

## MATRIX ANALYSIS AND APPLICATIONS

This balanced and comprehensive study presents the theory, methods and applications of matrix analysis in a new theoretical framework, allowing readers to understand second-order and higher-order matrix analysis in a completely new light.

Alongside the core subjects in matrix analysis, such as singular value analysis, the solution of matrix equations and eigenanalysis, the author introduces new applications and perspectives that are unique to this book. The very topical subjects of gradient analysis and optimization play a central role here. Also included are subspace analysis, projection analysis and tensor analysis, subjects which are often neglected in other books. Having provided a solid foundation to the subject, the author goes on to place particular emphasis on the many applications matrix analysis has in science and engineering, making this book suitable for scientists, engineers and graduate students alike.

XIAN-DA ZHANG is Professor Emeritus in the Department of Automation, at Tsinghua University, Beijing. He was a Distinguished Professor at Xidian University, Xi'an, China – a post awarded by the Ministry of Education of China, and funded by the Ministry of Education of China and the Cheung Kong Scholars Programme – from 1999 to 2002. His areas of research include signal processing, pattern recognition, machine learning and related applied mathematics. He has published over 120 international journal and conference papers, and 7 books in Chinese. He taught the graduate course “Matrix Analysis and Applications” at Tsinghua University from 2004 to 2011.

# MATRIX ANALYSIS AND APPLICATIONS

XIAN-DA ZHANG  
*Tsinghua University, Beijing*



CAMBRIDGE  
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India  
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781108417419](http://www.cambridge.org/9781108417419)

DOI: 10.1017/9781108277587

© Xian-Da Zhang 2017

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2017

Printed in the United Kingdom by Clays, St Ives plc

*A catalogue record for this publication is available from the British Library.*

ISBN 978-1-108-41741-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

To John Zhang, Ellen Zhang and Andrew Wei

# Contents

<i>Preface</i>	page xvii
<i>Notation</i>	xxi
<i>Abbreviations</i>	xxxii
<i>Algorithms</i>	xxxiv
<b>PART I MATRIX ALGEBRA</b>	<b>1</b>
<b>1 Introduction to Matrix Algebra</b>	<b>3</b>
1.1 Basic Concepts of Vectors and Matrices	3
1.1.1 Vectors and Matrices	3
1.1.2 Basic Vector Calculus	6
1.1.3 Basic Matrix Calculus	8
1.1.4 Linear Independence of Vectors	11
1.1.5 Matrix Functions	11
1.2 Elementary Row Operations and Applications	13
1.2.1 Elementary Row Operations	13
1.2.2 Gauss Elimination Methods	16
1.3 Sets, Vector Subspaces and Linear Mapping	20
1.3.1 Sets	20
1.3.2 Fields and Vector Spaces	22
1.3.3 Linear Mapping	24
1.4 Inner Products and Vector Norms	27
1.4.1 Inner Products of Vectors	27
1.4.2 Norms of Vectors	28
1.4.3 Similarity Comparison Between Vectors	32
1.4.4 Banach Space, Euclidean Space, Hilbert Space	35
1.4.5 Inner Products and Norms of Matrices	36
1.5 Random Vectors	40
1.5.1 Statistical Interpretation of Random Vectors	41
1.5.2 Gaussian Random Vectors	44

1.6	Performance Indexes of Matrices	47
1.6.1	Quadratic Forms	47
1.6.2	Determinants	49
1.6.3	Matrix Eigenvalues	52
1.6.4	Matrix Trace	54
1.6.5	Matrix Rank	56
1.7	Inverse Matrices and Pseudo-Inverse Matrices	59
1.7.1	Definition and Properties of Inverse Matrices	59
1.7.2	Matrix Inversion Lemma	60
1.7.3	Inversion of Hermitian Matrices	61
1.7.4	Left and Right Pseudo-Inverse Matrices	63
1.8	Moore–Penrose Inverse Matrices	65
1.8.1	Definition and Properties	65
1.8.2	Computation of Moore–Penrose Inverse Matrix	69
1.9	Direct Sum and Hadamard Product	71
1.9.1	Direct Sum of Matrices	71
1.9.2	Hadamard Product	72
1.10	Kronecker Products and Khatri–Rao Product	75
1.10.1	Kronecker Products	75
1.10.2	Generalized Kronecker Products	77
1.10.3	Khatri–Rao Product	78
1.11	Vectorization and Matricization	79
1.11.1	Vectorization and Commutation Matrix	79
1.11.2	Matricization of a Vector	82
1.11.3	Properties of Vectorization Operator	83
1.12	Sparse Representations	84
1.12.1	Sparse Vectors and Sparse Representations	84
1.12.2	Sparse Representation of Face Recognition	86
	Exercises	87
<b>2</b>	<b>Special Matrices</b>	<b>95</b>
2.1	Hermitian Matrices	95
2.2	Idempotent Matrix	96
2.3	Permutation Matrix	98
2.3.1	Permutation Matrix and Exchange Matrix	98
2.3.2	Generalized Permutation Matrix	101
2.4	Orthogonal Matrix and Unitary Matrix	104
2.5	Band Matrix and Triangular Matrix	107
2.5.1	Band Matrix	107
2.5.2	Triangular Matrix	107
2.6	Summing Vector and Centering Matrix	109
2.6.1	Summing Vector	110

