorname{vec}(A \odot BXC)}{\partial \operatorname{vec}(X)'}\right) = (\det B)^n (\det C)^m \prod_{i=1}^{m} \prod_{j=1}^{n} a_{ij}.$$

(15) X $(m \times n)$, $A = [a_{ij}]$ $(n \times m)$, B $(n \times n)$, C $(m \times m)$:

$$\det\left(\frac{\partial \operatorname{vec}(A \odot BX'C)}{\partial \operatorname{vec}(X)'}\right)$$
$$= (-1)^{mn(m-1)(n-1)/4} (\det B)^m (\det C)^n \prod_{i=1}^{n} \prod_{j=1}^{m} a_{ij}.$$

(16) X, $A = [a_{ij}]$, $B = [b_{ij}]$ $(m \times n)$:

$$\det\left(\frac{\partial \operatorname{vec}(A \odot X \odot B)}{\partial \operatorname{vec}(X)'}\right) = \prod_{i=1}^{m} \prod_{j=1}^{n} a_{ij} b_{ij}.$$

(17) X $(m \times n)$, $A = [a_{ij}]$, $B = [b_{ij}]$ $(n \times m)$:

$$\det\left(\frac{\partial \operatorname{vec}(A \odot X' \odot B)}{\partial \operatorname{vec}(X)'}\right) = (-1)^{mn(m-1)(n-1)/4} \prod_{i=1}^{n} \prod_{j=1}^{m} a_{ij} b_{ij}.$$

Note: The results of this subsection follow from the rules of the previous sections of this chapter and the rules for determinants (see also Magnus (1988, Chapter 8)).

10.8.2 Nonlinear Transformations

(1) (Chain rule)
$x, y(x), z(y)$ $(m \times 1)$:

$$\det\left(\frac{\partial z(y(x))}{\partial x'}\right) = \det\left(\frac{\partial z(y)}{\partial y'}\right) \det\left(\frac{\partial y(x)}{\partial x'}\right).$$

(2) (Chain rule)
$X, Y(X), Z(Y)$ $(m \times n)$:

$$\det\left(\frac{\partial \operatorname{vec} Z(Y(X))}{\partial \operatorname{vec}(X)'}\right) = \det\left(\frac{\partial \operatorname{vec} Z(Y)}{\partial \operatorname{vec}(Y)'}\right) \det\left(\frac{\partial \operatorname{vec} Y(X)}{\partial \operatorname{vec}(X)'}\right).$$

(3) X $(m \times m)$ symmetric with eigenvalues $\lambda_1, \ldots, \lambda_m$:

$$\det\left(\frac{\partial \operatorname{vech}(X^i)}{\partial \operatorname{vech}(X)'}\right) = i^m (\det X)^{i-1} \prod_{k>l} \mu_{kl},$$

where

$$\mu_{kl} = \begin{cases} (\lambda_k^i - \lambda_l^i)/(\lambda_k - \lambda_l) & \text{if} \quad \lambda_k \neq \lambda_l \\ i\lambda_k^{i-1} & \text{if} \quad \lambda_k = \lambda_l \end{cases} .$$

(4) $X = [x_{kl}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(X^i)}{\partial \mathrm{vech}(X)'}\right) = i^m (\det X)^{i-1} \prod_{k>l} \mu_{kl},$$

where

$$\mu_{kl} = \begin{cases} (x_{kk}^i - x_{ll}^i)/(x_{kk} - x_{ll}) & \text{if} \quad x_{kk} \neq x_{ll} \\ i x_{kk}^{i-1} & \text{if} \quad x_{kk} = x_{ll} \end{cases} .$$

(5) X, A $(m \times m)$ symmetric, $\lambda_1, \ldots, \lambda_m$ eigenvalues of XA :

$$\det\left(\frac{\partial \mathrm{vech}(XAX)}{\partial \mathrm{vech}(X)'}\right) = 2^m \det(A) \det(X) \prod_{i>j} (\lambda_i + \lambda_j).$$

(6) $X = [x_{ij}]$, $A = [a_{ij}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(XAX)}{\partial \mathrm{vech}(X)'}\right) = 2^m \det(A) \det(X) \prod_{i>j} (a_{ii} x_{ii} + a_{jj} x_{jj}).$$

(7) $X = [x_{ij}]$ $(m \times m)$ lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(X'X)}{\partial \mathrm{vech}(X)'}\right) = 2^m \prod_{i=1}^m x_{ii}^i,$$

$$\det\left(\frac{\partial \mathrm{vech}(XX')}{\partial \mathrm{vech}(X)'}\right) = 2^m \prod_{i=1}^m x_{ii}^{m-i+1}.$$

(8) X $(m \times m)$ nonsingular:

$$\det\left(\frac{\partial \mathrm{vec}(X^{-1})}{\partial \mathrm{vec}(X)'}\right) = (-1)^m (\det X)^{-2m}.$$

(9) X $(m \times m)$ symmetric nonsingular:

$$\det\left(\frac{\partial \mathrm{vech}(X^{-1})}{\partial \mathrm{vech}(X)'}\right) = (-1)^{m(m+1)/2} (\det X)^{-(m+1)}.$$

(10) X $(m \times m)$ nonsingular, lower triangular:

$$\det\left(\frac{\partial \mathrm{vech}(X^{-1})}{\partial \mathrm{vech}(X)'}\right) = (-1)^{m(m+1)/2} (\det X)^{-(m+1)}.$$

(11) X $(m \times m)$ symmetric nonsingular:

$$\det\left(\frac{\partial \operatorname{vech}(X^{adj})}{\partial \operatorname{vech}(X)'}\right) = (-1)^{m(m+1)/2}(1-m)(\det X)^{(m+1)(m-2)/2}.$$

(12) X $(m \times m)$ nonsingular, lower triangular:

$$\det\left(\frac{\partial \operatorname{vech}(X^{adj})}{\partial \operatorname{vech}(X)'}\right) = (-1)^{m(m+1)/2}(1-m)(\det X)^{(m+1)(m-2)/2}.$$

Note: Most results of this subsection are given in Magnus (1988, Chapter 8). They follow from rules of the previous sections and the results for determinants.

10.9 Matrix Valued Functions of a Scalar Variable

(1) $x \in \mathbb{R}$, $A(x) = A$ $(m \times n)$ constant: $\dfrac{\mathrm{d}A}{\mathrm{d}x} = O_{m\times n}$.

(2) (Linearity)
$x \in \mathbb{R}$, $A(x), B(x)$ $(m \times n), c_1, c_2 \in \mathbb{R}$:

$$\frac{\mathrm{d}[c_1A(x) + c_2B(x)]}{\mathrm{d}x} = c_1\frac{\mathrm{d}A(x)}{\mathrm{d}x} + c_2\frac{\mathrm{d}B(x)}{\mathrm{d}x}.$$

(3) (Product rule)
$x \in \mathbb{R}$, $A(x)$ $(m \times n)$, $B(x)$ $(n \times p)$:

$$\frac{\mathrm{d}A(x)B(x)}{\mathrm{d}x} = A(x)\frac{\mathrm{d}B(x)}{\mathrm{d}x} + \frac{\mathrm{d}A(x)}{\mathrm{d}x}B(x).$$

(4) $x \in \mathbb{R}$, $A(x)$ $(m \times n)$, $B(x)$ $(n \times p)$, $C(x)$ $(p \times q)$:

$$\frac{\mathrm{d}A(x)B(x)C(x)}{\mathrm{d}x}$$
$$= A(x)B(x)\frac{\mathrm{d}C(x)}{\mathrm{d}x} + A(x)\frac{\mathrm{d}B(x)}{\mathrm{d}x}C(x) + \frac{\mathrm{d}A(x)}{\mathrm{d}x}B(x)C(x).$$

(5) $x \in \mathbb{R}$, $A(x)$ $(m \times m)$ nonsingular:

$$\frac{\mathrm{d}A(x)^{-1}}{\mathrm{d}x} = -A(x)^{-1}\frac{\mathrm{d}A(x)}{\mathrm{d}x}A(x)^{-1}.$$

(6) (Ratio rule)
$x \in \mathbb{R}$, $A(x)$ $(m \times m)$ nonsingular, $B(x)$ $(n \times m)$:

$$\frac{\mathrm{d}B(x)A(x)^{-1}}{\mathrm{d}x} = \frac{\mathrm{d}B(x)}{\mathrm{d}x}A(x)^{-1} - B(x)A(x)^{-1}\frac{\mathrm{d}A(x)}{\mathrm{d}x}A(x)^{-1}.$$

(7) (Generalized inverse rule)
$x \in \mathbb{R}$, $A(x)$ $(m \times n)$, $A(x)^-$ some generalized inverse of $A(x)$:

$$A(x)\frac{\mathrm{d}A(x)^-}{\mathrm{d}x}A(x) = -A(x)A(x)^- \frac{\mathrm{d}A(x)}{\mathrm{d}x} A(x)^- A(x).$$

Note: The first four rules follow from the corresponding rules for real valued functions by considering typical elements of the matrices involved. Rule (5) is obtained by applying the product rule to $AA^{-1} = I_m$, (6) follows from (5) and the product rule and the result in (7) follows by applying the product rule to $AA^-A = A$ and multiplying by AA^- from the left.

11

Polynomials, Power Series and Matrices

11.1 Definitions and Notations

11.1.1 Definitions and Notation Related to Polynomials

Polynomial: For given $p_0, p_1, \ldots, p_n \in \mathbb{C}$, a function $p : \mathbb{C} \to \mathbb{C}$ defined by $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ is a (complex) polynomial. (Notation: $p(x)$ or $p(\cdot)$).

Real polynomial: For given $p_0, p_1, \ldots, p_n \in \mathbb{R}$, a function $p : \mathbb{R} \to \mathbb{R}$ defined by $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ is a real polynomial.

Degree of a polynomial: A polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ with $p_n \neq 0$ is said to have degree n.

Root of a polynomial: A number λ for which $p(\lambda) = p_0 + p_1 \lambda + \cdots + p_n \lambda^n = 0$ is called a root of the polynomial $p(\cdot)$.

Multiplicity of the root of a polynomial: A polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ with distinct roots $\lambda_1, \ldots, \lambda_r$ can be written in the form

$$p(x) = p_n (x - \lambda_1)^{k_1} \cdots (x - \lambda_r)^{k_r} = p_n \prod_{i=1}^{r} (x - \lambda_i)^{k_i}$$

where the k_i are positive integers. k_i is called the multiplicity of the root λ_i, $i = 1, \ldots, r$.

Convergent polynomial: A polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ with roots $\lambda_1, \ldots, \lambda_r$ is said to be convergent if $|\lambda_i|_{\mathrm{abs}} < 1$, $i = 1, \ldots, r$.

Divergent polynomial: A polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ with roots $\lambda_1, \ldots, \lambda_r$ is said to be divergent if $|\lambda_i|_{\mathrm{abs}} > 1$ for at least one $i \in \{1, \ldots, r\}$.

Asymptotically stable polynomial: A real polynomial $p(x) = p_0 + p_1x + \cdots + p_nx^n$ with roots $\alpha_1 + \mathrm{i}\beta_1, \ldots, \alpha_r + \mathrm{i}\beta_r$ is said to be asymptotically stable if $\alpha_i < 0$, $i = 1, \ldots, r$.

Unstable polynomial: A real polynomial $p(x) = p_0 + p_1x + \cdots + p_nx^n$ with roots $\alpha_1 + \mathrm{i}\beta_1, \ldots, \alpha_r + \mathrm{i}\beta_r$ is said to be unstable if $\alpha_i > 0$ for at least one $i \in \{1, \ldots, r\}$.

Linear polynomial: A polynomial of degree 1 is linear.

Monic polynomial: A polynomial $p(x) = p_0 + p_1x + \cdots + p_nx^n$ with $p_n = 1$ is called monic.

Factor or **divisor of a polynomial:** If the polynomial $p(x)$ can be written as the product of two polynomials $q(x)$ and $r(x)$, where $q(x)$ has degree > 0, $p(x) = q(x)r(x)$, the polynomial $q(x)$ is said to be a factor or a divisor of the polynomial $p(x)$.

Common factor or **divisor of polynomials:** If the polynomials $p(x)$ and $q(x)$ both have a factor $r(x)$, that is, there exist polynomials $p_1(x)$ and $q_1(x)$ such that $p(x) = r(x)p_1(x)$ and $q(x) = r(x)q_1(x)$, $r(x)$ is said to be a common factor or divisor of the polynomials $p(x)$ and $q(x)$.

Greatest common divisor of two polynomials: Among all monic polynomials which are common divisors of the polynomials $p(x)$ and $q(x)$, the one with the largest degree is said to be the greatest common divisor or factor of $p(x)$ and $q(x)$.

Relatively prime or **coprime polynomials:** The polynomials $p(x)$ and $q(x)$ are relatively prime or coprime if they do not have a common factor.

Dividing polynomial: A polynomial $q(x)$ divides a polynomial $p(x)$ if $q(x)$ is a factor of $p(x)$.

Least common multiple of two polynomials: The least common multiple of the polynomials $p(x)$ and $q(x)$ is the polynomial $r(x)$ with smallest degree such that $p(x)$ and $q(x)$ divide $r(x)$.

Resultant: A function $f : \mathbb{C}^{m+n+1} \longrightarrow \mathbb{C}$ is called a resultant if $f(p_0, p_1, \ldots, p_{m-1}, q_0, q_1, \ldots, q_n) \neq 0 \iff p(x) = p_0 + p_1x + \cdots + p_{m-1}x^{m-1} + x^m$ and $q(x) = q_0 + q_1x + \cdots + q_nx^n$ are relatively prime polynomials.

11.1.2 Matrices Related to Polynomials

Companion matrix: The companion matrix of a monic polynomial $p(x) = p_0 + p_1 x + \cdots + p_{m-1}x^{m-1} + x^m$ is the $(m \times m)$ matrix

$$\begin{bmatrix} -p_{m-1} & -p_{m-2} & \cdots & -p_1 & -p_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & \cdots & -p_{m-2} & -p_{m-1} \end{bmatrix}$$

or

$$\begin{bmatrix} -p_{m-1} & 1 & 0 & \cdots & 0 \\ -p_{m-2} & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -p_1 & 0 & 0 & & 1 \\ -p_0 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & \cdots & 0 & 0 & -p_0 \\ 1 & & 0 & 0 & -p_1 \\ & \ddots & & \vdots & \vdots \\ 0 & & 1 & 0 & -p_{m-2} \\ 0 & \cdots & 0 & 1 & -p_{m-1} \end{bmatrix}.$$

Sylvester matrix: For the polynomials $p(x) = p_0 + p_1 x + \cdots + p_{m-1}x^{m-1} + x^m$ and $q(x) = q_0 + q_1 x + \cdots + q_{n-1}x^{n-1} + q_n x^n$, the corresponding Sylvester matrix is

$$S = \begin{bmatrix} 1 & p_{m-1} & p_{m-2} & \cdots & p_0 & 0 & \cdots & 0 \\ 0 & 1 & p_{m-1} & \cdots & p_1 & p_0 & & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & p_{m-1} & \cdots & p_1 & p_0 \\ q_n & q_{n-1} & q_{n-2} & \cdots & q_0 & 0 & \cdots & 0 \\ 0 & q_n & q_{n-1} & \cdots & q_1 & q_0 & & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & q_n & q_{n-1} & \cdots & q_1 & q_0 \end{bmatrix}$$

$$((m+n) \times (m+n))$$

where the upper part consists of n rows and the lower part consists of m rows. For example, for $p(x) = p_0 + p_1 x + p_2 x^2 + x^3$ and

$q(x) = q_0 + q_1 x + q_2 x^2$, the corresponding Sylvester matrix is

$$S = \begin{bmatrix} 1 & p_2 & p_1 & p_0 & 0 \\ 0 & 1 & p_2 & p_1 & p_0 \\ q_2 & q_1 & q_0 & 0 & 0 \\ 0 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & q_2 & q_1 & q_0 \end{bmatrix}.$$

Hurwitz matrix: The Hurwitz matrix corresponding to the polynomial $p(x) = p_0 + p_1 x + \cdots + p_m x^m$ is

$$H = \begin{bmatrix} p_{m-1} & p_{m-3} & p_{m-5} & \cdots & p_{-m+1} \\ p_m & p_{m-2} & p_{m-4} & \cdots & p_{-m+2} \\ 0 & p_{m-1} & p_{m-3} & \cdots & p_{-m+3} \\ 0 & p_m & p_{m-2} & \cdots & p_{-m+4} \\ 0 & 0 & p_{m-1} & \cdots & p_{-m+5} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & & p_0 \end{bmatrix} \quad (m \times m)$$

where $p_i = 0$ for $i < 0$. For example, for $m = 4$,

$$H = \begin{bmatrix} p_3 & p_1 & 0 & 0 \\ p_4 & p_2 & p_0 & 0 \\ 0 & p_3 & p_1 & 0 \\ 0 & p_4 & p_2 & p_0 \end{bmatrix}$$

is the Hurwitz matrix corresponding to $p(x)$.

Derogatory matrix: An $(m \times m)$ matrix A is derogatory if a polynomial $p(x) = p_0 + p_1 x + \cdots + p_r x^r$ of degree $r < m$ exists such that

$$p(A) = p_0 I_m + p_1 A + \cdots + p_r A^r = O_{m \times m}.$$

11.1.3 Polynomials and Power Series Related to Matrices

Characteristic polynomial: The polynomial defined by

$$p(x) = \det(x I_m - A)$$

is the characteristic polynomial of the $(m \times m)$ matrix A. It is denoted by $p_A(\cdot)$.

Annihilating polynomial: A polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$ is said to annihilate the $(m \times m)$ matrix A if

$$p(A) = p_0 I_m + p_1 A + \cdots + p_n A^n = O_{m \times m}.$$

Minimal polynomial of a matrix: The monic polynomial with minimal degree that annihilates an $(m \times m)$ matrix A is called the minimal polynomial of A. It is denoted by $q_A(\cdot)$.

Matrix polynomial: Given a polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$, the function $p : \mathbb{C}^{m\times m} \to \mathbb{C}^{m\times m}$ (or $p : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$) defined by

$$p(A) \equiv p_0 I_m + p_1 A + \cdots + p_n A^n$$

for $(m \times m)$ matrices A, is a matrix polynomial.

Matrix power series: Given a power series $p(x) = \sum_{n=0}^{\infty} p_n x^n$, the function $p : \mathbb{C}^{m\times m} \to \mathbb{C}^{m\times m}$ (or $p : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$) defined by

$$p(A) \equiv \sum_{n=0}^{\infty} p_n A^n$$

for $(m \times m)$ matrices A, is a matrix power series, provided the infinite sum exists.

Exponential function of a matrix: A $(m \times m)$

$$\exp(A) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

Sine of a matrix: A $(m \times m)$

$$\sin(A) \equiv \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1}.$$

Cosine of a matrix: A $(m \times m)$

$$\cos(A) \equiv \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} A^{2n}.$$

11.2 Results Relating Polynomials and Matrices

Reminder: For an $(m \times m)$ matrix A,

- $p_A(x) = \det(xI_m - A)$ denotes the characteristic polynomial.
- $q_A(x)$ denotes the minimal polynomial.

(1) $A\ (m \times m)$, $\lambda \in \mathbb{C}$:

(a) $q_A(x)$ is unique.

(b) degree$[q_A(x)] \leq m$.

(c) $p(x)$ is a polynomial with $p(A) = 0 \ \Rightarrow\ q_A(x)$ divides $p(x)$.

(d) $p_A(A) = 0$ (Cayley–Hamilton theorem).

(e) $q_A(x)$ divides $p_A(x)$.

(f) $p_A(\lambda) = 0 \iff \lambda$ is eigenvalue of A.

(g) $q_A(\lambda) = 0 \iff \lambda$ is eigenvalue of A.

(2) $p(x)$ a polynomial with companion matrix C: The eigenvalues of C are the roots of $p(x)$.

(3) $A\ (m \times m)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ having multiplicities $m_1, \ldots, m_n$, respectively:

$$q_A(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_n)^{r_n}, \quad \text{where} \quad r_i \leq m_i, \ i = 1, \ldots n,$$

and r_i is the order of the largest block of the Jordan decomposition of A corresponding to the eigenvalue λ_i (see Chapter 6 for the Jordan decomposition).

(4) $A\ (m \times m)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, $q(x) = (x - \lambda_1) \cdots (x - \lambda_n)$: A is diagonalizable $\iff q(A) = 0$.

(5) $A\ (m \times m)$:

(a) A is diagonalizable $\iff$ all roots of $q_A(x)$ have multiplicity 1.

(b) A is diagonalizable $\iff$ all linear factors of $q_A(x)$ are distinct.