	2.6.2	Centering Matrix	111
2.7		Vandermonde Matrix and Fourier Matrix	112
	2.7.1	Vandermonde Matrix	112
	2.7.2	Fourier Matrix	114
	2.7.3	Index Vectors	116
	2.7.4	FFT Algorithm	118
2.8		Hadamard Matrix	121
2.9		Toeplitz Matrix	123
	2.9.1	Symmetric Toeplitz Matrix	123
	2.9.2	Discrete Cosine Transform of Toeplitz Matrix	125
		Exercises	126
<b>3</b>		<b>Matrix Differential</b>	129
	3.1	Jacobian Matrix and Gradient Matrix	129
	3.1.1	Jacobian Matrix	130
	3.1.2	Gradient Matrix	132
	3.1.3	Calculation of Partial Derivative and Gradient	133
	3.2	Real Matrix Differential	139
	3.2.1	Calculation of Real Matrix Differential	139
	3.2.2	Jacobian Matrix Identification	141
	3.2.3	Jacobian Matrix of Real Matrix Functions	147
	3.3	Real Hessian Matrix and Identification	150
	3.3.1	Real Hessian Matrix	150
	3.3.2	Real Hessian Matrix Identification	152
	3.4	Complex Gradient Matrices	157
	3.4.1	Holomorphic Function and Complex Partial Derivative	157
	3.4.2	Complex Matrix Differential	161
	3.4.3	Complex Gradient Matrix Identification	169
	3.5	Complex Hessian Matrices and Identification	174
	3.5.1	Complex Hessian Matrices	175
	3.5.2	Complex Hessian Matrix Identification	177
		Exercises	179
		<b>PART II MATRIX ANALYSIS</b>	181
<b>4</b>		<b>Gradient Analysis and Optimization</b>	183
	4.1	Real Gradient Analysis	183
	4.1.1	Stationary Points and Extreme Points	184
	4.1.2	Real Gradient Analysis of $f(\mathbf{x})$	186
	4.1.3	Real Gradient Analysis of $f(\mathbf{X})$	188
	4.2	Gradient Analysis of Complex Variable Function	190
	4.2.1	Extreme Point of Complex Variable Function	190

	4.2.2	Complex Gradient Analysis	195
4.3		Convex Sets and Convex Function Identification	198
	4.3.1	Standard Constrained Optimization Problems	198
	4.3.2	Convex Sets and Convex Functions	200
	4.3.3	Convex Function Identification	203
4.4		Gradient Methods for Smooth Convex Optimization	205
	4.4.1	Gradient Method	205
	4.4.2	Conjugate Gradient Method	210
	4.4.3	Convergence Rates	215
4.5		Nesterov Optimal Gradient Method	217
	4.5.1	Lipschitz Continuous Function	217
	4.5.2	Nesterov Optimal Gradient Algorithms	220
4.6		Nonsmooth Convex Optimization	223
	4.6.1	Subgradient and Subdifferential	224
	4.6.2	Proximal Operator	228
	4.6.3	Proximal Gradient Method	232
4.7		Constrained Convex Optimization	237
	4.7.1	Lagrange Multiplier Method	237
	4.7.2	Penalty Function Method	238
	4.7.3	Augmented Lagrange Multiplier Method	240
	4.7.4	Lagrangian Dual Method	242
	4.7.5	Karush–Kuhn–Tucker Conditions	244
	4.7.6	Alternating Direction Method of Multipliers	248
4.8		Newton Methods	251
	4.8.1	Newton Method for Unconstrained Optimization	251
	4.8.2	Newton Method for Constrained Optimization	254
4.9		Original–Dual Interior-Point Method	260
	4.9.1	Original–Dual Problems	260
	4.9.2	First-Order Original–Dual Interior-Point Method	261
	4.9.3	Second-Order Original–Dual Interior-Point Method	263
		Exercises	268
<b>5</b>		<b>Singular Value Analysis</b>	<b>271</b>
	5.1	Numerical Stability and Condition Number	271
	5.2	Singular Value Decomposition (SVD)	274
	5.2.1	Singular Value Decomposition	274
	5.2.2	Properties of Singular Values	277
	5.2.3	Rank-Deficient Least Squares Solutions	280
	5.3	Product Singular Value Decomposition (PSVD)	283
	5.3.1	PSVD Problem	283
	5.3.2	Accurate Calculation of PSVD	284
	5.4	Applications of Singular Value Decomposition	285



	xi	
5.4.1	Static Systems	286
5.4.2	Image Compression	287
5.5	Generalized Singular Value Decomposition (GSVD)	289
5.5.1	Definition and Properties	289
5.5.2	Algorithms for GSVD	292
5.5.3	Two Application Examples of GSVD	294
5.6	Low-Rank-Sparse Matrix Decomposition	296
5.6.1	Matrix Decomposition Problems	297
5.6.2	Singular Value Thresholding	298
5.6.3	Robust Principal Component Analysis	300
5.7	Matrix Completion	302
5.7.1	Matrix Completion Problems	303
5.7.2	Matrix Completion Model and Incoherence	305
5.7.3	Singular Value Thresholding Algorithm	306
5.7.4	Fast and Accurate Matrix Completion	308
	Exercises	312
<b>6</b>	<b>Solving Matrix Equations</b>	<b>315</b>
6.1	Least Squares Method	316
6.1.1	Ordinary Least Squares Methods	316
6.1.2	Properties of Least Squares Solutions	317
6.1.3	Data Least Squares	320
6.2	Tikhonov Regularization and Gauss-Seidel Method	321
6.2.1	Tikhonov Regularization	321
6.2.2	Regularized Gauss-Seidel Method	324
6.3	Total Least Squares (TLS) Methods	328
6.3.1	TLS Problems	328
6.3.2	TLS Solution	329
6.3.3	Performances of TLS Solution	333
6.3.4	Generalized Total Least Squares	335
6.3.5	Total Least Squares Fitting	337
6.3.6	Total Maximum Likelihood Method	342
6.4	Constrained Total Least Squares	344
6.4.1	Constrained Total Least Squares Method	345
6.4.2	Harmonic Superresolution	347
6.4.3	Image Restoration	348
6.5	Subspace Method for Solving Blind Matrix Equations	350
6.6	Nonnegative Matrix Factorization: Optimization Theory	353
6.6.1	Nonnegative Matrices	353
6.6.2	Nonnegativity and Sparsity Constraints	355
6.6.3	Nonnegative Matrix Factorization Model	356
6.6.4	Divergences and Deformed Logarithm	361

6.7	Nonnegative Matrix Factorization: Optimization Algorithms	366
6.7.1	Multiplication Algorithms	366
6.7.2	Nesterov Optimal Gradient Algorithm	372
6.7.3	Alternating Nonnegative Least Squares	374
6.7.4	Quasi-Newton Method	377
6.7.5	Sparse Nonnegative Matrix Factorization	378
6.8	Sparse Matrix Equation Solving: Optimization Theory	381
6.8.1	$\ell_1$ -Norm Minimization	381
6.8.2	Lasso and Robust Linear Regression	384
6.8.3	Mutual Coherence and RIP Conditions	387
6.8.4	Relation to Tikhonov Regularization	389
6.8.5	Gradient Analysis of $\ell_1$ -Norm Minimization	390
6.9	Sparse Matrix Equation Solving: Optimization Algorithms	391
6.9.1	Basis Pursuit Algorithms	391
6.9.2	First-Order Augmented Lagrangian Algorithm	394
6.9.3	Barzilai–Borwein Gradient Projection Algorithm	394
6.9.4	ADMM Algorithms for Lasso Problems	397
6.9.5	LARS Algorithms for Lasso Problems	398
6.9.6	Covariance Graphical Lasso Method	400
6.9.7	Homotopy Algorithm	402
6.9.8	Bregman Iteration Algorithms	403
	Exercises	409
<b>7</b>	<b>Eigenanalysis</b>	<b>413</b>
7.1	Eigenvalue Problem and Characteristic Equation	413
7.1.1	Eigenvalue Problem	413
7.1.2	Characteristic Polynomial	415
7.2	Eigenvalues and Eigenvectors	416
7.2.1	Eigenvalues	416
7.2.2	Eigenvectors	419
7.3	Similarity Reduction	422
7.3.1	Similarity Transformation of Matrices	423
7.3.2	Similarity Reduction of Matrices	426
7.3.3	Similarity Reduction of Matrix Polynomials	430
7.4	Polynomial Matrices and Balanced Reduction	434
7.4.1	Smith Normal Forms	434
7.4.2	Invariant Factor Method	437
7.4.3	Conversion of Jordan Form and Smith Form	441
7.4.4	Finding Smith Blocks from Jordan Blocks	442
7.4.5	Finding Jordan Blocks from Smith Blocks	443
7.5	Cayley–Hamilton Theorem with Applications	446
7.5.1	Cayley–Hamilton Theorem	446