(c) A is diagonalizable $\iff [q_A(x) = 0 \ \Rightarrow\ \mathrm{d}q_A(x)/\mathrm{d}x \neq 0]$ (where $\mathrm{d}q_A(x)/\mathrm{d}x$ denotes the derivative of $q_A(x)$).

(6) $p(x)$ a monic polynomial:
C is a companion matrix of $p(x) \ \Rightarrow\ p(x) = q_C(x) = p_C(x)$.

(7) $A\ (m \times m)$. C companion matrix of $p_A(x)$:

(a) A is similar to $C \iff q_A(x) = p_A(x)$.

(b) A is similar to $C \iff$ every eigenvalue of A has geometric multiplicity 1.

(8) $A, B\ (m \times m)$:

(a) A, B similar $\Rightarrow\ q_A(x) = q_B(x)$.

(b) A, B similar $\Rightarrow\ p_A(x) = p_B(x)$.

(c) $p_A(x) = p_B(x) = q_A(x) = q_B(x) \ \Rightarrow\ A, B$ similar .

(9) A, B $(m \times m)$ similar, $p(x)$ a polynomial: $\quad p(A) = 0 \iff p(B) = 0.$

(10) A $(m \times m)$, B $(n \times n)$: $\quad q_{A \oplus B}(x)$ is the least common multiple of $q_A(x)$ and $q_B(x)$.

(11) (Hurwitz theorem)
A real polynomial $p(x)$ is asymptotically stable if and only if all principal minors of its Hurwitz matrix are positive.

Note: Most results of this section can be found in Horn & Johnson (1985). The last result is given in Barnett (1990).

11.3 Polynomial Matrices

11.3.1 Definitions

Polynomial matrix: An $(m \times n)$ matrix $P(x) = [p_{ij}(x)]$ whose typical elements are polynomials $p_{ij}(x) = p_{ij,0} + p_{ij,1}x + \cdots + p_{ij,r_{ij}}x^{r_{ij}}$, is a polynomial matrix. Alternative notation:

$$P(x) = P_0 + P_1 x + \cdots + P_r x^r,$$

where $r = \max_{i,j} r_{ij}$ and

$$P_k = \begin{bmatrix} p_{11,k} & \cdots & p_{1n,k} \\ \vdots & \ddots & \vdots \\ p_{m1,k} & \cdots & p_{mn,k} \end{bmatrix}, \quad k = 0, \ldots, r,$$

with $p_{ij,k} = 0$ for $k > r_{ij}$.

Real polynomial matrix: A polynomial matrix whose elements are real polynomials is said to be a real polynomial matrix.

Degree of a polynomial matrix: The degree of an $(m \times n)$ polynomial matrix $P(x) = [p_{ij}(x)]$ with typical elements $p_{ij}(x) = p_{ij,0} + p_{ij,1}x + \cdots + p_{ij,r_{ij}}x^{r_{ij}}$, is $\max_{i,j} r_{ij}$, where r_{ij} is the degree of $p_{ij}(x)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Equivalently, the degree of $P(x) = P_0 + P_1 x + \cdots + P_r x^r$ is r if $P_r \neq 0$.

Polynomial matrix operations: All matrix operations specified in Chapter 1 are defined analogously for polynomial matrices except where noted otherwise in the following.

Rank of a polynomial matrix: The rank of a polynomial matrix is the number of columns of the largest submatrix whose determinant is not identically zero.

Eigenvalues or **latent roots of a polynomial matrix:** The eigenvalues or latent roots of the polynomial matrix $P(x)$ are the roots of the polynomial $\det P(x)$.

Latent vector of a polynomial matrix: Let $P(x)$ be a polynomial matrix with eigenvalue λ. A vector v such that $P(\lambda)v = 0$ is called a latent vector of $P(x)$.

Regular polynomial matrix: The $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1x + \cdots + P_rx^r$ is called regular if P_r is nonsingular.

Monic polynomial matrix: The $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1x + \cdots + P_rx^r$ is monic if $P_r = I_m$.

Unimodular or **invertible polynomial matrix:** The $(m \times m)$ polynomial matrix $P(x)$ is said to be unimodular or invertible if $\det P(x) = \text{constant} \neq 0$, that is, $\det P(x)$ is a constant function.

Left and **right multiples of polynomial matrices:** The $(m \times n)$ polynomial matrix $P(x)$ satisfying $P(x) = Q(x)T(x)$ for polynomial matrices $Q(x)$ $(m \times h)$ and $T(x)$ $(h \times n)$ is called a left multiple of $T(x)$ and a right multiple of $Q(x)$.

Left quotient and **left divisor:** Consider polynomial matrices

$$P(x) = P_0 + P_1x + \cdots + P_rx^r \quad (m \times n), \quad P_r \neq 0,$$

and

$$Q(x) = Q_0 + Q_1x + \cdots + Q_sx^s \quad (m \times h), \quad \det(Q_s) \neq 0 \text{ and } s \leq r.$$

That is, the latter polynomial is regular. An $(h \times n)$ polynomial matrix $T(x)$ is a **left quotient** of $P(x)$ and $Q(x)$ if an $(m \times n)$ polynomial matrix $R(x)$ with degree less than s exists such that $P(x) = Q(x)T(x) + R(x)$. The polynomial matrix $R(x)$ is said to be a **left remainder**. $Q(x)$ is called a **left divisor** of $P(x)$ if $R(x) \equiv 0$ and in this case $P(x)$ is said to be **left divisible** by $Q(x)$.

Right quotient and **right divisor:** Consider polynomial matrices

$$P(x) = P_0 + P_1x + \cdots + P_rx^r \quad (m \times n), \quad P_r \neq 0,$$

and

$$Q(x) = Q_0 + Q_1x + \cdots + Q_sx^s \quad (m \times h), \quad \det(Q_s) \neq 0 \text{ and } s \leq r.$$

That is, the latter polynomial is regular. An $(m \times h)$ polynomial matrix $T(x)$ is a **right quotient** of $P(x)$ and $Q(x)$ if an $(m \times n)$

polynomial matrix $R(x)$ with degree less than s exists such that $P(x) = T(x)Q(x) + R(x)$. The polynomial matrix $R(x)$ is said to be a **right remainder**. $Q(x)$ is called a **right divisor** of $P(x)$ if $R(x) \equiv 0$ and in this case $P(x)$ is said to be **right divisible** by $Q(x)$.

Common left divisor: A polynomial matrix $D(x)$ is a **common left divisor** of the polynomial matrices $P(x)$ and $Q(x)$ if $D(x)$ is a left divisor of both $P(x)$ and $Q(x)$, that is, if polynomial matrices $P_1(x)$ and $Q_1(x)$ exist such that $P(x) = D(x)P_1(x)$ and $Q(x) = D(x)Q_1(x)$. $D(x)$ is called a **greatest common left divisor** of $P(x)$ and $Q(x)$ if for any other common left divisor $F(x)$ there exists a polynomial matrix $B(x)$ such that $D(x) = F(x)B(x)$. The polynomial matrices $P(x)$ and $Q(x)$ are said to be **relatively left prime** or **left coprime** if their greatest common left divisors are unimodular.

Common right divisor: A polynomial matrix $D(x)$ is a **common right divisor** of the polynomial matrices $P(x)$ and $Q(x)$ if $D(x)$ is a right divisor of both $P(x)$ and $Q(x)$, that is, if polynomial matrices $P_1(x)$ and $Q_1(x)$ exist such that $P(x) = P_1(x)D(x)$ and $Q(x) = Q_1(x)D(x)$. $D(x)$ is called a **greatest common right divisor** of $P(x)$ and $Q(x)$ if for any other common right divisor $F(x)$ there exists a polynomial matrix $B(x)$ such that $D(x) = B(x)F(x)$. The polynomial matrices $P(x)$ and $Q(x)$ are said to be **relatively right prime** or **right coprime** if their greatest common right divisors are unimodular.

Skew prime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be skew prime if det $P(x)$ and det $Q(x)$ are relatively prime polynomials.

Elementary operations for polynomial matrices: The following modifications of a polynomial matrix are called elementary operations:

(i) interchanging two rows or two columns,

(ii) multiplying any row or column by a nonzero number,

(iii) adding to one row another row multiplied by an arbitrary polynomial,

(iv) adding to one column another column multiplied by an arbitrary polynomial.

Elementary polynomial matrix: An $(m \times m)$ polynomial matrix is elementary if it may be obtained by applying a single elementary operation to I_m.

Equivalent polynomial matrices: Two polynomial matrices $P(x)$ and $Q(x)$ are said to be equivalent if unimodular polynomial matrices $U(x)$ and $V(x)$ exist such that $P(x) = U(x)Q(x)V(x)$.

Characteristic matrix: The $(m \times m)$ polynomial matrix $P(x) = xI_m - A$ is the characteristic matrix of an $(m \times m)$ matrix A.

Invariant factors or **invariant polynomials of a polynomial matrix:** The invariant factors or polynomials of an $(m \times m)$ polynomial matrix $P(x)$ of rank r are defined as

$$i_k(x) = \frac{d_k(x)}{d_{k-1}(x)}, \quad k = 1, \ldots, r,$$

where $d_0(x) \equiv 1$ and $d_k(x)$ is the monic greatest common divisor of all minors of order k of $P(x)$ for $k = 1, \ldots, r$.

Elementary divisors of a polynomial matrix: Let λ be a root of an invariant factor of a polynomial matrix $P(x)$. Then the linear polynomial $p(x) = x - \lambda$ is an elementary divisor of $P(x)$.

Block companion matrix of a polynomial matrix: The (block) companion matrix of an $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1 x + \cdots + P_{r-1}x^{r-1} + I_m x^r$ is the $(rm \times rm)$ matrix

$$\begin{bmatrix} -P_{r-1} & -P_{r-2} & \cdots & -P_1 & -P_0 \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & I_m & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & I_m & 0 \\ 0 & 0 & \cdots & 0 & I_m \\ -P_0 & -P_1 & \cdots & -P_{r-2} & -P_{r-1} \end{bmatrix}$$

or

$$\begin{bmatrix} -P_{r-1} & I_m & 0 & \ldots & 0 \\ -P_{r-2} & 0 & I_m & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -P_1 & 0 & 0 & & I_m \\ -P_0 & 0 & 0 & \ldots & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & \cdots & 0 & 0 & -P_0 \\ I_m & & 0 & 0 & -P_1 \\ & \ddots & & \vdots & \vdots \\ 0 & & I_m & 0 & -P_{r-2} \\ 0 & \cdots & 0 & I_m & -P_{r-1} \end{bmatrix}.$$

11.3.2 Results for Polynomial Matrices

The results for matrix operations listed in Chapter 2 remain valid for polynomial matrices. Here a few special results are given.

(1) $P(x)$ $(m \times m)$ polynomial matrix of degree r :

(a) The degree of det $P(x) \leq mr$.

(b) Elementary operations on $P(x)$ do not alter rk $P(x)$.

(c) Elementary operations on $P(x)$ do not alter its invariant factors.

(d) $P(x)$ is an elementary polynomial matrix
$\Rightarrow$ $P(x)$ is unimodular.

(e) $P(x)$ is unimodular

$$\Rightarrow \quad P(x)^{-1} = \frac{1}{\det P(x)} P(x)^{adj}$$

is a polynomial matrix of finite degree.

(2) $r \geq 1, P(x) = P_0 + P_1 x + \cdots + P_r x^r$ $(m \times m)$ polynomial matrix, $P_r \neq 0$: $P(x)$ is unimodular $\Rightarrow$ det $P_r = 0$.

(3) $P(x)$ $(m \times m)$ monic polynomial matrix: det $P(x) = i_1(x) \cdots i_m(x)$, where the $i_k(x)$ are the invariant factors of $P(x)$.

(4) $P(x)$ $(m \times m)$ monic polynomial matrix of degree r and corresponding block companion matrix C: $\det(xI_{rm} - C) = \det P(x)$.

(5) $P(x) = P_0 + P_1 x + \cdots + P_r x^r$, $Q(x) = Q_0 + Q_1 x + \cdots + Q_s x^s$ $(m \times m)$ polynomial matrices, $r \geq s$:

(a) $Q(x)$ is regular $\Rightarrow$ the left (right) quotients and remainders of $P(x)$ and $Q(x)$ are unique.

(b) A greatest common left (right) divisor of $P(x)$ and $Q(x)$ is unique up to postmultiplication (premultiplication) by an arbitrary unimodular polynomial matrix.

(6) $P(x), Q(x)$ $(m \times m)$ polynomial matrices:

(a) $P(x), Q(x)$ are unimodular $\Rightarrow$ $P(x)Q(x)$ and $Q(x)P(x)$ are unimodular.

(b) $P(x), Q(x)$ are left coprime $\Rightarrow$ $P(x)'$ and $Q(x)'$ are right coprime.

(c) $P(x), Q(x)$ are right coprime $\Rightarrow$ $P(x)'$ and $Q(x)'$ are left coprime.

(d) $P(x), Q(x)$ are left coprime $\Rightarrow$ there exist $(m \times m)$ polynomial matrices $T(x), S(x)$ such that $P(x)T(x) + Q(x)S(x) \equiv I_m$.

(e) $P(x), Q(x)$ are right coprime $\Rightarrow$ there exist $(m \times m)$ polynomial matrices $T(x), S(x)$ such that $T(x)P(x) + S(x)Q(x) \equiv I_m$.

(f) $P(x), Q(x)$ are left coprime $\Rightarrow$ $\text{rk}[P(x) : Q(x)] = m$.

(g) $P(x), Q(x)$ are right coprime

$$\Rightarrow \text{rk}\begin{bmatrix} P(x) \\ Q(x) \end{bmatrix} = m.$$

(h) $P(x), Q(x)$ are left coprime $\Rightarrow$ there exists a unimodular polynomial matrix $T(x)$ such that $[P(x) : Q(x)]T(x) = [I_m : 0]$.

(i) $P(x), Q(x)$ are right coprime $\Rightarrow$ there exists a unimodular polynomial matrix $T(x)$ such that

$$T(x)\begin{bmatrix} P(x) \\ Q(x) \end{bmatrix} = \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

(j) $P(x), Q(x)$ are left coprime $\Rightarrow$ $[P(x) : Q(x)]$ is equivalent to $[I_m : 0]$.

(k) $P(x), Q(x)$ are right coprime

$$\Rightarrow \begin{bmatrix} P(x) \\ Q(x) \end{bmatrix} \text{ is equivalent to } \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

(l) $P(x), Q(x)$ are skew prime $\Rightarrow$ $P(x), Q(x)$ are left and right coprime.

(7) $P(x), Q(x)$ $(m \times m)$ monic polynomial matrices of degrees r and s, respectively: $P(x), Q(x)$ are skew prime $\Longleftrightarrow$ the equation $X(x)P(x) + Q(x)Y(x) = I_m$ has a unique solution for polynomial matrices $X(x)$ and $Y(x)$ with degree $X(x) < s$ and degree $Y(x) < r$.

(8) $P(x), Q(x)$ $(m \times m)$ monic polynomial matrices of degrees r and s, respectively, C block companion matrix of $P(x)$: $P(x), Q(x)$ are skew prime $\Longleftrightarrow$

$$Q_0 \otimes I_{mr} + Q_1 \otimes C + \cdots + Q_{s-1} \otimes C^{s-1} + Q_s \otimes C^s$$

is nonsingular.

(9) $P(x) = P_0 + P_1 x$, $Q(x) = Q_0 + Q_1 x$ $(m \times m)$ polynomial matrices: P_1 and Q_1 are nonsingular $\Rightarrow$ $P(x)$ and $Q(x)$ are equivalent.

(10) $P(x), Q(x)$ $(m \times m)$ polynomial matrices: $P(x), Q(x)$ are equivalent $\Longleftrightarrow$ $P(x)$ and $Q(x)$ have the same invariant factors.

(11) (Smith normal (canonical) form)
$P(x)$ $(m \times m)$ polynomial matrix, $\text{rk}\, P(x) = r$: There exist

unimodular polynomial matrices $U(x)$ and $V(x)$ such that

$$P(x) = U(x) \begin{bmatrix} \begin{bmatrix} i_1(x) & & 0 \\ & \ddots & \\ 0 & & i_r(x) \end{bmatrix} & 0 \\ 0 & 0 \end{bmatrix} V(x)$$

where the $i_k(x)$ are the invariant factors of $P(x)$.

Note: The results listed in this subsection can be found, e.g., in Barnett (1990). For the Smith normal form see also Gantmacher (1959a).

Appendix A

Dictionary of Matrices and Related Terms

!!! Warning !!!

Many terms given in the following are defined differently in some of the matrix literature. The precise meaning of a given term in other literature should be checked carefully in each individual case.

Absolute value of a matrix: The absolute value of an $(m \times n)$ matrix $A = [a_{ij}]$ is defined as

$$|A|_{\text{abs}} \equiv [\ |a_{ij}|_{\text{abs}}\ ] = \begin{bmatrix} |a_{11}|_{\text{abs}} & |a_{12}|_{\text{abs}} & \cdots & |a_{1n}|_{\text{abs}} \\ |a_{21}|_{\text{abs}} & |a_{22}|_{\text{abs}} & \cdots & |a_{2n}|_{\text{abs}} \\ \vdots & \vdots & & \vdots \\ |a_{m1}|_{\text{abs}} & |a_{m2}|_{\text{abs}} & \cdots & |a_{mn}|_{\text{abs}} \end{bmatrix} \quad (m \times n)$$

where $|c|_{\text{abs}}$ denotes the modulus of the complex number $c = c_1 + ic_2$ defined as $|c|_{\text{abs}} = \sqrt{c_1^2 + c_2^2} = \sqrt{c\bar{c}}$. Here $\bar{c}$ is the complex conjugate of c. (For the properties of the absolute value of a matrix see Section 3.8.)

Addition of matrices: $A = [a_{ij}]$ $(m \times n)$, $B = [b_{ij}]$ $(m \times n)$

$$A + B \equiv [a_{ij} + b_{ij}] \quad (m \times n).$$

(For the rules see Section 2.1).

Adjoint of a matrix: For $m \geq 2$, the $(m \times m)$ matrix $A^{adj} = [\text{cof}(a_{ij})]'$ is the adjoint of the $(m \times m)$ matrix $A = [a_{ij}]$. Here $\text{cof}(a_{ij})$ is the cofactor of a_{ij}. For instance, for $m = 3$,

$$A^{adj} = \begin{bmatrix} \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ -\det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \\ \det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} & -\det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{bmatrix}'.$$

For $m = 1$, A^{adj} is defined to be 1. (See Section 3.4 for the properties of the adjoint.)

Adjugate of a matrix: The *adjoint* of a marix is sometimes called its adjugate.

Algebraic multiplicity of an eigenvalue: Let $\lambda_1, \ldots, \lambda_n$ be the distinct eigenvalues of the $(m \times m)$ matrix A. Then the *characteristic polynomial* of A can be represented as

$$p_A(\lambda) = p_0 + p_1\lambda + \cdots + p_{m-1}\lambda^{m-1} + \lambda^m = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_n)^{m_n}.$$

where the m_i are positive integers with $\sum_{i=1}^n m_i = m$. The number m_i is the multiplicity or algebraic multiplicity of the eigenvalue λ_i, $i = 1, \ldots, n$. (For more on eigenvalues see Chapter 5.)

Annihilating polynomial: A polynomial $p(x) = p_0 + p_1x + \cdots + p_nx^n$ is said to annihilate the $(m \times m)$ matrix A if $p(A) = p_0 I_m + p_1 A + \cdots + p_n A^n = O_{m\times m}$. (See Section 11.1.1.)

Antimetric matrix: An $(m \times m)$ matrix A is antimetric or skew-symmetric if $A = -A'$. (See *skew-symmetric matrix.*)

Arithmetic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called arithmetic matrix if complex numbers a and d exist such that $a_{ij} = a + (i + j - 2)d$. For example, for $m = 3$,

$$A = \begin{bmatrix} a & a+d & a+2d \\ a+d & a+2d & a+3d \\ a+2d & a+3d & a+4d \end{bmatrix}$$

is an arithmetic matrix.

Band matrix or **banded matrix:** For $r, s < m$, an $(m \times m)$ matrix

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} & 0 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} & a_{2,r+1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ a_{s,1} & a_{s,2} & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & a_{s+1,2} & & \ddots & \ddots & \ddots & & 0 \\ 0 & 0 & \ddots & & \ddots & \ddots & & a_{m-r+1,m} \\ 0 & 0 & \ddots & \ddots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & a_{m,m-s+1} & \cdots & \cdots & a_{mm} \end{bmatrix} = [a_{i,j}]$$

with $a_{i,j} = 0$ for $i - j \geq s$ and for $j - i \geq r$ is a band matrix or banded matrix with bandwidth $r + s + 1$.