	xiii	
7.5.2	Computation of Inverse Matrices	448
7.5.3	Computation of Matrix Powers	450
7.5.4	Calculation of Matrix Exponential Functions	452
7.6	Application Examples of Eigenvalue Decomposition	455
7.6.1	Pisarenko Harmonic Decomposition	455
7.6.2	Discrete Karhunen–Loeve Transformation	458
7.6.3	Principal Component Analysis	461
7.7	Generalized Eigenvalue Decomposition (GEVD)	463
7.7.1	Generalized Eigenvalue Decomposition	463
7.7.2	Total Least Squares Method for GEVD	467
7.7.3	Application of GEVD: ESPRIT	468
7.7.4	Similarity Transformation in GEVD	471
7.8	Rayleigh Quotient	474
7.8.1	Definition and Properties of Rayleigh Quotient	474
7.8.2	Rayleigh Quotient Iteration	475
7.8.3	Algorithms for Rayleigh Quotient	476
7.9	Generalized Rayleigh Quotient	478
7.9.1	Definition and Properties	478
7.9.2	Effectiveness of Class Discrimination	480
7.9.3	Robust Beamforming	482
7.10	Quadratic Eigenvalue Problems	484
7.10.1	Description of Quadratic Eigenvalue Problems	484
7.10.2	Solving Quadratic Eigenvalue Problems	486
7.10.3	Application Examples	490
7.11	Joint Diagonalization	495
7.11.1	Joint Diagonalization Problems	495
7.11.2	Orthogonal Approximate Joint Diagonalization	497
7.11.3	Nonorthogonal Approximate Joint Diagonalization	500
	Exercises	503
<b>8</b>	<b>Subspace Analysis and Tracking</b>	<b>511</b>
8.1	General Theory of Subspaces	511
8.1.1	Bases of Subspaces	511
8.1.2	Disjoint Subspaces and Orthogonal Complement	513
8.2	Column Space, Row Space and Null Space	516
8.2.1	Definitions and Properties	516
8.2.2	Subspace Basis Construction	520
8.2.3	SVD-Based Orthonormal Basis Construction	522
8.2.4	Basis Construction of Subspaces Intersection	525
8.3	Subspace Methods	526
8.3.1	Signal Subspace and Noise Subspace	526
8.3.2	Multiple Signal Classification (MUSIC)	529

xiv

	8.3.3	Subspace Whitening	531
8.4		Grassmann Manifold and Stiefel Manifold	532
	8.4.1	Equivalent Subspaces	532
	8.4.2	Grassmann Manifold	533
	8.4.3	Stiefel Manifold	535
8.5		Projection Approximation Subspace Tracking (PAST)	536
	8.5.1	Basic PAST Theory	537
	8.5.2	PAST Algorithms	540
8.6		Fast Subspace Decomposition	542
	8.6.1	Rayleigh–Ritz Approximation	542
	8.6.2	Fast Subspace Decomposition Algorithm	544
		Exercises	546
<b>9</b>		<b>Projection Analysis</b>	<b>551</b>
	9.1	Projection and Orthogonal Projection	551
		9.1.1 Projection Theorem	552
		9.1.2 Mean Square Estimation	554
	9.2	Projectors and Projection Matrices	556
		9.2.1 Projector and Orthogonal Projector	556
		9.2.2 Projection Matrices	558
		9.2.3 Derivatives of Projection Matrix	561
	9.3	Updating of Projection Matrices	562
		9.3.1 Updating Formulas for Projection Matrices	562
		9.3.2 Prediction Filters	564
		9.3.3 Updating of Lattice Adaptive Filter	567
	9.4	Oblique Projector of Full Column Rank Matrix	570
		9.4.1 Definition and Properties of Oblique Projectors	571
		9.4.2 Geometric Interpretation of Oblique Projectors	575
		9.4.3 Recursion of Oblique Projectors	578
	9.5	Oblique Projector of Full Row Rank Matrices	579
		9.5.1 Definition and Properties	579
		9.5.2 Calculation of Oblique Projection	581
		9.5.3 Applications of Oblique Projectors	583
		Exercises	585
		<b>PART III HIGHER-ORDER MATRIX ANALYSIS</b>	<b>587</b>
<b>10</b>		<b>Tensor Analysis</b>	<b>589</b>
	10.1	Tensors and their Presentation	589
		10.1.1 Tensors	589
		10.1.2 Tensor Representation	592
	10.2	Vectorization and Matricization of Tensors	597

10.2.1	Vectorization and Horizontal Unfolding	597
10.2.2	Longitudinal Unfolding of Tensors	601
10.3	Basic Algebraic Operations of Tensors	606
10.3.1	Inner Product, Norm and Outer Product	606
10.3.2	Mode- $n$ Product of Tensors	608
10.3.3	Rank of Tensor	612
10.4	Tucker Decomposition of Tensors	614
10.4.1	Tucker Decomposition (Higher-Order SVD)	615
10.4.2	Third-Order SVD	617
10.4.3	Alternating Least Squares Algorithms	621
10.5	Parallel Factor Decomposition of Tensors	625
10.5.1	Bilinear Model	625
10.5.2	Parallel Factor Analysis	627
10.5.3	Uniqueness Condition	635
10.5.4	Alternating Least Squares Algorithm	637
10.6	Applications of Low-Rank Tensor Decomposition	641
10.6.1	Multimodal Data Fusion	642
10.6.2	Fusion of Multimodal Brain Images	644
10.6.3	Process Monitoring	646
10.6.4	Note on Other Applications	648
10.7	Tensor Eigenvalue Decomposition	649
10.7.1	Tensor–Vector Products	649
10.7.2	Determinants and Eigenvalues of Tensors	651
10.7.3	Generalized Tensor Eigenvalues Problems	656
10.7.4	Orthogonal Decomposition of Symmetric Tensors	658
10.8	Preprocessing and Postprocessing	659
10.8.1	Centering and Scaling of Multi-Way Data	660
10.8.2	Compression of Data Array	661
10.9	Nonnegative Tensor Decomposition Algorithms	664
10.9.1	Multiplication Algorithm	664
10.9.2	ALS Algorithms	667
10.10	Tensor Completion	670
10.10.1	Simultaneous Tensor Decomposition and Completion	671
10.10.2	Smooth PARAFAC Tensor Completion	674
10.11	Software	676
	Exercises	678
	<i>References</i>	681
	<i>Index</i>	708

## Preface

Linear algebra is a vast field of fundamental importance in most areas of pure (and applied) mathematics, while matrices are a key tool for the researchers, scientists, engineers and graduate students majoring in the science and engineering disciplines.

From the viewpoint of applications, matrix analysis provides a powerful mathematical modeling and computational framework for posing and solving important scientific and engineering problems. It is no exaggeration to say that matrix analysis is one of the most creative and flexible mathematical tools and that it plays an irreplaceable role in physics, mechanics, signal and information processing, wireless communications, machine learning, computer vision, automatic control, system engineering, aerospace, bioinformatics, medical image processing and many other disciplines, and it effectively supports research in them all. At the same time, novel applications in these disciplines have spawned a number of new results and methods of matrix analysis, such as quadratic eigenvalue problems, joint diagonalization, sparse representation and compressed sensing, matrix completion, nonnegative matrix factorization, tensor analysis and so on.

### Goal of the Book

The main goal of this book is to help the reader develop the skills and background needed to recognize, formulate and solve linear algebraic problems by presenting systematically the theory, methods and applications of matrix analysis.

A secondary goal is to help the reader understand some recent applications, perspectives and developments in matrix analysis.

### Structure of the Book

In order to provide a balanced and comprehensive account of the subject, this book covers the core theory and methods in matrix analysis, and places particular emphasis on its typical applications in various science and engineering disciplines. The book consists of ten chapters, spread over three parts.

Part I is on matrix algebra: it contains Chapters 1 through 3 and focuses on the necessary background material. Chapter 1 is an introduction to matrix algebra that is devoted to basic matrix operations. This is followed by a description of the vec-

torization of matrices, the representation of vectors as matrices, i.e. matricization, and the application of sparse matrices to face recognition. Chapter 2 presents some special matrices used commonly in matrix analysis. Chapter 3 presents the matrix differential, which is an important tool in optimization.