Bidiagonal matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is bidiagonal if $a_{ij} = 0$ for $i > j$ and $j > i + 1$ or for $i > j + 1$ and $j > i$, that is, if

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 \\ 0 & a_{22} & a_{23} & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & a_{m-1,m} \\ 0 & \cdots & \cdots & 0 & a_{mm} \end{bmatrix}$$

or

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ 0 & a_{32} & a_{33} & & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & a_{m,m-1} & a_{mm} \end{bmatrix}.$$

Binary matrix: An $(m \times n)$ matrix is said to be binary if all elements are 0 or 1. For instance,

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

is a binary matrix.

Block band matrix: An $(m \times n)$ matrix

$$\begin{bmatrix} A_{11} & \cdots & A_{1p} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & 0 \\ A_{s1} & & & \ddots & & A_{q-p+1,q} \\ 0 & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \\ 0 & \cdots & 0 & A_{q,q-s+1} & \cdots & A_{qq} \end{bmatrix} = [A_{ij}]$$

consisting of $(m_i \times n_j)$ submatrices $A_{ij},\ i,j = 1,\ldots,q$, is a block band matrix if $A_{ij} = 0$ for $i - j \geq s$ and for $j - i \geq p$. Here $s, p < q$.

Block circulant matrix: An $(n \times n)$ matrix of the form

$$\begin{bmatrix} A_1 & A_2 & A_3 & \cdots & A_{m-1} & A_m \\ A_m & A_1 & A_2 & \cdots & A_{m-2} & A_{m-1} \\ A_{m-1} & A_m & A_1 & \cdots & A_{m-3} & A_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ A_3 & A_4 & A_5 & \cdots & A_1 & A_2 \\ A_2 & A_3 & A_4 & \cdots & A_m & A_1 \end{bmatrix},$$

where the A_i are $(k \times k)$ matrices, is a block circulant matrix.

Block companion matrix of a polynomial matrix: The (block) companion matrix of an $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1 x + \cdots + P_{r-1}x^{r-1} + I_m x^r$ is the $(rm \times rm)$ matrix

$$\begin{bmatrix} -P_{r-1} & -P_{r-2} & \cdots & -P_1 & -P_0 \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & I_m & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & I_m & 0 \\ 0 & 0 & \cdots & 0 & I_m \\ -P_0 & -P_1 & \cdots & -P_{r-2} & -P_{r-1} \end{bmatrix}$$

or

$$\begin{bmatrix} -P_{r-1} & I_m & 0 & \dots & 0 \\ -P_{r-2} & 0 & I_m & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -P_1 & 0 & 0 & & I_m \\ -P_0 & 0 & 0 & \dots & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & \cdots & 0 & 0 & -P_0 \\ I_m & & 0 & 0 & -P_1 \\ & \ddots & & \vdots & \vdots \\ 0 & & I_m & 0 & -P_{r-2} \\ 0 & \cdots & 0 & I_m & -P_{r-1} \end{bmatrix}.$$

(See Section 11.3 for more details.)

Block diagonal matrix: An $(m \times n)$ matrix

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{bmatrix}$$

with $(m_i \times n_j)$ submatrices $A_i, i = 1, \dots, p$, along the main diagonal and zeros elsewhere, is a block diagonal matrix. (See Section 9.11 for properties.)

Block Hankel matrix: For given $(n \times n)$ matrices $B_i, i = 1, \dots, 2m - 1$, the matrix

$$\begin{bmatrix} B_1 & B_2 & B_3 & \cdots & B_m \\ B_2 & B_3 & B_4 & \cdots & B_{m+1} \\ B_3 & B_4 & B_5 & & B_{m+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B_m & B_{m+1} & B_{m+2} & \cdots & B_{2m-1} \end{bmatrix} = [A_{ij}]$$

with $A_{ij} = B_{i+j-1}$ is called a block Hankel matrix.

Block matrix: An $(m \times n)$ matrix $A = [A_{ij}]$, consisting of $(m_i \times n_j)$ submatrices $A_{ij}, i = 1, \dots, p,\ j = 1, \dots, q$, is a partitioned or block matrix. (See Section 9.11 for the properties.)

Block Toeplitz matrix: A matrix of the form

$$\begin{bmatrix} A_1 & A_{q+1} & A_{q+2} & \cdots & A_{2q-2} & A_{2q-1} \\ A_2 & A_1 & A_{q+1} & \cdots & A_{2q-3} & A_{2q-2} \\ A_3 & A_2 & A_1 & & A_{2q-4} & A_{2q-3} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ A_{q-1} & A_{q-2} & A_{q-3} & \ddots & A_1 & A_{q+1} \\ A_q & A_{q-1} & A_{q-2} & \cdots & A_2 & A_1 \end{bmatrix} = [A_{ij}]$$

with $A_{ij} = A_{i+k,j+k}$ for all i, j, k, is a block Toeplitz matrix. Here the A_{ij} are $(p \times p)$ matrices.

Block triangular matrix: An $(m \times m)$ matrix $A = [A_{ij}]$, consisting of $(m_i \times n_j)$ submatrices A_{ij} is a block triangular matrix if $A_{ij} = 0$ for $i > j$ or if $A_{ij} = 0$ for $i < j$. (See also *lower block triangular matrix* and *upper block triangular matrix.*)

Block tridiagonal matrix: An $(m \times m)$ matrix $A = [A_{ij}]$, consisting of $(m_i \times n_j)$ submatrices A_{ij} is a block tridiagonal matrix if $A_{ij} = 0$ for $|i - j|_{\text{abs}} > 1$, that is,

$$A = \begin{bmatrix} A_{11} & A_{12} & 0 & \cdots & 0 & 0 \\ A_{21} & A_{22} & A_{23} & & 0 & 0 \\ 0 & A_{32} & A_{33} & & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & & \ddots & A_{p-1,p-1} & A_{p-1,p} \\ 0 & 0 & \cdots & & A_{p,p-1} & A_{p,p} \end{bmatrix}.$$

Bordered Gramian matrix: An $((m+n) \times (m+n))$ matrix

$$\begin{bmatrix} A & B \\ B^H & 0 \end{bmatrix},$$

with A being a positive semidefinite $(m \times m)$ matrix and B being an $(m \times n)$ matrix, is a bordered Gramian matrix.

Bordered matrix: An $((m+n) \times (m+p))$ matrix

$$\begin{bmatrix} A & B \\ C & 0 \end{bmatrix},$$

where A is $(m \times m)$, B is $(m \times p)$, C is $(n \times m)$, is a bordered matrix.

Brownian matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a Brownian matrix if

$$a_{ij} = \begin{cases} a_{i,j+1}, & j > i \\ a_{i+1,j}, & i > j \end{cases}$$

for $i, j = 1, \ldots, m$. For instance, for $m = 4$,

$$A = \begin{bmatrix} a_1 & a_2 & a_2 & a_2 \\ a_3 & a_4 & a_5 & a_5 \\ a_3 & a_6 & a_7 & a_8 \\ a_3 & a_6 & a_9 & a_{10} \end{bmatrix}$$

is a Brownian matrix.

Centro-Hermitian matrix: An $(m\times m)$ matrix $A = [a_{ij}]$ is called centro-Hermitian if $a_{ij} = \bar{a}_{m+1-i,m+1-j}$, $i,j = 1,\dots,m$. For example, the (4×4) matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_6 & a_7 & a_8 \\ \bar{a}_8 & \bar{a}_7 & \bar{a}_6 & \bar{a}_5 \\ \bar{a}_4 & \bar{a}_3 & \bar{a}_2 & \bar{a}_1 \end{bmatrix}$$

is centro-Hermitian.

Centro-symmetric matrix: An $(m\times m)$ matrix $A = [a_{ij}]$ is centro-symmetric if $a_{ij} = a_{m+1-i,m+1-j}$, $i,j = 1,\dots,m$. For example, the (4×4) matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_6 & a_7 & a_8 \\ a_8 & a_7 & a_6 & a_5 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix}$$

is centro-symmetric.

Characteristic determinant: The polynomial in λ defined by the determinant $\det(\lambda I_m - A)$ is the characteristic determinant or characteristic polynomial of the $(m\times m)$ matrix A. (See *characteristic polynomial.*)

Characteristic equation: The equation $\det(\lambda I_m - A) = 0$ is the characteristic equation of the $(m\times m)$ matrix A. (See Section 5.1 for more details.)

Characteristic matrix: The $(m\times m)$ polynomial matrix $P(x) = xI_m - A$ is the characteristic matrix of an $(m\times m)$ matrix A. (See Section 5.1 for more details.)

Characteristic polynomial: The polynomial in λ given by $\det(\lambda I_m - A)$ is the characteristic polynomial of the $(m\times m)$ matrix A. (See Section 5.1 for more details.)

Characteristic root or **value:** The roots of the characteristic polynomial of an $(m\times m)$ matrix A are the eigenvalues, the characteristic values, the characteristic roots or the latent roots of A. (See also *eigenvalue of a matrix* and *latent root* and refer to Chapter 5 for details.)

Characteristic root of a matrix in the metric of another matrix: For $(m\times m)$ matrices A and B, the roots of the polynomial $p(\lambda) = \det(\lambda B - A)$ are called eigenvalues or characteristic roots of A in the metric of B. (See also *eigenvalue of a matrix in the metric of another matrix* and refer to Section 5.1 for further details.)

Characteristic vector: An $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda v$, for an $(m \times m)$ matrix A and a complex number λ, is a characteristic vector or eigenvector of A corresponding to or associated with the characteristic value λ. (See also *eigenvector* and refer to Chapter 5 for properties.)

Characteristic vector of a matrix in the metric of another matrix: Given $(m \times m)$ matrices A and B, an $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda Bv$ for some complex number λ, is called a characteristic vector or eigenvector of A in the metric of B. (See also *eigenvector of a matrix in the metric of another matrix* and refer to Section 5.1 for further details.)

Check matrix: For $n > m$, an $(m \times n)$ matrix $A = [I_m : B]$ is a check matrix or parity check matrix if B is a binary matrix, that is, all elements of B are 0 or 1. For instance,

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

is a check matrix.

Circulant matrix: For $a_1, \dots, a_m \in \mathbb{C}$, an $(m \times m)$ matrix

$$\operatorname{circ}(a_1, \dots, a_m) \equiv \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{m-1} & a_m \\ a_m & a_1 & a_2 & \cdots & a_{m-2} & a_{m-1} \\ a_{m-1} & a_m & a_1 & \cdots & a_{m-3} & a_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_3 & a_4 & a_5 & \cdots & a_1 & a_2 \\ a_2 & a_3 & a_4 & \cdots & a_m & a_1 \end{bmatrix}$$

is a circulant matrix. (See Section 9.1 for its properties.)

Cofactor matrix: The $(m \times m)$ matrix $[\operatorname{cof}(a_{ij})]$ is the cofactor matrix of the $(m \times m)$ matrix $A = [a_{ij}]$. (See also Section 3.4.)

Cofactor of an element of a matrix: For an $(m \times m)$ matrix $A = [a_{ij}]$, the cofactor of the ijth element a_{ij} is

$$\operatorname{cof}(a_{ij}) \equiv (-1)^{i+j} \operatorname{minor}(a_{ij}),$$

where $\text{minor}(a_{ij})$ is the determinant of the matrix obtained by deleting the ith row and jth column from A,

$$\text{minor}(a_{ij}) \equiv \det \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,m} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,m} \end{bmatrix}.$$

(See also Section 3.4.)

Cogredient matrices: The $(m \times m)$ matrices A and B are cogredient if a *permutation matrix* P exists such that $A = P'BP$.

Column dimension of a matrix: The number of columns of a matrix is its column dimension.

Column rank: The column rank of a matrix A (col rk A or col rk(A)) is the maximum number of linearly independent columns of A. (See Section 4.3 for further details.)

Column regular matrix: An $(m \times n)$ matrix A is column regular if $\text{rk}(A) = n$.

Column stochastic matrix: An $(m \times m)$ matrices $A = [a_{ij}]$ is said to be column stochastic if $A \geq 0$, that is, $a_{ij} \geq 0$, $i, j = 1, \ldots, m$, and $\sum_{i=1}^{m} a_{ij} = 1$, $j = 1, \ldots, m$. (See also Section 9.9.)

Column vector: An $(m \times 1)$ matrix is a column vector.

Column vectorization of a matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ $(m \times n)$ can be vectorized by stacking its columns in one vector. Notation:

$$\text{vec } A = \text{vec}(A) = \text{col}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ a_{13} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1).$$

(See Section 7.2 for the rules.)

Common left divisor of polynomial matrices: A polynomial matrix $D(x)$ is a common left divisor of the polynomial matrices $P(x)$ and $Q(x)$ if $D(x)$ is a left divisor of both $P(x)$ and $Q(x)$, that is, if polynomial matrices $P_1(x)$ and $Q_1(x)$ exist such that $P(x) = D(x)P_1(x)$ and $Q(x) = D(x)Q_1(x)$. (See Section 11.3 for further details.)

Common right divisor of polynomial matrices: A polynomial matrix $D(x)$ is a common right divisor of the polynomial matrices $P(x)$ and $Q(x)$ if $D(x)$ is a right divisor of both $P(x)$ and $Q(x)$, that is, if polynomial matrices $P_1(x)$ and $Q_1(x)$ exist such that $P(x) = P_1(x)D(x)$ and $Q(x) = Q_1(x)D(x)$. (See Section 11.3 for further details.)

Commutation matrix: The $(mn \times mn)$ commutation matrix K_{mn} (or $K_{m,n}$) is defined such that $\text{vec}(A') = K_{mn}\text{vec}(A)$ for any $(m \times n)$ matrix A. For example,

$$K_{32} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(See Section 9.2 for the properties of commutation matrices.)

Commuting matrices: The $(m \times m)$ matrices A and B commute if $AB = BA$.

Companion matrix: The companion matrix of a monic polynomial $p(x) = p_0 + p_1 x + \cdots + p_{m-1}x^{m-1} + x^m$ is the $(m \times m)$ matrix

$$\begin{bmatrix} -p_{m-1} & -p_{m-2} & \cdots & -p_1 & -p_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & \cdots & -p_{m-2} & -p_{m-1} \end{bmatrix}$$

or

$$\begin{bmatrix} -p_{m-1} & 1 & 0 & \cdots & 0 \\ -p_{m-2} & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ -p_1 & 0 & 0 & & 1 \\ -p_0 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & \cdots & 0 & 0 & -p_0 \\ 1 & & 0 & 0 & -p_1 \\ & \ddots & \vdots & \vdots & \vdots \\ 0 & & 1 & 0 & -p_{m-2} \\ 0 & \cdots & 0 & 1 & -p_{m-1} \end{bmatrix}.$$

(See also Section 11.1.)

Compatible norms: Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be norms for $(m \times 1)$ and $(n \times 1)$ vectors, respectively. The norm $\|\cdot\|$ for $(m \times n)$ matrices is compatible with $\|\cdot\|_a$ and $\|\cdot\|_b$ if

$$\|Ax\|_a \leq \|A\|\,\|x\|_b$$

for all $(m \times n)$ matrices A and $(n \times 1)$ vectors x. A norm $\|\cdot\|$ for $(m \times m)$ square matrices is compatible with $\|\cdot\|_a$ if

$$\|Ax\|_a \leq \|A\|\,\|x\|_a$$

for all $(m \times m)$ matrices A and $(m \times 1)$ vectors x. (See Chapter 8 for details on norms.)

Complex conjugate matrix: For an $(m \times n)$ matrix $A = [a_{ij}]$, the complex conjugate matrix $\bar{A}$ is obtained by replacing all elements a_{ij} by their complex conjugates $\bar{a}_{ij}$, that is, $\bar{A} = [\bar{a}_{ij}]$. (See Section 3.2 for the properties.)

Complex matrix: The $(m \times n)$ matrix $A = [a_{ij}]$ is a complex matrix if the a_{ij} are complex numbers.

Compound matrix: $B = [b_{ij}]$ is a compound matrix of A if b_{ij} is a minor of a given size of A. For example, for a (3×3) matrix $A = [a_{ij}]$,

$$B = \begin{bmatrix} \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} & \det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix} \\ \det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} & \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} \\ \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} & \det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \end{bmatrix}$$

is a compound matrix of A.

Condition number: The condition number of an $(m \times n)$ matrix A is the ratio $\sigma_{max}/\sigma_{min}$, where σ_{max} and σ_{min} denote the largest and smallest singular values of A, respectively. (See Chapter 5 for details on singular values.)

Conditional inverse matrix: A generalized inverse matrix is sometimes called conditional inverse matrix. (See *generalized inverse matrix.*)

Conformable matrices: The matrices A and B are conformable if their product AB is defined, that is, if the number of columns of A is equal to the number of rows of B.

Congruent matrices: The $(m \times m)$ matrices A and B are congruent if a nonsingular $(m \times m)$ matrix C exists such that $B = C^H AC$. (See Chapter 6 for some congruent matrices.)

Conjugate matrix: The $(m \times n)$ matrix $\bar{A} = [\bar{a}_{ij}]$ is the conjugate or *complex conjugate* of the $(m \times n)$ matrix $A = [a_{ij}]$, where $\bar{a}_{ij}$ is the complex conjugate of a_{ij}. (See Section 3.2 for the properties.)

Conjugate transpose matrix: The conjugate transpose of the $(m \times n)$ matrix $A = [a_{ij}]$ is the $(n \times m)$ matrix $A^H \equiv \bar{A}' = [\bar{a}_{ij}]'$. (See Section 3.3 for the properties and further details.)

Conjunctive matrices: The $(m \times m)$ matrices A and B are said to be conjunctive if a unitary $(m \times m)$ matrix U exists such that $A = U^H BU$. (See Chapter 6 for some conjunctive matrices.)

Controllability matrix: For an $(m \times m)$ matrix A and an $(m \times n)$ matrix B the corresponding $(m \times mn)$ controllability matrix is defined as

$$C = [B : AB : \cdots : A^{m-1}B].$$

Convergent matrix: An $(m \times m)$ square matrix A is convergent or *stable* if $A^n \longrightarrow O_{m\times m}$ for $n \longrightarrow \infty$. (See Section 9.3 for the properties.)

Cosine of a matrix: The cosine of an $(m \times m)$ square matrix A is defined as

$$\cos(A) \equiv \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} A^{2n}.$$

(See Section 11.1 for more details.)

D-inverse: See *Drazin inverse.*

Decomposable matrix: An $(m \times m)$ matrix A is said to be decomposable or reducible if there exists a permutation matrix P such that

$$P'AP = \begin{bmatrix} A_{11} & O_{k\times(m-k)} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $(k \times k)$, A_{22} is $((m-k)\times(m-k))$ and A_{21} is $((m-k)\times k)$.

Defective matrix: An $(m \times m)$ matrix A is called defective if it has fewer than m linearly independent eigenvectors.

Deficient matrix: Alternative expression for *defective matrix*.

Definite matrix: A matrix is definite if it is positive definite or negative definite. (See also *positive definite matrix* and *negative definite matrix* and refer to Section 9.12 for the properties.)

Degree of a polynomial matrix: The degree of an $(m \times n)$ polynomial matrix $P(x) = [p_{ij}(x)]$ with typical elements $p_{ij}(x) = p_{ij,0} + p_{ij,1}x + \cdots + p_{ij,r_{ij}}x^{r_{ij}}$, is $\max_{i,j} r_{ij}$, where r_{ij} is the degree of $p_{ij}(x)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Equivalently, the degree of $P(x) = P_0 + P_1x + \cdots + P_rx^r$, is r if $P_r \neq 0$. (See Section 11.3 for details on polynomial matrices.)

Dependent vectors: See *linearly dependent vectors*.

Derivative matrix: Let $A(x) = [a_{ij}(x)]$ be an $(m \times n)$ matrix whose elements are differentiable functions $a_{ij} : \mathbb{R} \to \mathbb{R}$. The matrix of first order derivatives

$$\frac{\mathrm{d}A}{\mathrm{d}x} = \left[\frac{\mathrm{d}a_{ij}}{\mathrm{d}x}\right]$$

is called the derivative matrix of $A(x)$. (See Sections 10.1 and 10.9 for details.)

Derived norm: If $< \cdot, \cdot >$ is an inner product for $(m \times n)$ matrices, $\|A\| \equiv \, < A, A >^{1/2}$ is a norm for $(m \times n)$ matrices. $\|\cdot\|$ is said to be a norm derived from the inner product $< \cdot, \cdot >$. (See Chapter 8 for details.)

Derogatory matrix: An $(m \times m)$ matrix A is derogatory if a polynomial $p(x) = p_0 + p_1x + \cdots + p_rx^r$ of degree $r < m$ exists such that

$$p(A) = p_0I_m + p_1A + \cdots + p_rA^r = 0.$$

(See Section 11.1.)