Part II is on matrix analysis: this is the heart of the book, and deals with the topics that are most frequently needed. It covers both theoretical and practical aspects and consists of six chapters, as follows.

Chapter 4 is devoted to the gradient analysis of matrices, with applications in smooth and nonsmooth convex optimization, constrained convex optimization, Newton's algorithm and the original–dual interior-point method.

In Chapter 5 we describe the singular value analysis of matrices, including singular value decomposition, generalized singular value decomposition, low-rank sparse matrix decomposition and matrix completion.

Researchers, scientists, engineers and graduate students from a wide variety of disciplines often have to use matrices for modeling purposes and to solve the resulting matrix equations. Chapter 6 focuses on ways to solve such equations and includes the Tikhonov regularization method, the total least squares method, the constrained total least squares method, nonnegative matrix factorization and the solution of sparse matrix equations.

Chapter 7 deals with eigenvalue decomposition, matrix reduction, generalized eigenvalue decomposition, the Rayleigh quotient, the generalized Rayleigh quotient, quadratic eigenvalue problems and joint diagonalization.

Chapter 8 is devoted to subspace analysis methods and subspace tracking algorithms in adaptive signal processing.

Chapter 9 focuses on orthogonal and oblique projections with their applications.

Part III is on higher-order matrix analysis and consists simply of Chapter 10. In it, matrix analysis is extended from the second-order case to higher orders via a presentation of the basic algebraic operations, representation as matrices, Tuckey decomposition, parallel factor decomposition, eigenvalue decomposition of tensors, nonnegative tensor decomposition and tensor completion, together with applications.

### Features of the Book

The book introduces a novel theoretical framework for matrix analysis by dividing it into second-order matrix analysis (including gradient analysis, singular value analysis, eigenanalysis, subspace analysis and projection analysis) and higher-order matrix analysis (tensor analysis).

Gradient analysis and optimization play an important role in the book. This is a very topical subject and is central to many modern applications (such as communications, signal processing, pattern recognition, machine learning, radar, big data analysis, multimodal brain image fusion etc.) though quite classical in origin.

Some more contemporary topics of matrix analysis such as subspace analysis,

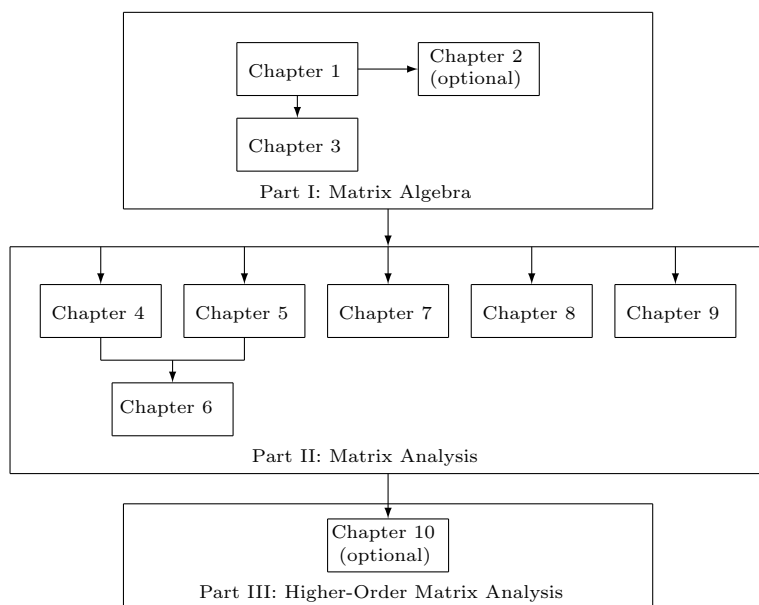
projection analysis and tensor analysis, and which are often missing from other books, are included in our text.

Particular emphasis is placed on typical applications of matrix methods in science and engineering. The 80 algorithms for which summaries are given should help readers learn how to conduct computer experiments using related matrix analysis in their studies and research.

In order to make these methods easy to understand and master, this book adheres to the principle of both interpreting physics problems in terms of mathematics, and mathematical results in terms of physical ideas. Thus some typical or important matrix analysis problems are introduced by modeling a problem from physics, while some important mathematical results are explained and understood by revealing their physical meaning.

### Reading the Book

The following diagram gives a schematic organization of this book to illustrate the chapter dependences.



Chapters 2 and 10 are optional. In particular, Chapter 10 is specifically devoted to readers involved in multi-channel or multi-way data analysis and processing.