Determinant of a matrix: The determinant of the $(m \times m)$ matrix $A = [a_{ij}]$ is defined as

$$\det A = \det(A) \equiv \sum (-1)^p a_{1i_1} a_{2i_2} \cdots a_{mi_m},$$

where the sum is taken over all products consisting of precisely one element from each row and each column of A multiplied by -1 or 1, if the permutation $i_1, \ldots, i_m$ is odd or even, respectively. (See Section 4.2 for the properties.)

Diagonal dominant matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is diagonal dominant if

$$|a_{ii}|_{\text{abs}} > \sum_{\substack{j=1 \\ j \neq i}}^{m} |a_{ij}|_{\text{abs}} \quad \text{for} \quad i = 1, 2, \ldots, m.$$

For example, the matrix

$$\begin{bmatrix} -5 & 1 & 2 \\ 1 & 3 & 1 \\ 5 & -5 & 11 \end{bmatrix}$$

is diagonal dominant.

Diagonal matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $i \neq j$ is a diagonal matrix. (See Section 9.4 for the properties.)

Diagonal of a matrix: A diagonal of an $(m \times m)$ matrix $A = [a_{ij}]$ consists of all elements a_{ij} for which $i - j$ equals a given integer.

Diagonal similar matrix: An $(m \times m)$ matrix A is diagonal similar if it is similar to a diagonal matrix, that is, if there exists a nonsingular $(m \times m)$ matrix P such that PAP^{-1} is diagonal. (See also *diagonalizable matrix* and Chapter 6.)

Diagonalizable matrix: An $(m \times m)$ matrix A is diagonalizable if a nonsingular $(m \times m)$ matrix P exists such that PAP^{-1} is a diagonal matrix. (See Chapter 6 for conditions.)

Dimension of a matrix: $(m \times n)$ is the dimension or *order* of a matrix with m rows and n columns.

Direct product: The direct product or *Kronecker product* or *tensor product* of two matrices $A = [a_{ij}]$ $(m \times n)$ and $B = [b_{ij}]$ $(p \times q)$ is

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \quad (mp \times nq).$$

(See Section 2.4 for the rules.)

Direct sum: The direct sum of two matrices $A = [a_{ij}]$ $(m \times m)$ and $B = [b_{ij}]$ $(n \times n)$ is

$$A \oplus B \equiv \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad ((m+n) \times (m+n)).$$

(See Section 2.6 for rules.)

Dominating nonnegative matrix: Given an $(m \times m)$ matrix $B = [b_{ij}]$, a nonnegative $(m \times m)$ matrix $A = [a_{ij}]$ is said to dominate B if $|b_{ij}|_{\mathrm{abs}} \leq a_{ij}$, $i, j = 1, \ldots, m$, that is, $|B|_{\mathrm{abs}} \leq A$. (See Section 9.9 for further details.)

Dominating principal diagonal: An $(m \times m)$ matrix $A = [a_{ij}]$ has a dominating principal diagonal if $d_i \in \mathbb{R}, d_i > 0$, $i = 1, \ldots, m$, exist such that

$$d_j |a_{jj}|_{\mathrm{abs}} > \sum_{\substack{i=1 \\ i \neq j}}^{m} d_i |a_{ij}|_{\mathrm{abs}}, \quad j = 1, \ldots, m.$$

For instance,

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

has a dominating principal diagonal.

Doubly centered matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called doubly centered if

$$\sum_{i=1}^{m} a_{ij} = 0 \quad \text{for} \quad j = 1, \ldots, m$$

and

$$\sum_{j=1}^{m} a_{ij} = 0 \quad \text{for} \quad i = 1, \ldots, m.$$

For example, the matrix

$$\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$$

is doubly centered.

Doubly quasi-stochastic matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is called doubly quasi-stochastic if $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, \ldots, m$ and $\sum_{i=1}^{m} a_{ij} = 1$ for $j = 1, \ldots, m$. (See also *doubly stochastic matrix* and refer to Section 9.9 for further details.)

Doubly stochastic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be doubly stochastic if $A \geq 0$, that is, $a_{ij} \geq 0$, $i, j = 1, \ldots, m$ and $\sum_{j=1}^{m} a_{ij} = 1$, $i = 1, \ldots, m$ and $\sum_{i=1}^{m} a_{ij} = 1$, $j = 1, \ldots, m$. For example, the $(m \times m)$ matrix

$$\begin{bmatrix} \frac{1}{m} & \cdots & \frac{1}{m} \\ \vdots & & \vdots \\ \frac{1}{m} & \cdots & \frac{1}{m} \end{bmatrix}$$

is doubly stochastic. (See Section 9.9 for further details.)

Drazin inverse: Let A be an $(m \times m)$ matrix with index r, that is, r is the smallest number such that $\text{rk}(A^r) = \text{rk}(A^{r+1})$. The Drazin inverse or *D-inverse* A^D of A is the unique solution of the set of equations

$$\begin{aligned} A^{r+1} A^D &= A^r \\ A^D A A^D &= A^D \\ A A^D &= A^D A. \end{aligned}$$

(See also *generalized inverse.*)

Duplication matrix: An $(m^2 \times \frac{1}{2}m(m+1))$ matrix D_m satisfying $\text{vec}(A) = D_m \text{vech}(A)$, for any symmetric $(m \times m)$ matrix A, is called a duplication matrix. For example,

$$D_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is a duplication matrix. (See Section 9.5 for the properties.)

Dyadic product: For vectors

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \quad (m \times 1), \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad (n \times 1)$$

the product

$$xy' \equiv \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{bmatrix}$$

is called the dyadic product.

Echelon form: An $(m \times n)$ matrix $A = [a_{ij}]$ is in echelon form if for any row i, either $a_{ij} = 0$, $j = 1, \ldots, n$, or there exists a $k \in \{1, \ldots, n\}$ such that $a_{ik} = 1$ and $a_{ij} = 0$ for $j < k$ and $a_{lk} = 0$ for $l \neq i$. For example,

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & a_{25} \\ 0 & 0 & 0 & 1 & a_{35} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is in echelon form. (See also Section 6.1.4.)

Effective inverse matrix: A generalized inverse matrix is sometimes called an effective inverse matrix. (See *generalized inverse matrix*.)

Eigenvalue of a matrix: The roots of the characteristic polynomial of an $(m \times m)$ matrix A, that is, the roots of $p(\lambda) = \det(\lambda I_m - A)$, are the eigenvalues of A. (See also *characteristic value* or *latent root* and refer to Chapter 5 for details.)

Eigenvalue of a matrix in the metric of another matrix: Given two $(m \times m)$ matrices A and B, the roots of the polynomial $p(\lambda) = \det(\lambda B - A)$ are called eigenvalues or characteristic roots of A in the metric of B. (See Section 5.1 for further details.)

Eigenvalue of a polynomial matrix: The eigenvalues or latent roots of a polynomial matrix $P(x)$ are the roots of the polynomial $\det P(x)$.

Eigenvector matrix: Let A be an $(m \times m)$ matrix with linearly independent eigenvectors $v_1, \ldots, v_m$ of length 1, that is, $v_i^H v_i = 1$ for $i = 1, \ldots, m$. Then $V = [v_1, \ldots, v_m]$ is called an eigenvector matrix or *modal matrix* of A.

Eigenvector of a matrix: An $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda v$ for an $(m \times m)$ matrix A and a complex number λ is an eigenvector or *characteristic vector* of A corresponding to or associated with the eigenvalue value λ. (See Chapter 5 for details.)

Eigenvector of a matrix in the metric of another matrix: Given two $(m \times m)$ matrices A and B, an $(m \times 1)$ vector $v \neq 0$ satisfying $Av = \lambda Bv$ for some complex number λ, is called an eigenvector or characteristic vector of A in the metric of B. (See Section 5.1 for further details.)

Elementary divisors of a polynomial matrix: Let λ be a root of an *invariant factor* of a polynomial matrix $P(x)$. Then the linear polynomial $p(x) = x - \lambda$ is an elementary divisor of $P(x)$. (See Section 11.3 for more details.)

Elementary matrix: An $(m \times m)$ matrix is called elementary if it is obtained by applying a single *elementary matrix operation* to the $(m \times m)$ identity matrix I_m.

Elementary matrix operations: The following modifications of a matrix are called elementary operations:

(i) interchanging two rows or two columns,

(ii) multiplying any row or column by a nonzero number,

(iii) adding a multiple of one row to another row,

(iv) adding a multiple of one column to another column.

Elementary operations for polynomial matrices: The following modifications of a polynomial matrix are called elementary operations:

(i) interchanging two rows or two columns,

(ii) multiplying any row or column by a nonzero number,

(iii) adding to one row another row multiplied by an arbitrary polynomial,

(iv) adding to one column another column multiplied by an arbitrary polynomial.

(See Section 11.3 for details on polynomial matrices.)

Elementary polynomial matrix: An $(m \times m)$ polynomial matrix is elementary if it may be obtained by applying a single *elementary* *polynomial matrix operation* to I_m. (See Section 11.3 for details on polynomial matrices.)

Elementwise product: See *Hadamard product*.

Elimination matrix: A $(\frac{1}{2}m(m+1) \times m^2)$ matrix L_m satisfying $\text{vech}(A) = L_m \text{vec}(A)$, for any $(m \times m)$ matrix A, is called an elimination matrix. For example,

$$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is an elimination matrix. (See Section 9.6 for the properties.)

EP-matrix: An $(m \times m)$ matrix A is said to be an EP-matrix if it commutes with its Moore–Penrose inverse A^+, that is, $AA^+ = A^+A$. (See Section 3.6.2 for details on the Moore–Penrose inverse.)

Equality of matrices: Two $(m \times n)$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are equal if $a_{ij} = b_{ij}$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$.

Equivalent matrices: The $(m \times n)$ matrices A and B are equivalent if nonsingular matrices C $(m \times m)$ and D $(n \times n)$ exist such that $B = CAD$. (For examples see Chapter 6.)

Equivalent polynomial matrices: Two polynomial matrices $P(x)$ and $Q(x)$ are said to be equivalent if *unimodular polynomial matrices* $U(x)$ and $V(x)$ exist such that $P(x) = U(x)Q(x)V(x)$. (See Section 11.3.)

Ergodic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be ergodic if it is *stochastic* (that is, $a_{ij} \geq 0$, $i, j = 1, \ldots, m$, and $\sum_{j=1}^m a_{ij} = 1$, $i = 1, \ldots, m$) and $\lambda = 1$ is the only eigenvalue with modulus 1. Moreover, if the multiplicity of $\lambda = 1$ is k, then there exist k linearly independent eigenvectors associated with λ. (See also Section 9.9.)

Exchange matrix: The $(m \times m)$ matrix

$$A = [a_{ij}] = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

with

$$a_{ij} = \begin{cases} 1 & \text{for } i+j = m+1 \\ 0 & \text{otherwise} \end{cases}$$

is called an exchange matrix, *reverse unit matrix* or *flip matrix*.

Exponential function of a matrix: The exponential function of an $(m \times m)$ matrix A is defined as

$$\exp(A) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} A^n .$$

Alternatively, this expression is called the *exponential matrix* of A. (See Section 11.1.)

Exponential matrix: See *exponential function of a matrix*.

Flip matrix: An *exchange matrix* is sometimes called a flip matrix.

Fourier matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a Fourier matrix if

$$a_{ij} = \frac{1}{\sqrt{m}} \omega^{(i-1)(j-1)}$$

with $\omega = \exp(-2\pi i/m)$. For example, for $m = 3$,

$$\frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega^4 \end{bmatrix}$$

is a Fourier matrix.

Frobenius matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a Frobenius matrix if one column only contains elements which are neither zero nor one and an identity matrix may be obtained by eliminating one row and one column. For instance,

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ a_{m1} & 0 & \cdots & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & \cdots & a_{1m} \\ 0 & 1 & & a_{2m} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} \end{bmatrix}$$

are Frobenius matrices.

g-circulant matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{m-1} & a_m \\ a_{m-g+1} & a_{m-g+2} & a_{m-g+3} & \cdots & a_{m-g-1} & a_{m-g} \\ a_{m-2g+1} & a_{m-2g+2} & a_{m-2g+3} & \cdots & \cdots & a_{m-2g} \\ \vdots & \vdots & \vdots & & & \vdots \\ a_{g+1} & a_{g+2} & a_{g+3} & \cdots & \cdots & a_g \end{bmatrix},$$

where the last g elements of a row are the first elements of the next row, is a g-circulant matrix. For example, a (4×4) 2-circulant matrix is of the form

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_3 & a_4 & a_1 & a_2 \\ a_1 & a_2 & a_3 & a_4 \\ a_3 & a_4 & a_1 & a_2 \end{bmatrix}.$$

(See also *circulant matrix.*)

g-inverse matrix: Short for *generalized inverse matrix.*

Generalized characteristic root or **value:** An *eigenvalue of a matrix in the metric of another matrix* is sometimes called a generalized characteristic root or generalized characteristic value.

Generalized eigenvalue: An *eigenvalue of a matrix in the metric of another matrix* is sometimes called a generalized eigenvalue.

Generalized eigenvector of a matrix: Let A be an $(m \times m)$ matrix and $k \in \mathbb{N}$, $k > 1$. An $(m \times 1)$ vector x is a generalized eigenvector of grade k or a *principal vector of grade k* corresponding to an eigenvalue λ if

$$(\lambda I_m - A)^k x = 0 \quad \text{and} \quad (\lambda I_m - A)^{k-1} x \neq 0.$$

Alternatively, an *eigenvector of a matrix in the metric of another matrix* is sometimes called a generalized eigenvector.

Generalized inverse matrix: An $(n \times m)$ matrix A^- is a generalized inverse of the $(m \times n)$ matrix A if it satisfies $AA^-A = A$. (See Section 3.6 for the properties.)

Generalized permutation matrix: An $(m \times m)$ matrix with at most one nonzero element in each row and each column and zeros elsewhere is said to be a generalized permutation matrix. For example,

$$\begin{bmatrix} 0 & a & 0 \\ b & 0 & 0 \\ 0 & 0 & c \end{bmatrix}$$

is a generalized permutation matrix. (See also *permutation matrix.*)

Geometric matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called a geometric matrix if $a_{11} = 1$ and $a_{ij} = \lambda^{i+j-2}$ for $i + j > 2$. Here λ is a complex number. For example, for $m = 3$,

$$A = \begin{bmatrix} 1 & \lambda & \lambda^2 \\ \lambda & \lambda^2 & \lambda^3 \\ \lambda^2 & \lambda^3 & \lambda^4 \end{bmatrix}$$

is a geometric matrix.

Geometric multiplicity of an eigenvalue: The geometric multiplicity of an eigenvalue λ of an $(m \times m)$ matrix A is the number of blocks of the form

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & \ddots & & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \dots & & \lambda \end{bmatrix}$$

with λ on the principal diagonal, in the *Jordan decomposition* of A. (See Section 6.1.)

Givens matrix: For $r, s \in \mathbb{N}$, $r < s \leq m$, an $(m \times m)$ matrix $G_{rs} = [g_{ij}]$ is said to be a Givens matrix or *rotation matrix* if $g_{rr} = g_{ss} = \cos\theta$, $g_{ii} = 1$ for $i = 1, \dots, m$, $i \neq r, s$, $g_{sr} = -\sin\theta$, $g_{rs} = \sin\theta$ and all other elements are zero. For example,

$$G_{13} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \quad G_{12} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

are Givens matrices.

Gradient matrix: The $(m \times n)$ matrix of first order partial derivatives of the vector valued function $y : U \subset \mathbb{R}^m \to \mathbb{R}^n$,

$$\frac{\partial y'}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_m} & \dots & \frac{\partial y_n}{\partial x_m} \end{bmatrix},$$

is the gradient matrix of $y(x) = (y_1(x), \dots, y_n(x))'$, where $x = (x_1, \dots, x_m)'$ is an $(m \times 1)$ vector. (See Chapter 10 for more details.)

Gradient vector: The $(m \times 1)$ vector of first order partial derivatives of the scalar valued function $f : U \subset \mathbb{R}^m \to \mathbb{R}$,

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{bmatrix},$$

is the gradient vector of $f(x)$, where $x = (x_1, \dots, x_m)'$ is an $(m \times 1)$ vector. (See Chapter 10 for more details.)

Gram matrix: For an $(m \times m)$ matrix A, $G = A^H A$ is the corresponding Gram matrix. (See also *Gramian matrix*.)

Gramian matrix: An $(m \times m)$ matrix A is a Gramian or positive semidefinite matrix if it is *Hermitian* and $x^H Ax \geq 0$ for all $(m \times 1)$ vectors x. (See also *positive semidefinite matrix* and refer to Section 9.12 for the properties.)

Greatest common left divisor: A polynomial matrix $G(x)$ is a greatest common left divisor of the polynomial matrices $P(x)$ and $Q(x)$ if $G(x)$ is a *common left divisor* of both $P(x)$ and $Q(x)$ (that is, $P(x) = G(x)P_1(x)$ and $Q(x) = G(x)Q_1(x)$ for suitable matrix polynomials $P_1(x)$ and $Q_1(x)$) and for any other common left divisor $D(x)$ there exists a polynomial matrix $B(x)$ such that $G(x) = D(x)B(x)$. (See Section 11.3.)

Greatest common right divisor: A polynomial matrix $G(x)$ is a greatest common right divisor of the polynomial matrices $P(x)$ and $Q(x)$ if $G(x)$ is a *common right divisor* of both $P(x)$ and $Q(x)$ (that is, $P(x) = P_1(x)G(x)$ and $Q(x) = Q_1(x)G(x)$ for suitable matrix polynomials $P_1(x)$ and $Q_1(x)$) and for any other common right divisor $D(x)$ there exists a polynomial matrix $B(x)$ such that $G(x) = B(x)D(x)$. (See Section 11.3.)

Group inverse of a matrix: If the $(m \times m)$ matrix A satisfies $\mathrm{rk}(A) = \mathrm{rk}(A^2)$, the matrix $A^\#$ satisfying $A^2 A^\# = A$, $A^\# A A^\# = A^\#$ and $AA^\# = A^\# A$ is called the group inverse of A. (See also *generalized inverse matrix*.)

Hadamard matrix: A Hadamard matrix $H_k, k = 0, 1, \dots$, is a $(2^k \times 2^k)$ matrix which has only elements 1 and -1 and which is obtained recursively as

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes H_{k-1}, \quad k = 1, 2, \dots,$$

starting with $H_0 = 1$. For instance,

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Alternatively, a matrix with dominating principal diagonal is sometimes called a Hadamard matrix. (See *dominating principal diagonal matrix.*)

Hadamard product: The Hadamard product or *Schur product* or *elementwise product* of the two matrices $A = [a_{ij}]$ $(m \times n)$ and $B = [b_{ij}]$ $(m \times n)$ is defined as

$$A \odot B \equiv [a_{ij} b_{ij}] \quad (m \times n).$$

(See Section 2.5 for the rules.)

Half-vectorization operator: For a symmetric $(m \times m)$ matrix $A = [a_{ij}]$ the half-vectorization operator vech stacks all columns from the principal diagonal downwards, that is,

$$\operatorname{vech} A = \operatorname{vech}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{33} \\ \vdots \\ a_{mm} \end{bmatrix} \quad (\tfrac{1}{2}m(m+1) \times 1).$$

(See Section 7.3 for the rules.)

Hamiltonian matrix: A $(2m \times 2m)$ matrix H is a Hamiltonian matrix if

$$H = \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} H' \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix}.$$

For instance, for matrices A $(m \times m)$, B $(m \times n)$ and C $(r \times m)$,

$$H = \begin{bmatrix} A & -BB' \\ -C'C & A' \end{bmatrix}$$

is a Hamiltonian matrix.

Hankel matrix: For $\alpha_1, \ldots, \alpha_{2m-1} \in \mathbb{C}$, an $(m \times m)$ matrix of the form

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_m \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha_{m+1} \\ \alpha_3 & \alpha_4 & \alpha_5 & & \alpha_{m+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \alpha_m & \alpha_{m+1} & \alpha_{m+2} & \cdots & \alpha_{2m-1} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = \alpha_{i+j-1}$ is a Hankel matrix.

Harmonic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called a harmonic matrix if $a_{ij} = \frac{1}{a+(i+j-2)d}$ for complex numbers a and d such that the denominators are nonzero. For example, for $m = 3$,

$$A = \begin{bmatrix} \frac{1}{a} & \frac{1}{a+d} & \frac{1}{a+2d} \\ \frac{1}{a+d} & \frac{1}{a+2d} & \frac{1}{a+3d} \\ \frac{1}{a+2d} & \frac{1}{a+3d} & \frac{1}{a+4d} \end{bmatrix}$$

is a harmonic matrix. (See *Hilbert matrix* for a special case.)

Helmert matrix: An $(m \times m)$ matrix of the form

$$\begin{bmatrix} m^{-\frac{1}{2}} & m^{-\frac{1}{2}} & m^{-\frac{1}{2}} & \cdots & m^{-\frac{1}{2}} & m^{-\frac{1}{2}} \\ 2^{-\frac{1}{2}} & -2^{-\frac{1}{2}} & 0 & \cdots & 0 & 0 \\ 6^{-\frac{1}{2}} & 6^{-\frac{1}{2}} & -2 \cdot 6^{-\frac{1}{2}} & \cdots & 0 & 0 \\ \vdots & & & & & \\ \frac{1}{[m(m-1)]^{\frac{1}{2}}} & \frac{1}{[m(m-1)]^{\frac{1}{2}}} & \cdots & \cdots & \frac{1}{[m(m-1)]^{\frac{1}{2}}} & \frac{-(m-1)}{[m(m-1)]^{\frac{1}{2}}} \end{bmatrix}$$

is called a Helmert matrix. For instance, for $m = 3$,

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

is a Helmert matrix.

Hermite canonical form: An $(m \times m)$ matrix $A = [a_{ij}]$ is in Hermite canonical form if

$$a_{ij} = \begin{cases} 0 \text{ or } 1 & \text{if } i = j \\ 0 & \text{if } i > j \\ 0 & \text{for } j = 1, \ldots, m, i \neq j, \text{ if } a_{ii} = 0 \\ 0 & \text{for } i = 1, \ldots, m, i \neq j, \text{ if } a_{jj} = 1 \end{cases}$$

For example,

$$\begin{bmatrix} 1 & a_{12} & 0 & a_{14} & a_{15} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & a_{34} & a_{35} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is in Hermite canonical form. (See also Section 6.1.4.)

Hermitian adjoint: Alternative expression for *conjugate transpose*.

Hermitian form: Given a *Hermitian* $(m \times m)$ matrix A, the function $Q : \mathbb{C}^m \rightarrow \mathbb{R}$ defined by $Q(x) = x^H Ax$ is called a Hermitian form.

Hermitian matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ with $a_{ij} = \bar{a}_{ji}$ is a Hermitian matrix. Here $\bar{a}_{ji}$ denotes the complex conjugate of a_{ji}. In other words, A is Hermitian if $\bar{A}' \equiv A^H = A$. (See Section 9.7 for the properties and more details.)

Hessenberg matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a Hessenberg matrix or upper Hessenberg matrix if $a_{ij} = 0$ for $i > j + 1$, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,m-1} & a_{1,m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,m-1} & a_{2,m} \\ 0 & a_{32} & a_{33} & & a_{3,m-1} & a_{3,m} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & a_{m-1,m-1} & a_{m-1,m} \\ 0 & 0 & 0 & \cdots & a_{m,m-1} & a_{m,m} \end{bmatrix}.$$

(See also Section 6.2.3.)

Hessian matrix: The $(m \times m)$ matrix of second order partial derivatives

$$\frac{\partial^2 f}{\partial x \partial x'} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_m \partial x_m} \end{bmatrix}$$

is the Hessian matrix of the scalar valued function $f(x)$, where $x = (x_1, \ldots, x_m)'$ is an $(m \times 1)$ vector. (See Chapter 10 for more details.)

Hilbert matrix: The $(m \times m)$ matrix $A = [a_{ij}]$ is a Hilbert matrix if $a_{ij} = 1/(i + j - 1)$, that is, if

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{m} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m} & \frac{1}{m+1} & \cdots & \frac{1}{2m-1} \end{bmatrix}.$$

Householder transformation: The Householder transformation of a real $(m \times 1)$ vector $x \neq 0$ is the $(m \times m)$ matrix

$$H = I_m - \frac{2xx'}{x'x}.$$

The Householder transformation of a complex $(m \times 1)$ vector $x \neq 0$ is

$$H = I_m - \frac{2xx^H}{x^H x}.$$

Hurwitz matrix: The Hurwitz matrix corresponding to the polynomial $p(x) = p_0 + p_1 x + \cdots + p_m x^m$ is

$$H = \begin{bmatrix} p_{m-1} & p_{m-3} & p_{m-5} & \cdots & p_{-m+1} \\ p_m & p_{m-2} & p_{m-4} & \cdots & p_{-m+2} \\ 0 & p_{m-1} & p_{m-3} & \cdots & p_{-m+3} \\ 0 & p_m & p_{m-2} & \cdots & p_{-m+4} \\ 0 & 0 & p_{m-1} & \cdots & p_{-m+5} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & & p_0 \end{bmatrix} \quad (m \times m)$$

where $p_i = 0$ for $i < 0$. For example,

$$H = \begin{bmatrix} p_3 & p_1 & 0 & 0 \\ p_4 & p_2 & p_0 & 0 \\ 0 & p_3 & p_1 & 0 \\ 0 & p_4 & p_2 & p_0 \end{bmatrix}$$

is the Hurwitz matrix corresponding to $p(x) = p_0 + p_1 x + p_2 x^2 + p_3 x^3 + p_4 x^4$. (See also Section 11.1.)

Idempotent matrix: An $(m \times m)$ matrix A is idempotent if $A^2 = A$. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is idempotent. (See Section 9.8 for the properties.)

Identity matrix: An $(m \times m)$ matrix

$$I_m = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 1$ for $i = 1, \ldots, m$ and $a_{ij} = 0$ for $i \neq j$ is an identity or unit matrix. (See also *unit matrix*.)

Imaginary matrix: A $(m \times n)$ is an imaginary matrix if it can be represented as $A = \mathrm{i}B$, where B is a real matrix and $\mathrm{i} = \sqrt{-1}$. that is, all elements of A are imaginary numbers or zero.

Imprimitive matrix: An *irreducible* $(m \times m)$ matrix A is said to be imprimitive if it has more than one eigenvalue whose modulus is equal to the *spectral radius* of A.

Indefinite matrix: A *Hermitian* $(m \times m)$ matrix A is indefinite if it is not definite. In other words, A is indefinite if $(m \times 1)$ vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ exist with $\boldsymbol{x}_1^H A \boldsymbol{x}_1 > 0$ and $\boldsymbol{x}_2^H A \boldsymbol{x}_2 < 0$.

Independent vectors: See *linearly independent vectors.*

Index of a square matrix: The smallest nonnegative integer i such that $\mathrm{rk}(A^i) = \mathrm{rk}(A^{i+1})$ is the index of the square matrix A.

Induced matrix norm: Given a *vector norm* $\|\cdot\|$ for $(m \times 1)$ vectors,

$$\|A\|_{lub} \equiv \sup\left\{ \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \ : \ \boldsymbol{x} \ (m \times 1), \ \boldsymbol{x} \neq 0 \right\}$$

is the matrix norm for $(m \times m)$ matrices induced by $\|\cdot\|$. Alternatively it is called the sup norm or *operator norm* or lub (*least upper bound*) *norm*. (For more details on norms see Chapter 8.)

Inertia of a square matrix: The inertia of an $(m \times m)$ matrix A is the triple of integers (a, b, c), where a, b, c are the numbers of eigenvalues (counted with their *algebraic multiplicities*) with positive real part, negative real part and zero real part, respectively. (See also Section 5.1.)

Inner product: A function $< \cdot, \cdot >$ attaching a complex number $< A, B >$ to any two $(m \times n)$ matrices A, B is called an inner product if for any arbitrary $(m \times n)$ matrices A, B, C and $c \in \mathbb{C}$ the following conditions are satisfied:

(i) $< A, A > \ > 0 \quad \text{if} \quad A \neq 0,$

(ii) $< A+B, C > \; = \; < A, C > + < B, C >$,

(iii) $< cA, B > \; = c < A, B >$,

(iv) $< A, B > \; = \; \overline{< B, A >}$.

(See Chapter 8 for more details.)

Inner product over the real numbers: A function $< \cdot, \cdot >$ is an inner product over the field of real numbers if it attaches a real number to a pair of real matrices in such a way that for any real $(m \times n)$ matrices A, B, C and $c \in \mathbb{R}$ the following conditions are satisfied:

(i) $< A, A > \; > 0 \quad \text{if} \quad A \neq 0$,

(ii) $< A+B, C > \; = \; < A, C > + < B, C >$,

(iii) $< cA, B > \; = \; c < A, B >$,

(iv) $< A, B > \; = \; < B, A >$.

(See Chapter 8 for further details.)

Integer matrix: $A = [a_{ij}]$ $(m \times n)$ is an integer matrix if all elements are integers, that is, $a_{ij} \in \mathbb{Z}$.

Integral matrix: Let $A(t) = [a_{ij}(t)]$ be an $(m \times n)$ matrix of integrable functions. Then $\int A \, \mathrm{d}t \equiv [\int a_{ij}(t) \mathrm{d}t]$ is the integral matrix of $A(t)$.

Invariant factors of a polynomial matrix: The invariant factors or invariant polynomials of an $(m \times m)$ polynomial matrix $P(x)$ of rank r are defined as

$$i_k(x) = \frac{d_k(x)}{d_{k-1}(x)}, \quad k = 1, \ldots, r,$$

where $d_0(x) \equiv 1$ and $d_k(x)$ is the monic greatest common divisor of all minors of order k of $P(x)$ for $k = 1, \ldots, r$. (See Section 11.1.1 for more details.)

Invariant polynomials of a polynomial matrix: See *invariant factors of a polynomial matrix*.

Inverse matrix: An $(m \times m)$ matrix A^{-1} is the inverse of the $(m \times m)$ matrix A if $AA^{-1} = A^{-1}A = I_m$. (See Section 3.5 for the properties and further details.)

Invertible matrix: An $(m \times m)$ matrix A is invertible or *nonsingular* or *regular* if $\det(A) \neq 0$ and thus A^{-1} exists. (See Section 3.5 for the properties.)

Invertible polynomial matrix: The $(m \times m)$ polynomial matrix $P(x)$ is said to be *unimodular* or invertible if $\det P(x) = \text{constant} \neq 0$, that is, $\det P(x)$ is a constant function. For instance, for $a \in \mathbb{R}$,

$$P(x) = I_2 + \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} x$$

has determinant $\det P(x) = 1$ and, hence, $P(x)$ is unimodular. (See Section 11.3 for more details.)

Involutory matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called involutory or *unipotent* if $A^2 = I_m$. For example,

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

is involutory.

Irreducible matrix: An $(m \times m)$ matrix A which is not reducible is called irreducible, that is, A is irreducible if there does not exist a *permutation matrix* P such that

$$P'AP = \begin{bmatrix} A_{11} & O_{k\times(m-k)} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $(k \times k)$, A_{22} is $((m-k)\times(m-k))$ and A_{21} is $((m-k)\times k)$.

Isometry for a norm: An $(m \times m)$ matrix A is called an isometry for a norm $\|\cdot\|$ for $(m \times 1)$ vectors if $\|Ax\| = \|x\|$ for all $(m \times 1)$ vectors x. (See Chapter 8.)

Jacobian matrix: The $(n \times m)$ matrix of first order partial derivatives

$$\frac{\partial y}{\partial x'} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}$$

is the Jacobian matrix of the $(n \times 1)$ vector function $y(x) = (y_1(x), \ldots, y_n(x))'$, where $x = (x_1, \ldots, x_m)'$ is an $(m \times 1)$ vector. (See Chapter 10.)

Jordan canonical form or **Jordan normal form:** A block diagonal matrix

$$\begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_p \end{bmatrix}$$

with

$$\Lambda_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix} \quad (m_i \times m_i)$$

$i = 1, \ldots, p$, is a Jordan canonical form or Jordan normal form or *Jordan matrix*. For example,

$$\Lambda = \begin{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} & & 0 \\ & [2] & \\ 0 & & \begin{bmatrix} 3.2 & 1 \\ 0 & 3.2 \end{bmatrix} \end{bmatrix}$$

is a Jordan canonical form with

$$\Lambda_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad \Lambda_2 = [2], \quad \Lambda_3 = \begin{bmatrix} 3.2 & 1 \\ 0 & 3.2 \end{bmatrix}.$$

(See Chapter 6 for more details.)

Jordan matrix: A *Jordan canonical form* is sometimes called a Jordan matrix.

Kronecker matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ with

$$a_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ c & \text{for } i = j = k \\ 1 & \text{for } i = j \neq k \end{cases}$$

for some $k \in \{1, \ldots, m\}$, is a Kronecker matrix. For instance,

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & c & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

is a Kronecker matrix.

Kronecker product: The Kronecker product of two matrices $A = [a_{ij}]$ $(m \times n)$ and $B = [b_{ij}]$ $(p \times q)$ is defined as

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \quad (mp \times nq).$$

(See also *direct product* or *tensor product* and refer to Section 2.4 for the rules.)

Kronecker sum: Let A and B be $(m \times m)$ and $(n \times n)$ matrices, respectively. The matrix $D = A \otimes I_n + I_m \otimes B$ is sometimes called Kronecker sum of A and B.

Latent root: The roots of the characteristic polynomial of an $(m \times m)$ matrix A, that is, the roots of $p(\lambda) = \det(\lambda I_m - A)$, are the latent roots of A. (See *characteristic value* or *eigenvalue*).

Latent vector of a polynomial matrix: Let $P(x)$ be a polynomial matrix with eigenvalue λ. A vector v such that $P(\lambda)v = 0$ is called a latent vector of $P(x)$. (See Section 11.3.)

Least upper bound norm: Given a norm $\|\cdot\|$ for $(m \times 1)$ vectors,

$$\|A\|_{lub} \equiv \sup\left\{\frac{\|Ax\|}{\|x\|} : x\ (m \times 1),\ x \neq 0\right\}$$

is a norm for $(m \times m)$ matrices called the least upper bound norm induced by $\|\cdot\|$. Alternatively it is called the sup norm or *operator norm* or *induced norm*. (See Chapter 8 for more details.)

Left coprime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be *relatively left prime* or left coprime if their *greatest common left divisors* are *unimodular*. (See Section 11.3.)

Left divisibility of a polynomial matrix: A polynomial matrix $P(x)$ is divisible on the left by a polynomial matrix $Q(x)$ if $Q(x)$ is a left divisor of $P(x)$, that is, if for some polynomial matrix $T(x)$ with degree greater than 0, $P(x) = Q(x)T(x)$. (See Section 11.3.)

Left divisor of a polynomial matrix: A polynomial matrix $Q(x)$ is a left divisor of the polynomial matrix $P(x)$ if a polynomial matrix $T(x)$ exists such that $P(x) = Q(x)T(x)$. (See Section 11.3.)

Left eigenvector: A $(1 \times m)$ vector $v^H \neq 0$ satisfying $v^H A = \lambda v^H$ for an $(m \times m)$ matrix A and a complex number λ is a left eigenvector of A corresponding to or associated with the eigenvalue λ.

Left inverse matrix: Let A be an $(m \times n)$ matrix. An $(n \times m)$ matrix B satisfying $BA = I_n$ is called a left inverse of A. (See also *generalized inverse matrix* and refer to Section 3.6 for further details.)

Left multiple of a polynomial matrix: The $(m \times n)$ polynomial matrix $P(x)$ is a left multiple of an $(h \times n)$ polynomial matrix $T(x)$ if there exists an $(m \times h)$ polynomial matrix $Q(x)$ such that $P(x) = Q(x)T(x)$. (See Section 11.3.)

Left quotient of polynomial matrices: Let $P(x) = P_0 + P_1 x + \cdots + P_r x^r$ $(m \times n)$, $P_r \neq 0$, and $Q(x) = Q_0 + Q_1 x + \cdots + Q_s x^s$ $(m \times h)$ regular, $s \leq r$. An $(h \times n)$ polynomial matrix $T(x)$ is a left quotient of $P(x)$ and $Q(x)$ if an $(m \times n)$ polynomial matrix $R(x)$ with degree less than s exists such that $P(x) = Q(x)T(x) + R(x)$. (See Section 11.3.)

Left remainder: Let $P(x)$ and $Q(x)$ be polynomial matrices of degrees r and s, respectively, $s \leq r$. If $T(x)$ is a *left quotient* of $P(x)$ and $Q(x)$ with $P(x) = Q(x)T(x) + R(x)$, then $R(x)$ is said to be a left remainder. (See Section 11.3.)

Length of a vector: The length or modulus of an $(m \times 1)$ vector $x = (x_1, \ldots, x_m)'$ is $\sqrt{x^H x} = (x_1\bar{x}_1 + \cdots + x_m\bar{x}_m)^{1/2}$. (See also Chapter 8.)

Leontief matrix: An $(m \times m)$ matrix of the form

$$\begin{bmatrix} 1 & -a_{12} & \cdots & -a_{1m} \\ -a_{21} & 1 & \cdots & -a_{2m} \\ \vdots & & \ddots & \vdots \\ -a_{m1} & -a_{m2} & \cdots & 1 \end{bmatrix}$$

with unit main diagonal and nonpositive elements elsewhere, is called a Leontief matrix.

Leslie matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 0 & a_1 & 0 & \cdots & 0 \\ 0 & 0 & a_2 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & & a_{m-1} \\ b_1 & b_2 & b_3 & \cdots & b_m \end{bmatrix}$$

where $a_i \neq 0$, $i = 1, \ldots, m-1$, is called a Leslie matrix.

Liapunov matrix equation: Let A be a real $(m \times m)$ matrix and X, Q be *real symmetric positive definite* matrices. The matrix equation

$$A'X + XA = -Q$$

is called a Liapunov (matrix) equation.

Linear combination of matrices: Let A_i, $i = 1, \ldots, k$, be $(m \times n)$ matrices and $c_i \in \mathbb{C}$, $i = 1, \ldots, k$. $c_1 A_1 + \cdots + c_k A_k$ is a linear combination of the matrices $A_1, \ldots, A_k$.

Linearly dependent vectors: The $(m \times 1)$ vectors $x_1, \ldots, x_n$ are linearly dependent if for some $i \in \{1, \ldots, n\}$ scalars $c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n$ exist such that $x_i = c_1 x_1 + \cdots + c_{i-1} x_{i-1} + c_{i+1} x_{i+1} + \cdots + c_n x_n$. (See also Section 4.3.)

Linearly independent vectors: The $(m \times 1)$ vectors $x_1, \ldots, x_n$ are linearly independent if they are not linearly dependent. In other words, $x_1, \ldots, x_n$ are linearly independent if, for $c_1, \ldots, c_n \in \mathbb{C}$, $c_1 x_1 + \cdots + c_n x_n = 0$ implies that $c_1 = \cdots = c_n = 0$. (See also Section 4.3.)

Loewner matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is called a Loewner matrix if its elements are of the form

$$a_{ij} = \frac{c_i - d_j}{f_i - h_j}, \quad i, j = 1, \ldots, m,$$

where $c_1, \ldots, c_m, d_1, \ldots, d_m, f_1, \ldots, f_m, h_1, \ldots, h_m \in \mathbb{C}$ and f_i and h_j are distinct.

Lower block triangular matrix: An $(m \times n)$ matrix

$$\begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A_{p1} & \cdots & \cdots & A_{pp} \end{bmatrix} = [A_{ij}],$$

where the A_{ij} are $(m_i \times n_j)$ matrices with $\sum_{i=1}^{p} m_i = m$, $\sum_{j=1}^{p} n_j = n$ and $A_{ij} = 0$ for $j > i$, is lower block triangular.

Lower Hessenberg matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a lower Hessenberg matrix if $a_{ij} = 0$ for $j > i + 1$, that is, if

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & & 0 & 0 \\ a_{31} & a_{32} & a_{33} & & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{m-1,1} & a_{m-1,2} & & \ddots & a_{m-1,m-1} & a_{m-1,m} \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,m-1} & a_{m,m} \end{bmatrix}.$$

Lower triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ a_{m1} & \cdots & \cdots & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $j > i$, is a lower triangular matrix. (See Section 9.14 for the properties.)

M-matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is said to be an M-matrix or *Minkowski matrix* if A is nonsingular, $a_{ij} \leq 0$ for $i \neq j$ and $A^{-1} > 0$.

Main diagonal of a matrix: See *principal diagonal.*

Matrix: An $(m \times n)$ matrix A is an array of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = [a_{ij}].$$

Alternative notations:

$$A\ (m \times n), \qquad \underset{(m \times n)}{A}, \qquad A = [a_{ij}]\ (m \times n).$$

Matrix function: A function $f : \mathbb{C}^{m \times n} \to \mathbb{C}$ (or $\mathbb{R}^{m \times n} \to \mathbb{R}$), $A \mapsto f(A)$ is called a matrix function or function of a matrix. Examples are the *determinant* and the *trace*, where $n = m$.