### Intended Readership

Linear algebra and matrix analysis are used in a very wide range of subjects including physics, statistics, computer science, economics, information science and



technology (including signal and image processing, communications, automation control, system engineering and pattern recognition), artificial intelligence, bioinformatics, biomedical engineering, to name just a selection. This book is dedicated to providing individuals in those disciplines with a solid foundation of the fundamental skills needed to develop and apply linear algebra and matrix analysis methods in their work.

The only background required of the reader is a good knowledge of advanced calculus, so the book will be suitable for graduate students in science and engineering.

### Acknowledgments

The contents of this book reflect the author's collaboration with his own graduate students Jian Li, Zi-Zhe Ding, Yong-Tao Su, Xi-Lin Li, Heng Yang, Xi-Kai Zhao, Qi Lv, Qiu-Beng Gao, Li Zhang, Jian-Jiang Ding, Lu Wu, Feng Zhu, Ling Zhang, Dong-Xia Chang, De-Guang Xie, Chun-Yu Peng, Dao-Ming Zhang, Kun Wang, Xi-Yuan Wang, Zhong Chen, Tian-Xiang Luan, Liang Zheng, Yong Zhang, Yan-Yi Rao, all at Tsinghua University, and Shun-Tian Lou, Xiao-Long Zhu, Ji-Ming Ye, Fang-Ming Han, Xiao-Jun Li, Jian-Feng Chen at Xidian University.

Since 2004 we have taught graduate courses on matrix analysis and applications at Tsinghua University. Over the years I have benefited from keen interest, feedback and suggestions from many people, including my own graduate students, and students in our courses. I wish to thank Dr. Fang-Ming Han for his contribution to co-teaching and then teaching these courses, and Xi-Lin Li, Lu Wu, Dong-Xia Chang, Kun Wang, Zhong Chen, Xi-Yuan Wang and Yan-Yi Rao for their assistance with the teaching.

Kun Wang, Zhong Chen, Liang Zheng and Xi-Yuan Wang kindly provided some illustrations in the book.

I am grateful to the countless researchers in linear algebra, matrix analysis, information science and technology for their original contributions and to the anonymous reviewers for their critical comments and suggestions, which have greatly improved the text.

I am most grateful to the Commissioning Editor, David Liu, the Content Manager, Esther Miguéliz, and the copyeditor, Susan Parkinson, for their patience, understanding, suggestions and high-quality content management and copyediting in the course of the book's writing and publication.

This book uses some of the contents and materials of my book *Matrix Analysis and Applications* (Second Edition in Chinese, Tsinghua University Press, 2013).

Finally, I am grateful to my wife Xiao-Ying Tang, my son Yuan-Sheng Zhang, my daughter-in-law Lin Yan, my daughter Ye-Wei Zhang, my son-in-law Wei Wei for their support and encouragement in this project.

## Notation

### Sets

$\mathbb{R}$	real numbers
$\mathbb{R}^n$	real $n$ -vectors ( $n \times 1$ real matrices)
$\mathbb{R}^{m \times n}$	real $m \times n$ matrices
$\mathbb{R}[x]$	real polynomials
$\mathbb{R}[x]^{m \times n}$	real $m \times n$ polynomial matrices
$\mathbb{R}^{I \times J \times K}$	real third-order tensors
$\mathbb{R}^{I_1 \times \cdots \times I_N}$	real $N$ th-order tensor
$\mathbb{R}_+$	nonnegative real numbers, nonnegative orthant
$\mathbb{R}_{++}$	positive real numbers
$\mathbb{C}$	complex numbers
$\mathbb{C}^n$	complex $n$ -vectors
$\mathbb{C}^{m \times n}$	complex $m \times n$ matrices
$\mathbb{C}[x]$	complex polynomials
$\mathbb{C}[x]^{m \times n}$	complex $m \times n$ polynomial matrices
$\mathbb{C}^{I \times J \times K}$	complex third-order tensors
$\mathbb{C}^{I_1 \times \cdots \times I_N}$	complex $N$ th-order tensors
$\mathbb{K}$	real or complex numbers
$\mathbb{K}^n$	real or complex $n$ -vectors
$\mathbb{K}^{m \times n}$	real or complex $m \times n$ matrices
$\mathbb{K}^{I \times J \times K}$	real or complex third-order tensors
$\mathbb{K}^{I_1 \times \cdots \times I_N}$	real or complex $N$ th-order tensors
$\mathbb{Z}$	integers
$\mathbb{Z}_+$	nonnegative integers