Matrix multiplication: The product of two matrices $A = [a_{ij}]$ $(m \times n)$ and $B = [b_{ij}]$ $(n \times p)$ is defined as

$$AB \equiv \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right] \quad (m \times p).$$

(See Section 2.2 for the rules.)

Matrix norm: A function $\|\cdot\|$ attaching a nonnegative real number $\|A\|$ to an $(m \times m)$ matrix A is a matrix norm if the following four conditions are satisfied for all complex $(m \times m)$ matrices A, B and $c \in \mathbb{C}$:

(i) $\|A\| > 0 \quad \text{if} \quad A \neq 0$,

(ii) $\|cA\| = |c|_{\text{abs}} \, \|A\|$,

(iii) $\|A + B\| \leq \|A\| + \|B\|$ (triangle inequality),

(iv) $\|AB\| \leq \|A\| \, \|B\|$.

Here $|c|_{\text{abs}}$ denotes the modulus of c, that is, $|c|_{\text{abs}} = \sqrt{c\bar{c}}$ with $\bar{c}$ being the complex conjugate of c. Instead of defining a matrix norm for all complex $(m \times m)$ matrices, it may be defined for real matrices only. If (i), (ii), (iii) and (iv) hold for all real $(m \times m)$ matrices and real numbers c, $\|\cdot\|$ is called a matrix norm over the field of real numbers ($\mathbb{R}$). In that case $|c|_{\text{abs}}$ is the absolute value of c. (See Chapter 8.)

Matrix polynomial: Given a polynomial $p(x) = p_0 + p_1 x + \cdots + p_n x^n$, the function $p : \mathbb{C}^{m \times m} \to \mathbb{C}^{m \times m}$ (or $p : \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$) defined by

$$p(A) = p_0 I_m + p_1 A + \cdots + p_n A^n$$

for $(m \times m)$ matrices A, is a matrix polynomial. (See Section 11.1.)

Matrix power series: Given a power series $p(x) = \sum_{n=0}^{\infty} p_n x^n$, the function $p : \mathbb{C}^{m \times m} \to \mathbb{C}^{m \times m}$ (or $p : \mathbb{R}^{m \times m} \to \mathbb{R}^{m \times m}$) defined by

$$p(A) = \sum_{n=0}^{\infty} p_n A^n$$

for $(m \times m)$ matrices A, is a matrix power series, provided the infinite sum exists. (See Section 11.1.)

Matrix product: The product of two or more matrices is a matrix product. (See *matrix multiplication*.)

Metzler matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is a Metzler matrix if $a_{ii} < 0$ and $a_{ij} \geq 0$ for $i \neq j$.

Minimal polynomial of a matrix: The monic polynomial $q(x)$ with minimum degree that annihilates an $(m \times m)$ matrix A, that is, for which $q(A) = 0$, is called the minimal polynomial of A. It is denoted by $q_A(\cdot)$. (See Section 11.1.3.)

Minkowski matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is said to be a Minkowski matrix or *M-matrix* if A is nonsingular, $a_{ij} \leq 0$ for $i \neq j$ and $A^{-1} \geq 0$.

Minkowski–Metzler matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is a Minkowski–Metzler matrix if $a_{ii} > 0$ for $i = 1, \ldots, m$ and $a_{ij} \leq 0$ for $i \neq j$. (See *Leontief matrix* for a special case.)

Minor of a matrix: The determinant of a square submatrix of an $(m \times m)$ matrix A is a minor of A.

Minor of an element of a matrix: For an $(m \times m)$ matrix $A = [a_{ij}]$ the minor of an element a_{ij} is the determinant of the matrix obtained by deleting the ith row and jth column from A.

$$\text{minor}(a_{ij}) \equiv \det \begin{bmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,m} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,m} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,m} \end{bmatrix}.$$

(See also Section 3.4.)

Modal matrix: Let A be an $(m \times m)$ matrix with linearly independent eigenvectors $v_1, \ldots, v_m$ of length 1, that is, $v_i^H v_i = 1$ for $i = 1, \ldots, m$. Then $V = [v_1, \ldots, v_m]$ is called a modal matrix or *eigenvector matrix* of A. (See Chapter 5 for results on eigenvectors.)

Modulus of a vector: The modulus or length of an $(m \times 1)$ vector $x = (x_1, \ldots, x_m)'$ is $\sqrt{x^H x} = (\bar{x}_1 x_1 + \cdots + \bar{x}_m x_m)^{1/2}$.

Monic polynomial matrix: The $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1 x + \cdots + P_r x^r$ is called monic if $P_r = I_m$. (See Section 11.3.)

Monomial matrix: An $(m \times m)$ matrix A is said to be monomial if there exists a *permutation matrix* P and a *diagonal matrix* D such that $A = PD$.

Moore–Penrose (generalized) inverse matrix: The $(n \times m)$ matrix A^+ is the Moore–Penrose (generalized) inverse of the $(m \times n)$ matrix A if it satisfies the following four conditions: (i) $AA^+A = A$, (ii) $A^+AA^+ = A^+$, (iii) $(AA^+)^H = AA^+$, (iv) $(A^+A)^H = A^+A$. (See Section 3.6.2 for the properties.)

MP-inverse of a matrix: Short for *Moore–Penrose inverse matrix.*

Multiplication by a scalar: The product of an $(m \times n)$ matrix $A = [a_{ij}]$ and a number c (a scalar) is defined as

$$cA \equiv [ca_{ij}] \ (m \times n)$$

and

$$Ac \equiv [ca_{ij}] \ (m \times n).$$

(See Section 2.3 for the rules.)

Multiplication of matrices: See *matrix multiplication.*

Multiplicative norm: A norm $\|\cdot\|$ for $(m \times m)$ matrices is called a multiplicative or *submultiplicative* or *matrix norm* if for any $(m \times m)$ matrices A, B, $\|AB\| \leq \|A\|\,\|B\|$. (See Chapter 8 for more details on norms.)

Multiplicity of an eigenvalue: The multiplicity of an eigenvalue of an $(m \times m)$ matrix A is the multiplicity of the corresponding root of the *characteristic equation* of A. (See Section 5.1.)

Negative definite Hermitian form: A *Hermitian form* $x^H Ax$, defined by a Hermitian $(m \times m)$ matrix A, is negative definite if A is negative definite. (See Section 9.12 for more on definite matrices.)

Negative definite matrix: An $(m \times m)$ matrix A is negative definite if it is Hermitian (or real symmetric) and $x^H Ax < 0$ (or $x'Ax < 0$) for any (real) $(m \times 1)$ vector $x \neq 0$. (See Section 9.12 for the properties.)

Negative definite quadratic form: A *quadratic form* $x'Ax$, defined by a real symmetric $(m \times m)$ matrix A, is negative definite if A is negative definite. (See Section 9.12 for more on definite matrices.)

Negative matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is negative if $a_{ij} < 0$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$. (See also Section 9.9.)

Negative semidefinite Hermitian form: A *Hermitian form* $x^H Ax$, defined by a Hermitian $(m \times m)$ matrix A, is negative semidefinite if A is negative semidefinite. (See Section 9.12 for more on definite matrices.)

Negative semidefinite matrix: An $(m \times m)$ matrix A is negative semidefinite if it is Hermitian (or real symmetric) and $x^H Ax \leq 0$ (or $x'Ax \leq 0$) for any (real) $(m \times 1)$ vector x. (See Section 9.12 for the properties.)

Negative semidefinite quadratic form: A *quadratic form* $x'Ax$, defined by a real symmetric $(m \times m)$ matrix A, is negative semidefinite if A is negative semidefinite. (See Section 9.12 for more on definite matrices.)

Nilpotent matrix: An $(m \times m)$ matrix A is nilpotent if there exists a positive integer i such that $A^i = 0$.

Nonderogatory matrix: An $(m \times m)$ matrix A is nonderogatory if every eigenvalue of A has *geometric multiplicity* 1. (See Chapter 5 for more on eigenvalues.)

Nonnegative definite matrix: An $(m \times m)$ matrix A is nonnegative definite or positive semidefinite if it is Hermitian (or real symmetric) and $x^H Ax \geq 0$ (or $x'Ax \geq 0$) for any (real) $(m \times 1)$ vector x. (See *positive semidefinite matrix* and refer to Section 9.12 for the properties.)

Nonnegative matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is nonnegative if $a_{ij} \geq 0$ for $i = 1, \ldots, m,\ j = 1, \ldots, n$. (See Section 9.9.)

Nonpositive definite matrix: An $(m \times m)$ matrix A is nonpositive definite or negative semidefinite if it is Hermitian (or real symmetric) and $x^H Ax \leq 0$ (or $x'Ax \leq 0$) for any (real) $(m \times 1)$ vector x. (See *negative semidefinite matrix* and refer to Section 9.12 for the properties.)

Nonpositive matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is nonpositive if $a_{ij} \leq 0$ for $i = 1, \ldots, m,\ j = 1, \ldots, n$. (See also Section 9.9.)

Nonsingular matrix: An $(m \times m)$ matrix A is nonsingular or *invertible* or *regular* if $\det(A) \neq 0$ and thus A^{-1} exists. (See Section 3.5 for the properties.)

Norm: A function $\|\cdot\|$ attaching a nonnegative real number $\|A\|$ to an $(m \times n)$ matrix A is a norm if the following three conditions are satisfied for all complex $(m \times n)$ matrices A, B and complex numbers c:

(i) $\|A\| > 0 \quad \text{if} \quad A \neq 0,$

(ii) $\|cA\| = |c|_{\text{abs}}\, \|A\|,$

(iii) $||A + B|| \leq ||A|| + ||B||$ (triangle inequality).

Here $|c|_{\text{abs}}$ denotes the modulus of c, that is, $|c|_{\text{abs}} = \sqrt{c\bar{c}}$ with $\bar{c}$ being the complex conjugate of c. Instead of defining a norm for all complex $(m \times n)$ matrices, it may be defined for real matrices only. If (i), (ii) and (iii) hold for all real $(m \times n)$ matrices and real numbers c, $|| \cdot ||$ is called a norm over the field of real numbers ($\mathbb{R}$). In that case $|c|_{\text{abs}}$ is the absolute value of c. (See Chapter 8 for details.)

Normal form of a matrix: An $(m \times m)$ matrix A can be reduced by elementary column and row operations to a matrix $N = [d_{ij}]$ with $d_{ii} = 1$ or 0, $d_{ij} = 0$ for $i \neq j$. The matrix N is said to be the normal form of A. For instance,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are normal forms of matrices.

Normal matrix: An $(m \times m)$ matrix A is normal if $A^H A = AA^H$.

Null matrix: An $(m \times n)$ matrix is a null matrix, denoted by $O_{m \times n}$ or simply by 0, if all its elements are zero. (See also *zero matrix*.)

Null vector: An $(m \times 1)$ or $(1 \times m)$ vector is a null or *zero vector* if all its elements are zero.

Numerical range of a matrix: For an $(m \times m)$ matrix A, the set $\{x^H Ax : x \ (m \times 1), x^H x = 1\}$ is sometimes called the numerical range of A.

Operator norm: Given a *norm* $|| \cdot ||$ for $(m \times 1)$ vectors,

$$||A||_{lub} \equiv \sup \left\{ \frac{||Ax||}{||x||} \ : \ x \ (m \times 1), \ x \neq 0 \right\}$$

defines a norm for $(m \times m)$ matrices called operator norm induced by $|| \cdot ||$. Alternatively it is called sup norm or *least upper bound norm* or *induced norm*. (See Chapter 8 for details on norms.)

Order of a matrix: $(m \times n)$ is the order or *dimension* of a matrix A with m rows and n columns. Sometimes the degree of the *characteristic polynomial* of A is called its order.

Orthogonal complement matrix: For $m > n$, the $(m \times (m-n))$ matrix $A_\perp$ is an orthogonal complement of the $(m \times n)$ matrix A if $[A : A_\perp]$ is invertible and $A^H A_\perp = 0$.

Orthogonal matrix: An $(m \times m)$ matrix A is orthogonal if $A' = A^{-1}$. (See Section 9.10 for its properties.)

Orthogonal projector: A *Hermitian idempotent* matrix is said to be an orthogonal projector, that is, A $(m \times m)$ is an orthogonal projector if $A = A^H = A^2$. For instance, given an $(m \times n)$ matrix B with $\text{rk}(B) = m$, the matrix $B^H(BB^H)^{-1}B$ is an orthogonal projector. (See Section 9.7 for more on Hermitian matrices and Section 9.8 for the properties of idempotent matrices.)

Orthogonal vectors: The $(m \times 1)$ vectors x and y are orthogonal if $x^H y = 0$.

Orthogonally equivalent matrices: The $(m \times n)$ matrices A and B are orthogonally equivalent if there exists a real orthogonal $(m \times m)$ matrix U and a real orthogonal $(n \times n)$ matrix V, such that $A = U'BV$. (See also Section 6.2.)

Orthogonally similar matrices: The real $(m \times m)$ matrices A and B are orthogonally similar if an orthogonal $(m \times m)$ matrix U exists such that $B = UAU'$. (See also Section 6.2.)

Orthonormal vectors: The $(m \times 1)$ vectors x and y are orthonormal if $x^H y = 0$ and $x^H x = y^H y = 1$.

Orthostochastic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be orthostochastic if there exists a real orthogonal matrix $Q = [q_{ij}]$ such that $a_{ij} = q_{ij}^2$ for $i, j = 1, \ldots, m$.

Orthosymmetric matrix: A *Hankel matrix* is sometimes called an orthosymmetric matrix.

Oscillatory matrix: A real $(m \times m)$ matrix A is called oscillatory if all its minors of all possible orders are nonnegative and there exists a positive integer k such that all minors of A^k are positive.

Parity check matrix: For $n > m$ an $(m \times n)$ matrix $A = [I_m : B]$ is a *check matrix* or parity check matrix if B is a *binary matrix*, that is, all elements of B are 0 or 1.

Partitioned matrix: An $(m \times n)$ matrix $A = [A_{ij}]$ consisting of $(m_i \times n_j)$ submatrices A_{ij}, $i = 1, \ldots, p$, $j = 1, \ldots, q$, is a partitioned matrix. (See Section 9.11 for the properties.)

Permanent of a matrix: The permanent of an $(m \times n)$ matrix $A = [a_{ij}]$ with $m \leq n$ is the quantity defined by

$$\sum \prod_{i=1}^{m} a_{ij},$$

where the sum is taken over all one-to-one functions from $\{1, 2, \ldots, m\}$ to $\{1, 2, \ldots, n\}$.

Permutation equivalent matrices: The $(m \times m)$ matrices A and B are said to be permutation equivalent if there exist *permutation matrices* P and Q such that $A = PBQ$.

Permutation matrix: An $(m \times m)$ matrix with exactly one 1 in each row and each column and zeros elsewhere is a permutation matrix. For instance,

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is a permutation matrix. (See Section 9.2 for the properties of special permutation matrices.)

Persymmetric matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is persymmetric if it is symmetric with respect to the diagonal from the upper right hand corner to the lower left hand corner, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,m-1} & a_{1m} \\ a_{21} & a_{22} & \cdots & & a_{1,m-1} \\ \vdots & & & & \vdots \\ a_{m-1,1} & & \cdots & a_{22} & a_{12} \\ a_{m1} & a_{m-1,1} & \cdots & a_{21} & a_{11} \end{bmatrix}$$

is persymmetric.

Polynomial matrix: An $(m \times n)$ matrix $P(x) = [p_{ij}(x)]$ whose typical elements are polynomials $p_{ij}(x) = p_{ij,0} + p_{ij,1}x + \cdots + p_{ij,r_{ij}}x^{r_{ij}}$, is a polynomial matrix. Alternative notation: $P(x) = P_0 + P_1 x + \cdots + P_r x^r$, where $r = \max_{i,j} r_{ij}$ and

$$P_k = \begin{bmatrix} p_{11,k} & \cdots & p_{1n,k} \\ \vdots & \ddots & \vdots \\ p_{m1,k} & \cdots & p_{mn,k} \end{bmatrix}, \quad k = 0, 1, \ldots, r,$$

with $p_{ij,k} = 0$ for $k > r_{ij}$. (See Section 11.3 for details.)

Positive definite Hermitian form: A *Hermitian form* $x^H Ax$, defined by a Hermitian $(m \times m)$ matrix A, is positive definite if A is positive definite. (See Section 9.12 for the properties of positive definite matrices.)

Positive definite matrix: An $(m \times m)$ matrix A is positive definite if it is Hermitian (or real symmetric) and $x^H Ax > 0$ (or $x'Ax > 0$) for any (real) $(m \times 1)$ vector $x \neq 0$. (See Section 9.12 for the properties.)

Positive definite quadratic form: A *quadratic form* $x'Ax$, defined by a real symmetric $(m \times m)$ matrix A, is positive definite if A is positive definite. (See Section 9.12 for the properties of positive definite matrices.)

Positive matrix: A real $(m \times n)$ matrix $A = [a_{ij}]$ is positive if $a_{ij} > 0$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$. (See Section 9.9 for the properties.)

Positive semidefinite Hermitian form: A *Hermitian form* $x^H Ax$, defined by a Hermitian $(m \times m)$ matrix A, is positive semidefinite if A is positive semidefinite. (See Section 9.12 for the properties of positive semidefinite matrices.)

Positive semidefinite matrix: An $(m \times m)$ matrix A is positive semidefinite if it is Hermitian (or real symmetric) and $x^H Ax \geq 0$ (or $x'Ax \geq 0$) for any (real) $(m \times 1)$ vector x. (See Section 9.12 for the properties.)

Positive semidefinite quadratic form: A *quadratic form* $x'Ax$, defined by a real symmetric $(m \times m)$ matrix A, is positive semidefinite if A is positive semidefinite. (See Section 9.12 for the properties of positive semidefinite matrices.)

Postmultiplication: Let A, B be $(m \times n)$ and $(n \times p)$ matrices, respectively. In the matrix product AB, the matrix A is said to be postmultiplied by B.

Power of a matrix: The ith power of the $(m \times m)$ matrix A, denoted by A^i, is defined as follows:

$$A^i = \begin{cases} \prod_{j=1}^{i} A & \text{for positive integers } i \\ I_m & \text{for } i = 0 \\ (\prod_{j=1}^{-i} A)^{-1} & \text{for negative integers } i, \text{ if } \det(A) \neq 0 \end{cases}$$

If A can be written as

$$A = U \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} U^H$$

for some *unitary matrix* U (see Chapter 6 for conditions), then the power of A is defined for any $\alpha \in \mathbb{R}, \alpha > 0$, as follows:

$$A^\alpha \equiv U \begin{bmatrix} \lambda_1^\alpha & & 0 \\ & \ddots & \\ 0 & & \lambda_m^\alpha \end{bmatrix} U^H.$$

This definition applies, for instance, for *Hermitian matrices*. (See Section 3.7 for some properties of powers of matrices).

Premultiplication: Let A, B be $(m \times n)$ and $(n \times p)$ matrices respectively. In the matrix product AB, the matrix B is said to be premultiplied by A.

Primitive matrix: An $(m \times m)$ matrix A is said to be primitive if it is *irreducible* and it has only one eigenvalue with modulus equal to its *spectral radius*.

Principal diagonal of a matrix: The elements $a_{ii}, i = 1, \ldots, m$, constitute the principal diagonal or *main diagonal* of an $(m \times m)$ matrix $A = [a_{ij}]$.

Principal minor: The determinant

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

of a *principal submatrix* of the $(m \times m)$ matrix $A = [a_{ij}]$ is a principal minor of A for $k = 1, \ldots, m-1$.

Principal submatrix: The matrices

$$\begin{bmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{bmatrix}, \quad k = 1, \ldots, m-1,$$

are the principal submatrices of the $(m \times m)$ matrix $A = [a_{ij}]$.

Principal vector of a matrix: Alternative expression for *generalized eigenvector of a matrix*.

Product matrix: The product AB of the $(m \times n)$ matrix A and the $(n \times p)$ matrix B is the product matrix of A and B. (See Section 2.2 for properties.)

Product of matrices: See *matrix multiplication*.

Projection matrix: An $(m \times m)$ matrix P_A is a projection matrix for the $(m \times n)$ matrix A if $P_A = P_A^2 = P_A^H$, $P_A A = A$ and $\text{rk}(P_A) = \text{rk}(A)$. For instance, if $\text{rk}(A) = n$, $P_A = A(A^H A)^{-1} A^H$ is a projection matrix for A. (See Sections 9.7 and 9.8 for properties.)

Pseudo-inverse matrix: A *generalized inverse matrix* is sometimes called a pseudo-inverse matrix. (See Section 3.6 for some properties.)

Quadratic form: Given a real symmetric $(m \times m)$ matrix A, the function $Q : \mathbb{R}^m \to \mathbb{R}$ defined by $Q(x) = x'Ax$ is called a quadratic form. (See Section 9.13 for the properties of symmetric matrices.)

Quadratic matrix: An $(m \times n)$ matrix with $m = n$ is a quadratic matrix. (See also *square matrix*.)

Rank of a matrix: The rank of a matrix A, denoted by rk A or $\text{rk}(A)$, is the maximum number of linearly independent rows or columns of A. (See Section 4.3.)

Rank of a polynomial matrix: The rank of a polynomial matrix is the number of columns of the largest submatrix whose determinant is not identically zero. (See Section 11.3.)

Rational matrix: An $(m \times n)$ matrix $R(x) = [r_{ij}(x)]$ whose elements $r_{ij}(x)$ are rational functions, that is, ratios of polynomials, is a rational matrix.

Rayleigh ratio: The Rayleigh ratio for an $(m \times m)$ square matrix A and an $(m \times 1)$ vector $x \neq 0$ is

$$\frac{x^H A x}{x^H x}.$$

(See Sections 5.2 and 5.3 for some of its properties.)

Real matrix: A matrix is called real if all its elements are real numbers.

Real polynomial matrix: A polynomial matrix is called real if all its elements are real polynomials. (See Section 11.3.)

Rectangular matrix: An $(m \times n)$ matrix with $m \neq n$ is sometimes called a rectangular matrix.

Reducible matrix: An $(m \times m)$ matrix A is said to be reducible or *decomposable* if there exists a *permutation matrix* P such that

$$P'AP = \begin{bmatrix} A_{11} & O_{k\times(m-k)} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $(k \times k)$, A_{22} is $((m-k)\times(m-k))$ and A_{21} is $((m-k)\times k)$.

Reflexive generalized inverse matrix: The $(n \times m)$ matrix A_r^- is a reflexive generalized inverse of the $(m\times n)$ matrix A if $AA_r^-A = A$ and $A_r^-AA_r^- = A_r^-$. (See Section 3.6 for properties of generalized inverse matrices.)

Regular matrix: An $(m \times m)$ matrix A is regular or *nonsingular* or *invertible* if $\det(A) \neq 0$ and thus A^{-1} exists. (See Section 3.5 for details.)

Regular polynomial matrix: The $(m \times m)$ polynomial matrix $P(x) = P_0 + P_1x + \cdots + P_rx^r$ is regular if P_r is a nonsingular matrix. (See Section 11.3.)

Relatively left prime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be relatively left prime or *left coprime* if their *greatest common left divisors* are *unimodular*. (See Section 11.3.)

Relatively right prime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be relatively right prime or *right coprime* if their *greatest common right divisors* are *unimodular*. (See Section 11.3.)

Resultant matrix: An $(m \times m)$ matrix A is a resultant matrix if $\det(A)$ is a resultant. (See Section 11.1.1 for the definition of a *resultant*.)

Reverse unit matrix: The $(m \times m)$ matrix

$$A = [a_{ij}] = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

with

$$a_{ij} = \begin{cases} 1 & \text{for } i+j = m+1 \\ 0 & \text{otherwise} \end{cases}$$

is called an *exchange matrix*, reverse unit matrix or *flip matrix*.

Riccati equation: Let A, B be $(m \times m)$ matrices, P $(m \times m)$ real symmetric, R $(m \times m)$ real symmetric, positive definite and Q real symmetric, positive semidefinite. The equation

$$PBR^{-1}B'P - A'P - PA - Q = 0$$

is called a Riccati equation.

Right coprime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be relatively right prime or right coprime if their *greatest common right divisors* are *unimodular*. (See Section 11.3.)

Right divisibility of a polynomial matrix: A polynomial matrix $P(x)$ is divisible on the right by a polynomial matrix $Q(x)$ if $Q(x)$ is a *right divisor* of $P(x)$, that is, if for some polynomial matrix $T(x)$ with degree greater than 0, $P(x) = T(x)Q(x)$. (See Section 11.3.)

Right divisor of a polynomial matrix: A polynomial matrix $Q(x)$ is a right divisor of the polynomial matrix $P(x)$ if a polynomial matrix $T(x)$ exists such that $P(x) = T(x)Q(x)$. (See Section 11.3.)

Right eigenvector: An eigenvector or characteristic vector is sometimes called a right eigenvector. (See *eigenvector of a matrix*.)

Right inverse matrix: Let A be an $(m \times n)$ matrix. An $(n \times m)$ matrix B satisfying $AB = I_m$ is called a right inverse of A. (See also Sections 3.5 and 3.6.)

Right multiple of a polynomial matrix: The $(m \times n)$ polynomial matrix $P(x)$ is a right multiple of an $(m \times h)$ polynomial matrix $T(x)$ if there exists an $(h \times n)$ polynomial matrix $Q(x)$ such that $P(x) = T(x)Q(x)$. (See Section 11.3.)

Right quotient of polynomial matrices: Let $P(x) = P_0 + P_1x + \cdots + P_rx^r$ $(m \times n), P_r \neq 0$, and $Q(x) = Q_0 + Q_1x + \cdots + Q_sx^s$ $(h \times n)$ regular, with $s \leq r$. An $(m \times h)$ polynomial matrix $T(x)$ is a right quotient of $P(x)$ and $Q(x)$ if an $(m \times n)$ polynomial matrix $R(x)$ with degree less than s exists such that $P(x) = T(x)Q(x) + R(x)$. (See Section 11.3.)

Right remainder: Let $P(x)$ and $Q(x)$ be polynomial matrices of degrees r and s, respectively, $s \leq r$. If $T(x)$ is a *right quotient* of $P(x)$ and $Q(x)$ with $P(x) = T(x)Q(x) + R(x)$, then $R(x)$ is said to be a right remainder. (See Section 11.3.)

Rotation matrix: For $r, s \in \mathbb{N}$, $r < s \leq m$, an $(m \times m)$ matrix $G_{rs} = [g_{ij}]$ is said to be a rotation matrix or *Givens matrix* if $g_{rr} = g_{ss} = \cos\theta$, $g_{ii} = 1$ for $i = 1, \ldots, m$, $i \neq r, s$, $g_{sr} = -\sin\theta$, $g_{rs} = \sin\theta$ and all other elements are zero. For example,

$$G_{13} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \quad G_{12} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

are rotation matrices.

Routh matrix: A *tridiagonal* $(m \times m)$ matrix $A = [a_{ij}]$ is called Routh matrix if

$$\begin{aligned} &a_{11} = -b_1, \quad a_{ii} = 0, \\ &a_{i-1,i} = \sqrt{b_i}, \quad a_{i,i-1} = -\sqrt{b_i}, \quad i = 2, 3, \ldots, m, \end{aligned}$$

where $b_1, \ldots, b_m$ are nonzero real numbers. For example,

$$A = \begin{bmatrix} -3 & \sqrt{3} & 0 & \cdots & \cdots & 0 \\ -\sqrt{3} & 0 & \sqrt{2.5} & & & 0 \\ 0 & -\sqrt{2.5} & 0 & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & & & \ddots & 0 & \sqrt{5} \\ 0 & \cdots & \cdots & \cdots & -\sqrt{5} & 0 \end{bmatrix}$$

is a Routh matrix.

Row dimension of a matrix: The number of rows of a matrix is its row dimension.

Row rank of a matrix: The row rank of a matrix A, denoted by row rk A or row rk(A), is the maximum number of linearly independent rows of A. (See Section 4.3.)

Row regular matrix: An $(m \times n)$ matrix A is row regular if rk$(A) = m$.

Row stochastic matrix: Alternative expression for *stochastic matrix*.

Row vector: A $(1 \times n)$ matrix is a row vector.

Row vectorization: Stacking the rows of a matrix in a row vector is called row vectorization of the matrix. Notation: For $A = [a_{ij}]$ $(m \times n)$

$$\text{rvec}(A) = [a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, a_{31}, \ldots, a_{mn}] \quad (1 \times mn).$$

(See Section 7.1.)

Scalar matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is sometimes said to be a scalar matrix if, for some $c \in \mathbb{C}$,

$$a_{ij} = \begin{cases} c & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases},$$

that is, $A = \operatorname{diag}(c, \ldots, c) = cI_m$.

Scalar multiplication: See *multiplication by a scalar*.

Scalar product: The scalar product of two vectors

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad (m \times 1)$$

is defined as

$$x^H y \equiv \sum_{i=1}^{m} \bar{x}_i y_i,$$

where $\bar{x}_i$ is the complex conjugate of x_i. (See also *inner product*.)

Schur complement of a matrix: For matrices $A\,(m \times m)$ nonsingular, $B\,(m \times n)$, $C\,(n \times m)$, $D\,(n \times n)$, the $(n \times n)$ matrix $S = D - CA^{-1}B$ is called the Schur complement of A in

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Schur product: See *Hadamard product*.

Schur-stochastic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be Schur-stochastic or *unitary stochastic* if there exists a unitary $(m \times m)$ matrix $U = [u_{ij}]$ such that $a_{ij} = |u_{ij}|^2_{\text{abs}}$ for $i, j = 1, \ldots, m$.

Secondary diagonal: The elements $a_{1+i,m-i}$, $i = 0, \ldots, m-1$, of an $(m \times m)$ matrix $A = [a_{ij}]$ constitute the secondary diagonal of A. In other words, the secondary diagonal is the diagonal from the upper right hand to the lower left hand corner of A.

Selfadjoint matrix: A Hermitian $(m \times m)$ matrix A is sometimes called selfadjoint, that is, A is selfadjoint if $A^H = A$. (See Section 9.7 for the properties.)

Semicirculant matrix: An $(m \times m)$ matrix of the form

$$\begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{m-1} & a_m \\ 0 & a_1 & a_2 & \cdots & a_{m-2} & a_{m-1} \\ 0 & 0 & a_1 & \cdots & a_{m-3} & a_{m-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & & a_1 & a_2 \\ 0 & 0 & 0 & \cdots & 0 & a_1 \end{bmatrix}$$

is said to be semicirculant. (See also *circulant matrix.*)

Semidefinite matrix: A matrix is semidefinite if it is *positive semidefinite* or *negative semidefinite*. (See Section 9.12 for some properties.)

Seminegative matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is seminegative if $a_{ij} \leq 0$ for all $i, j = 1, \ldots, m$ and at least one element is strictly negative. (See also Section 9.9.)

Seminorm: A function $||\cdot||$ attaching a nonnegative real number $||A||$ to an $(m \times n)$ matrix A is a seminorm if the following three conditions are satisfied for all complex $(m \times n)$ matrices A, B and complex numbers c :

(i) $||A|| \geq 0$,

(ii) $||cA|| = |c|_{\text{abs}}\, ||A||$,

(iii) $||A + B|| \leq ||A|| + ||B||$ (triangle inequality).

Here $|c|_{\text{abs}}$ denotes the modulus of c, that is, $|c|_{\text{abs}} = \sqrt{c\bar{c}}$ with $\bar{c}$ being the complex conjugate of c. Instead of defining a seminorm for all complex $(m \times n)$ matrices, it may be defined for real matrices only. If (i), (ii) and (iii) hold for all real $(m \times n)$ matrices and real numbers c, $||\cdot||$ is called a seminorm over the field of real numbers ($\mathbb{R}$). In that case $|c|_{\text{abs}}$ is the absolute value of c. (See Chapter 8 for further details.)

Semiorthogonal matrix: An $(m \times n)$ matrix A is semiorthogonal if $AA' = I_m$ or $A'A = I_n$. (See also *orthogonal matrix.*)

Semipositive matrix: A real $(m \times m)$ matrix $A = [a_{ij}]$ is semipositive if $a_{ij} \geq 0$ for all $i, j = 1, \ldots, m$ and at least one element is strictly positive. (See Section 9.9.)

Shift matrix: The $(m \times m)$ *circulant matrix*

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

is called a shift matrix. (See Section 9.1 for some properties.)

Similar matrices: The $(m \times m)$ matrices A and B are similar if an invertible matrix P exists such that $B = PAP^{-1}$. (See also Chapter 6.)

Simple eigenvalue: A root of the characteristic polynomial of an $(m \times m)$ matrix A with multiplicity 1 is a simple eigenvalue of A. (See Section 5.1.)

Simple matrix: An $(m \times m)$ matrix A is said to be simple if it is similar to a diagonal matrix, that is, if there exists an invertible matrix P such that PAP^{-1} is a diagonal matrix. (See Chapter 6 for examples.)

Simultaneously diagonalizable matrices: The $(m \times m)$ matrices $A_1, \ldots, A_n$ are said to be simultaneously diagonalizable if a nonsingular $(m \times m)$ matrix P exists such that PA_iP^{-1} is diagonal for $i = 1, \ldots, n$. (See Sections 6.1.2 and 6.2.3.)

Sine of a matrix: The sine of an $(m \times m)$ matrix A is defined as

$$\sin(A) \equiv \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1}.$$

(See Section 11.1.3.)

Singular matrix: An $(m \times m)$ matrix A is singular if $\det(A) = 0$.

Singular value: The singular values of an $(m \times n)$ matrix A are the nonnegative square roots of the eigenvalues of AA^H, if $m \leq n$, and of $A^H A$, if $m \geq n$. (See Sections 5.1 and 5.5.)

Skew-circulant matrix: For $a_1, \ldots, a_m \in \mathbb{C}$, an $(m \times m)$ matrix of the form

$$\begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{m-1} & a_m \\ -a_m & a_1 & a_2 & \cdots & a_{m-2} & a_{m-1} \\ -a_{m-1} & -a_m & a_1 & \cdots & a_{m-3} & a_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -a_3 & -a_4 & -a_5 & \cdots & a_1 & a_2 \\ -a_2 & -a_3 & -a_4 & \cdots & -a_m & a_1 \end{bmatrix}$$

is said to be a skew-circulant matrix. (See also *circulant matrix*.)

Skew-Hermitian matrix: An $(m \times m)$ matrix

$$A = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1m} \\ -\bar{a}_{12} & 0 & \cdots & a_{2m} \\ \vdots & & \ddots & \vdots \\ -\bar{a}_{1m} & -\bar{a}_{2m} & \cdots & 0 \end{bmatrix} = [a_{ij}],$$

with $a_{ij} = -\bar{a}_{ji}$, is a skew-Hermitian matrix. Here $\bar{a}_{ij}$ denotes the complex conjugate of a_{ij}. In other words, A is skew-Hermitian if $A^H = -A$. (See Section 9.7.)

Skew prime polynomial matrices: The polynomial matrices $P(x)$ and $Q(x)$ are said to be skew prime if $\det P(x)$ and $\det Q(x)$ are relatively prime polynomials. (See Section 11.1.1 for the definition of *relatively prime polynomials* and Section 11.3 for more on skew prime polynomial matrices.)

Skew-symmetric matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 0 & a_{12} & \cdots & a_{1m} \\ -a_{12} & 0 & & a_{2m} \\ \vdots & & \ddots & \vdots \\ -a_{1m} & -a_{2m} & \cdots & 0 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 0$ and $a_{ij} = -a_{ji}$, is skew-symmetric. In other words, an $(m \times m)$ matrix A is skew-symmetric if $A' = -A$. (See Section 9.13.)

Skew-symmetric part of a matrix: The matrix $\frac{1}{2}(A - A')$ is the skew-symmetric part of an $(m \times m)$ matrix A.

Sparse matrix: A matrix is sparse if only a small fraction of its elements are nonzero.

Spectral radius: $\rho(A) \equiv \max_i |\lambda_i|_{\text{abs}}$ is the spectral radius of an $(m \times m)$ matrix A with eigenvalues $\lambda_1, \ldots, \lambda_m$. (See Section 5.4.)

Spectrum of a matrix: The set of eigenvalues of a matrix A is called the spectrum of A. (See Chapter 5.)

Square matrix: An $(m \times n)$ matrix with $m = n$ is a square matrix or *quadratic matrix*.

Square root matrix: An $(m \times m)$ matrix B is a square root of the $(m \times m)$ matrix A if $BB = A$. It is denoted by $A^{1/2}$. (See Section 6.1.4.)

Stable matrix: An $(m \times m)$ square matrix A is stable or *convergent* if $A^n \to 0$ for $n \to \infty$. (See Section 9.3.)

Stieltjes matrix: A symmetric nonsingular *M-matrix* is called a Stieltjes matrix.

Stochastic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be a stochastic or *row stochastic* or *transition matrix* if $0 \leq a_{ij} \leq 1$, $i, j = 1, \ldots, m$, and $\sum_{j=1}^{m} a_{ij} = 1$, $i = 1, \ldots, m$. For instance,

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

is a stochastic matrix. (See Section 9.9.)

Strictly lower triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m,m-1} & 0 \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $j \geq i$ is a strictly lower triangular matrix. (See Section 9.14 for the properties.)

Strictly triangular matrix: A matrix is strictly triangular if it is *strictly lower triangular* or *strictly upper triangular*. (See Section 9.14 for the properties.)

Strictly upper triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 0 & a_{12} & \cdots & a_{1m} \\ & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_{m-1,m} \\ 0 & \cdots & \cdots & 0 \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $i \geq j$ is strictly upper triangular. (See Section 9.14 for the properties.)

Striped matrix: A matrix is called striped if it is either a *Hankel* or a *Toeplitz matrix*.

Submatrix: A $(p \times q)$ matrix $B = [a_{i_k j_l}]$ with $i_1 < \cdots < i_p$ and $j_1 < \cdots < j_q$ is a submatrix of the $(m \times n)$ matrix $A = [a_{ij}]$.

Submultiplicative norm: See *multiplicative norm* or *matrix norm.*

Subtraction of matrices: The difference between two matrices $A = [a_{ij}]$ $(m \times n)$ and $B = [b_{ij}]$ $(m \times n)$ is defined as

$$A - B \equiv [a_{ij} - b_{ij}] \; (m \times n).$$

(See Section 2.1 for the rules.)

Sylvester matrix: Let $p(x) = p_0 + p_1 x + \cdots + p_{m-1}x^{m-1} + x^m$, $q(x) = q_0 + q_1 x + \cdots + q_{n-1}x^{n-1} + q_n x^n$ be polynomials. The Sylvester matrix corresponding to $p(\cdot)$ and $q(\cdot)$ is

$$S = \begin{bmatrix} 1 & p_{m-1} & p_{m-2} & \dots & p_0 & 0 & \dots & 0 \\ 0 & 1 & p_{m-1} & \dots & p_1 & p_0 & & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & p_{m-1} & \dots & p_1 & p_0 \\ & & & & & & & \\ q_n & q_{n-1} & q_{n-2} & \dots & q_0 & 0 & \dots & 0 \\ 0 & q_n & q_{n-1} & \dots & q_1 & q_0 & & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & q_n & q_{n-1} & \dots & q_1 & q_0 \end{bmatrix}$$

$$((m+n) \times (m+n))$$

where the upper part consists of n rows and the lower part consists of m rows. For example, for $p(x) = p_0 + p_1 x + p_2 x^2 + x^3$ and $q(x) = q_0 + q_1 x + q_2 x^2$,

$$S = \begin{bmatrix} 1 & p_2 & p_1 & p_0 & 0 \\ 0 & 1 & p_2 & p_1 & p_0 \\ q_2 & q_1 & q_0 & 0 & 0 \\ 0 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & q_2 & q_1 & q_0 \end{bmatrix}$$

(See Section 11.1.2.)

Symmetric matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{12} & a_{22} & \cdots & a_{2m} \\ \vdots & & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = a_{ji}$, $i, j = 1, \ldots, m$, is symmetric. In other words, an $(m \times m)$ matrix A is symmetric if $A' = A$. (See Section 9.13 for the properties.)

Symmetric part of a matrix: The matrix $\frac{1}{2}(A + A')$ is the symmetric part of an $(m \times m)$ matrix A.

Symmetric r-Toeplitz matrix: For $r \leq m$ and $\alpha_1, \ldots, \alpha_r \in \mathbb{C}$, an $(m \times m)$ matrix of the form

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_r & 0 & \cdots & 0 \\ \alpha_2 & \alpha_1 & \cdots & \alpha_{r-1} & \alpha_r & & 0 \\ \vdots & \vdots & \ddots & & \ddots & \ddots & \vdots \\ \alpha_r & \alpha_{r-1} & & \ddots & & \ddots & \alpha_r \\ 0 & \alpha_r & & & \ddots & & \alpha_{r-1} \\ \vdots & & \ddots & & & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_r & \cdots & \cdots & \alpha_1 \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = a_{i+k,j+k}$ and $a_{ij} = a_{ji}$ for all i, j, k and $a_{ij} = 0$ for $|i - j|_{\text{abs}} \geq r$, is a symmetric r-Toeplitz matrix. (See also *Toeplitz matrix.*)

Symmetric Toeplitz matrix: For $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$, an $(m \times m)$ matrix of the form

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{m-1} & \alpha_m \\ \alpha_2 & \alpha_1 & \alpha_2 & \cdots & \alpha_{m-2} & \alpha_{m-1} \\ \alpha_3 & \alpha_2 & \alpha_1 & & \alpha_{m-3} & \alpha_{m-2} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha_{m-1} & \alpha_{m-2} & & \ddots & \alpha_1 & \alpha_2 \\ \alpha_m & \alpha_{m-1} & \cdots & \cdots & \alpha_2 & \alpha_1 \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = a_{i+k,j+k}$ and $a_{ij} = a_{ji}$ for all i, j, k, is a symmetric Toeplitz matrix. (See also *Toeplitz matrix.*)

Tensor product: See *Kronecker product* or *direct product.*

Toeplitz matrix: For $\alpha_1, \ldots, \alpha_{2m-1} \in \mathbb{C}$, an $(m \times m)$ matrix of the form

$$\begin{bmatrix} \alpha_1 & \alpha_{m+1} & \alpha_{m+2} & \cdots & \alpha_{2m-2} & \alpha_{2m-1} \\ \alpha_2 & \alpha_1 & \alpha_{m+1} & \cdots & \alpha_{2m-3} & \alpha_{2m-2} \\ \alpha_3 & \alpha_2 & \alpha_1 & & \alpha_{2m-4} & \alpha_{2m-3} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha_{m-1} & \alpha_{m-2} & \alpha_{m-3} & \ddots & \alpha_1 & \alpha_{m+1} \\ \alpha_m & \alpha_{m-1} & \alpha_{m-2} & \cdots & \alpha_2 & \alpha_1 \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = a_{i+k,j+k}$ for all i, j, k, is a Toeplitz matrix.

Totally nonnegative matrix: A real square matrix is called totally nonnegative if all its minors (of all possible orders) are nonnegative.

Totally positive matrix: A real square matrix is called totally positive if all its minors (of all possible orders) are positive.

Trace of a matrix: The trace of an $(m \times m)$ matrix $A = [a_{ij}]$ is

$$\operatorname{tr} A = \operatorname{tr}(A) \equiv a_{11} + \cdots + a_{mm} = \sum_{i=1}^{m} a_{ii}.$$

(See Section 4.1 for the properties.)

Transition matrix (of a Markov chain): A nonnegative $(m \times m)$ matrix $A = [a_{ij}]$ is a transition matrix or a *stochastic matrix* if $0 \leq a_{ij} \leq 1$, $i, j = 1, \ldots, m$, and $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, \ldots, m$. For example,

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

is a transition matrix. (See Section 9.9.)

Transpose matrix: The transpose of the $(m \times n)$ matrix $A = [a_{ij}]$ is the $(n \times m)$ matrix

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}' \equiv \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

(see Section 3.1 for its properties). Note that the transpose of a matrix A is sometimes denoted by A^T.

Transposition matrix: An $(m \times m)$ matrix is a transposition matrix if it is obtained from the unit matrix I_m by interchanging two rows or two columns. For example,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is a transposition matrix.

Triangular matrix: A *lower* or *upper triangular matrix* is a triangular matrix. (See Section 9.14 for the properties.)

Tridiagonal matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is tridiagonal if $a_{ij} = 0$ for $|i - j|_{\text{abs}} > 1$, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & & 0 & 0 \\ 0 & a_{32} & a_{33} & & 0 & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & & \ddots & a_{m-1,m-1} & a_{m-1,m} \\ 0 & 0 & \cdots & & a_{m,m-1} & a_{m,m} \end{bmatrix}.$$

(See also *band matrix*.)

Tripotent matrix: An $(m \times m)$ matrix A is tripotent if $A^3 = A$. (See also *idempotent matrix*.)

Unimodular integer matrix: An $(m \times m)$ *integer matrix* A is unimodular if $\det(A) = \pm 1$.

Unimodular polynomial matrix: The $(m \times m)$ polynomial matrix $P(x)$ is said to be unimodular or *invertible* if $\det P(x) = \text{constant} \neq 0$, that is, $\det P(x)$ is a constant function. (See Section 11.3.)

Unipotent matrix: An $(m \times m)$ matrix A is said to be unipotent or *involutory* if $A^2 = I_m$. For example, for an $(n \times n)$ matrix B,

$$\begin{bmatrix} I_n & B \\ 0 & -I_n \end{bmatrix}$$

is unipotent.

Unit lower triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & 1 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 1$ for $i = 1, \ldots, m$, and $a_{ij} = 0$ for $i < j$, is a unit lower triangular matrix. (See also *lower triangular matrix.*)

Unit matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is a unit matrix or *identity matrix*, denoted by I_m, if $a_{ij} = 0$ for $i \neq j$ and $a_{ii} = 1$ for $i = 1, \ldots, m$.

Unit triangular matrix: A *unit lower* or *unit upper triangular matrix* is a unit triangular matrix.

Unit upper triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 1 & a_{12} & \cdots & a_{1m} \\ 0 & 1 & & a_{2m} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [a_{ij}]$$

with $a_{ii} = 1$ for $i = 1, \ldots, m$, and $a_{ij} = 0$ for $i > j$, is unit upper triangular. (See also *upper triangular matrix.*)

Unit vector: An $(m \times 1)$ vector a is a unit vector if $||a||_2 = \sqrt{a^H a} = 1$.

Unitarily equivalent matrices: The $(m \times n)$ matrices A and B are unitarily equivalent if there exists a unitary $(m \times m)$ matrix U and a unitary $(n \times n)$ matrix V, such that $A = U^H B V$. (See also Section 9.15.)

Unitarily similar matrices: The $(m \times m)$ matrices A and B are unitarily similar if a unitary $(m \times m)$ matrix U exists such that $B = UAU^H$. (See also Section 9.15.)

Unitary matrix: An $(m \times m)$ matrix A is unitary if $A^H = A^{-1}$. (See Section 9.15 for the properties.)

Unitary stochastic matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is said to be unitary stochastic or *Schur-stochastic* if there exists a unitary $(m \times m)$ matrix $U = [u_{ij}]$ such that $a_{ij} = |u_{ij}|^2_{\text{abs}}$ for $i, j = 1, \ldots, m$.

Upper block triangular matrix: An $(m \times n)$ matrix

$$\begin{bmatrix} A_{11} & \cdots & \cdots & A_{1p} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{pp} \end{bmatrix} = [A_{ij}],$$

where the A_{ij} are $(m_i \times n_j)$ matrices with $\sum_{i=1}^{p} m_i = m$, $\sum_{j=1}^{p} n_j = n$ and $A_{ij} = 0$ for $i > j$, is upper block triangular. (See also *upper triangular matrix* and *partitioned matrix*.)

Upper Hessenberg matrix: An $(m \times m)$ matrix $A = [a_{ij}]$ is an upper Hessenberg matrix or simply a *Hessenberg matrix* if $a_{ij} = 0$ for $i > j + 1$, that is,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,m-1} & a_{1,m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,m-1} & a_{2,m} \\ 0 & a_{32} & a_{33} & & a_{3,m-1} & a_{3,m} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & a_{m-1,m-1} & a_{m-1,m} \\ 0 & 0 & 0 & \cdots & a_{m,m-1} & a_{m,m} \end{bmatrix}.$$

(See also Section 6.2.3.)

Upper triangular matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} a_{11} & \cdots & \cdots & a_{1m} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{mm} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = 0$ for $i > j$ is upper triangular. (See Section 9.14 for the properties.)

Vandermonde determinant: For $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$, the determinant

$$\det \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_m \\ \vdots & \vdots & & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \ldots & \lambda_m^{m-1} \end{bmatrix} = \prod_{j<i} (\lambda_i - \lambda_j).$$

is called a Vandermonde determinant. In other words, the determinant of a *Vandermonde matrix* is a Vandermonde determinant.

Vandermonde matrix: An $(m \times m)$ matrix

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_m \\ \vdots & \vdots & & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \cdots & \lambda_m^{m-1} \end{bmatrix} = [a_{ij}]$$

with $a_{ij} = \lambda_j^{i-1}$ for complex numbers λ_j, $j = 1, \ldots, m$, is a Vandermonde matrix.

vec operator: For an $(m \times n)$ matrix $A = [a_{ij}]$, the vec operator is defined as

$$\operatorname{vec} A = \operatorname{vec}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ a_{13} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1),$$

that is, vec stacks the columns of A in a column vector. (See Section 7.2 for the rules.)

vech operator: For an $(m \times m)$ matrix $A = [a_{ij}]$, the vech operator is defined as

$$\operatorname{vech} A = \operatorname{vech}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{22} \\ \vdots \\ a_{m2} \\ a_{33} \\ \vdots \\ a_{mm} \end{bmatrix} \quad (\tfrac{1}{2}m(m+1) \times 1),$$

that is, vech stacks the columns of A from the principal diagonal downwards in a column vector. (See Section 7.3 for the rules.)

Vector norm: A function $\|\cdot\|$ attaching a nonnegative real number $\|x\|$ to an $(m \times 1)$ vector x is a norm or a vector norm if the following three conditions are satisfied for all complex $(m \times 1)$ vectors x, y and complex numbers c :

(i) $\|x\| > 0 \quad \text{if} \quad x \neq 0,$

(ii) $\|cx\| = |c|_{\text{abs}}\, \|x\|,$

(iii) $\|x + y\| \leq \|x\| + \|y\| \qquad$ (triangle inequality).

Here $|c|_{\text{abs}}$ denotes the modulus of c, that is, $|c|_{\text{abs}} = \sqrt{c\bar{c}}$ with $\bar{c}$ being the complex conjugate of c. Instead of defining a norm for all complex $(m \times 1)$ vectors, it may be defined for real vectors only. If (i), (ii) and (iii) hold for all real $(m \times 1)$ vectors and real numbers c, $\|\cdot\|$ is called a norm or vector norm over the field of real numbers ($\mathbb{R}$). In that case $|c|_{\text{abs}}$ is the absolute value of c. (See Chapter 8.)

Vector polynomial: The function $p : \mathbb{C}^m \to \mathbb{C}$ (or $\mathbb{R}^m \to \mathbb{R}$) defined by

$$\begin{aligned} p(x_1, \ldots, x_m) &= p_0 + \sum_{i=1}^{m} p_{1,i} x_i + \sum_{i=1}^{m} \sum_{j=1}^{m} p_{2,ij} x_i x_j \\ &\quad + \cdots + \sum_{i=1}^{m} \cdots \sum_{j=1}^{m} p_{d,i \ldots j} x_i \cdots x_j \end{aligned}$$

is a vector polynomial of degree d. For example,

$$p(x_1, x_2) = 2 + 4x_1 - x_2 + 5x_1^2 + \tfrac{1}{2} x_1 x_2$$

is a vector polynomial of degree 2.

Vector product: The vector product of the 3-dimensional vectors $x = (x_1, x_2, x_3)'$ and $y = (y_1, y_2, y_3)'$ is

$$x \times y = [(x_2 y_3 - x_3 y_2), -(x_1 y_3 - x_3 y_1), (x_1 y_2 - x_2 y_1)]'.$$

Vectorization of a matrix: Stacking the rows or columns of a matrix $A = [a_{ij}]$ in a vector is called vectorization of A. Notation:

$$\operatorname{vec} A = \operatorname{vec}(A) = \operatorname{col}(A) \equiv \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ a_{13} \\ \vdots \\ a_{mn} \end{bmatrix} \quad (mn \times 1),$$

$$\operatorname{rvec}(A) = [\operatorname{vec}(A')]',$$

$$\operatorname{row}(A) = \operatorname{vec}(A') = \operatorname{rvec}(A)'.$$

(See Chapter 7 for details.)

Zero matrix: An $(m \times n)$ matrix is a zero or *null matrix*, denoted by $O_{m\times n}$ or simply by 0, if all its elements are zero.

Zero vector: An $(m \times 1)$ or $(1 \times m)$ vector is a zero or *null vector* if all its elements are zero.

References

Note: There exists a tremendous number of books on matrices. Also a substantial number of books on linear algebra and other topics such as econometrics and multivariate statistical methods contain chapters, sections or appendices on matrices. The number of articles with special matrix results is even greater. In the following only those books and articles cited in the text are listed in addition to a number of general and some more special matrix books.

Agaian, S.S. (1975). *Hadamard Matrices and Their Applications.* Springer-Verlag, Berlin.

Aitken, A.C. (1956). *Determinants and Matrices.* 9th ed. Oliver and Boyd, Edinburgh.

Anderson T.W. (1984). *An Introduction to Multivariate Statistical Analysis.* 2nd edn, John Wiley, New York.

Barnett. S. (1979). *Matrix Methods for Engineers and Scientists.* McGraw-Hill, London.

Barnett, S. (1983). *Polynomials and Linear Control Systems.* Marcel Dekker, New York.

Barnett, S. (1984). *Matrices in Control Theory,* (2nd edn). Krieger, Florida.

Barnett, S. (1990). *Matrices: Methods and Applications.* Clarendon Press, Oxford.

Barnett, S. & C. Storey (1970). *Matrix Methods in Stability Theory.* Nelson, London.

Basilevsky, A. (1983). *Applied Matrix Algebra in the Statistical Sciences.* North-Holland, New York.

Beckenbach E.F. & R. Bellman (1961). *Inequalities.* Springer-Verlag, Berlin.

Bellman, R. (1970). *Introduction to Matrix Analysis.* 2d. ed. McGraw-Hill, New York.

Ben-Israel, A. & T.N.E. Greville (1974). *Generalized Inverses: Theory and Applications.* Wiley, New York.

Berman, A., M. Neumann & R.J. Stern (1989). *Non-negative Matrices in Systems Theory.* Wiley, New York.

Berman, A. & R.J. Plemmons (1979). *Non-negative Matrices in the Mathematical Sciences.* Academic Press, New York.

Björck, A., R.J. Plemmons & H. Schneider (eds) (1981). *Large Scale Matrix Problems.* North Holland, Amsterdam.

Boullion, T.L. & P.L. Odell (1971). *Generalized Inverse Matrices.* Wiley. New York.

Bradley, I. & R.L. Meek (1986). *Matrices and Society.* Penguin, Harmondsworth, Middlesex.

Bronson, R. (1989). *Schaum's Outline of Theory and Problems of Matrix Operations.* McGraw-Hill, New York.

Bunch, J.R. & D.J. Rose (eds) (1976). *Sparse Matrix Computations.* Academic Press, New York.

Campbell, S.L. & C.D. Meyer Jr. (1979). *Generalized Inverses of Linear Transformations.* Pitman, London.

Chatelin, F. (1993). *Eigenvalues of Matrices.* Wiley, New York.

Collar, A.R. & A. Simpson (1987). *Matrices and Engineering Dynamics.* Ellis Horwood, Chichester.

Cullen, C.G. (1966). *Matrices and Linear Transformations.* Addison-Wesley, Reading, Mass.

Davis, P.J. (1979). *Circulant Matrices.* Wiley, New York.

Donoghue, W.F. Jr. (1974). *Monotone Matrix Functions and Analytic Continuation.* Springer-Verlag, Berlin.

Duff, I.S., A.M. Erisman & J.K. Reid (1987). *Direct Methods for Sparse Matrices.* Clarendon Press, Oxford.

Fiedler, M. (1986). *Special Matrices and Their Applications in Numerical Mathematics.* Martinus Nijhoff, Dordrecht.

Franklin, J. (1968). *Matrix Theory.* Prentice-Hall, Englewood Cliffs, N.J.

Gantmacher, F.R. (1959a). *The Theory of Matrices.* 2 vols. Chelsea, New York.

Gantmacher, F.R. (1959b). *Applications of the Theory of Matrices.* Interscience, New York.

Gohberg, I., P. Lancaster & L. Rodman (1982). *Matrix Polynomials.* Academic Press, New York.

Gohberg, I., P. Lancaster & L. Rodman (1983). *Matrices and Indefinite Scalar Products.* Birkhäuser-Verlag, Boston.

Golub, G.H. & C.F. Van Loan (1989). *Matrix Computations.* (2nd edn). Johns Hopkins University Press, Baltimore.

Graham, A. (1981). *Kronecker Products and Matrix Calculus with Applications.* Horwood, Chichester, U.K.

Graybill, F.A. (1983). *Matrices with Applications to Statistics.* (2nd edn). Wadsworth, Belmont, California.

Heinig, G. & R. Rost (1984). *Algebraic Methods for Toeplitz-like Matrices and Operators.* Akademie-Verlag, Berlin.

Horn, R.A. & C.R. Johnson (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

Horn, R.A. & C.R. Johnson (1989). *Topics in Matrix Analysis.* Cambridge University Press, Cambridge.

Householder, A.S. (1964). *The Theory of Matrices in Numerical Analysis.* Blaisdell, New York.

Johansen, S. (1995). *Likelihood Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford.

Kim, K.H. (1982). *Boolean Matrix Theory and Applications.* Marcel Dekker, New York.

Kowalsky, H. (1972). *Lineare Algebra.* 6th ed. de Gruyter, Berlin.

Krishnamurthy, E.V. (1985). *Error-free Polynomial Matrix Computations.* Springer-Verlag, New York.

Lancaster, P. (1966). *Lambda-matrices and Vibrating Systems.* Pergamon, Oxford.

Lancaster, P. (1969). *Theory of Matrices.* Academic Press, New York.

Lancaster, P. & M. Tismenetsky (1985). *The Theory of Matrices With Applications.* 2d ed. Academic Press, New York.
Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis.* Springer. Berlin.
MacDuffee, C.C. (1946). *The Theory of Matrices.* Chelsea, New York.
Magnus, R.J. (1988). *Linear Structures.* Griffin, London.
Magnus, J.R. & H. Neudecker (1988). *Matrix Differential Calculus With Applications in Statistics and Econometrics.* Wiley, Chichester.
Marcus, M. & H. Minc (1964). *A Survey of Matrix Theory and Matrix Inequalities.* Allyn and Bacon, Boston.
Miller, K.S. (1987). *Some Eclectic Matrix Theory.* Krieger, Florida.
Minc, H. (1988). *Nonnegative Matrices.* Wiley, New York.
Namboodiri, K. (1984). *Matrix Algebra: An Introduction.* Sage Publications, Beverly Hills.
Nashed, M.Z. (ed.) (1976). *Generalized Inverses and Applications.* Academic Press, New York.
Nering, E. (1963). *Linear Algebra and Matrix Theory.* 2d ed. Wiley, New York.
Neudecker, H. & L. Shuangzhe (1993). "Matrix Trace Inequalities Involving Simple, Kronecker, and Hadamard Products." *Econometric Theory*, 9, 690.
Neudecker, H. & L. Shuangzhe (1995). "Matrix Trace Inequalities Involving Simple, Kronecker, and Hadamard Products – Solution." *Econometric Theory*, 11, 669 - 670.
Newman, M. (1972). *Integral Matrices.* Academic Press, New York.
Ortega, J.M. (1987). *Matrix Theory, a Second Course.* Plenum Press, New York.
Perlis, S. (1952). *Theory of Matrices.* Addison-Wesley, Reading, Mass.
Pissanetsky, S. (1984). *Sparse Matrix Technology.* Academic Press, London.
Pollock, D.S.G. (1979). *The Algebra of Econometrics.* Wiley, Chichester.
Pringle, R.M. & A.A. Rayner (1971). *Generalized Inverse Matrices with Applications to Statistics.* Griffin, London.
Rao, C.R. (1973). *Linear Statistical Inference and its Applications.* 2nd edn, John Wiley, New York.
Rao, C.R. & S.K. Mitra (1971). *Generalized Inverse of Matrices and its Applications.* Wiley, New York.
Rogers, G.S. (1980). *Matrix Derivatives.* Marcel Dekker, New York.
Searle, S.R. (1982). *Matrix Algebra Useful for Statistics.* Wiley, New York.
Seber, G.A.F. (1984). *Multivariate Observations.* Wiley, New York.
Seneta, E. (1973). *Nonnegative Matrices.* Wiley, New York.
Stewart, G.W. (1973). *Introduction to Matrix Computations.* Academic Press, New York.
Suprenenko, D.A. & R.I. Tyshkevich (1968). *Commutative Matrices.* Academic Press, New York.
Turnbull, H.W. (1950). *The Theory of Determinants, Matrices and Invariants.* Blackie, London.
Turnbull, H.W. & A.C. Aitken (1932). *An Introduction to the Theory of Canonical Matrices.* Blackie, London.
Varga, R.S. (1962). *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs, N.J.
Wilkinson, J.H. (1965). *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford.
Zurmühl, R. & S. Falk (1984). *Matrizen und ihre Anwendungen, Teil 1: Grundlagen.* Springer, Berlin.
Zurmühl, R. & S. Falk (1986). *Matrizen und ihre Anwendungen. Teil 2: Numerische Methoden.* Springer, Berlin.

Index