**Sets** (*continued*)

$\mathbb{S}^{n \times n}$	symmetric $n \times n$ matrices
$\mathbb{S}_+^{n \times n}$	symmetric positive semi-definite $n \times n$ matrices
$\mathbb{S}_{++}^{n \times n}$	symmetric positive definite $n \times n$ matrices
$\mathbb{S}^{[m, n]}$	symmetric $m$ th-order $n$ -dimensional tensors $\mathcal{A}^{I_1 \times \cdots \times I_m}, I_1 = \cdots = I_n$
$\mathbb{S}_+^{[m, n]}$	symmetric $m$ th-order $n$ -dimensional nonnegative tensors
$\forall$	for all
$x \in A$	$x$ belongs to the set $A$ , i.e. $x$ is an element of $A$
$x \notin A$	$x$ is not an element of the set $A$
$U \mapsto V$	$U$ maps to $V$
$U \rightarrow W$	$U$ transforms to $W$
$\ni$	such that
$\exists$	exists
$A \Rightarrow B$	$A$ implies $B$
$A \subseteq B$	$A$ is a subset of $B$
$A \subset B$	$A$ is a proper subset of $B$
$A \cup B$	union of sets $A$ and $B$
$A \cap B$	intersection of sets $A$ and $B$
$A + B$	sum set of sets $A$ and $B$
$A - B$	set-theoretic difference of sets $A$ and $B$
$X \setminus A$	complement of the set $A$ in the set $X$
$X_1 \times \cdots \times X_n$	Cartesian product of sets $X_1, \dots, X_n$
$\mathcal{L}$	linear manifold
$\text{Gr}(n, r)$	Grassmann manifold
$\text{St}(n, r)$	Stiefel manifold
$O_r$	orthogonal group
$S^\perp$	orthogonal complement of the subspace $S$
$\mathcal{K}^m(\mathbf{A}, \mathbf{f})$	order- $m$ Krylov subspace generated by $\mathbf{A}$ and $\mathbf{f}$
$\text{Col}(\mathbf{A})$	column space of the matrix $\mathbf{A}$
$\text{Ker}(\mathbf{A})$	kernel space of the matrix $\mathbf{A}$
$\text{Null}(\mathbf{A})$	null space of the matrix $\mathbf{A}$
$\text{nullity}(\mathbf{A})$	nullity of the matrix $\mathbf{A}$
$\text{Range}(\mathbf{A})$	range space of the matrix $\mathbf{A}$

**Sets** (*continued*)

$\text{Row}(\mathbf{A})$	row space of the matrix $\mathbf{A}$
$\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m)$	span of vectors $\mathbf{a}_1, \dots, \mathbf{a}_m$

**Vectors**

$\mathbf{x}^*$	conjugate of the vector $\mathbf{x}$
$\mathbf{x}^T$	transpose of the vector $\mathbf{x}$
$\mathbf{x}^H$	conjugate transpose (Hermitian conjugate) of the vector $\mathbf{x}$
$\mathcal{L}(\mathbf{u})$	linear transform of the vector $\mathbf{u}$
$\ \mathbf{x}\ _0$	$\ell_0$ -norm: the number of nonzero entries in the vector $\mathbf{x}$
$\ \mathbf{x}\ _1$	$\ell_1$ -norm of the vector $\mathbf{x}$
$\ \mathbf{x}\ _2$	Euclidean norm of the vector $\mathbf{x}$
$\ \mathbf{x}\ _p$	$\ell_p$ -norm or Hölder norm of the vector $\mathbf{x}$
$\ \mathbf{x}\ _*$	nuclear norm of the vector $\mathbf{x}$
$\ \mathbf{x}\ _\infty$	$\ell_\infty$ -norm of the vector $\mathbf{x}$
$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y}$	inner product of vectors $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{x} \circ \mathbf{y} = \mathbf{x} \mathbf{y}^H$	outer product of vectors $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{x} \perp \mathbf{y}$	orthogonality of vectors $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{x} > \mathbf{0}$	positive vector, with components $x_i > 0, \forall i$
$\mathbf{x} \geq \mathbf{0}$	nonnegative vector, with components $x_i \geq 0, \forall i$
$\mathbf{x} \geq \mathbf{y}$	vector elementwise inequality $x_i \geq y_i, \forall i$
$\text{unvec}(\mathbf{x})$	matricization of the column vector $\mathbf{x}$
$\text{unrvec}(\mathbf{x})$	row matricization of the column vector $\mathbf{x}$
$\theta_i^{(m)}, \mathbf{y}_i^{(m)}$	Rayleigh–Ritz (RR) values, RR vectors
$(\theta_i^{(m)}, \mathbf{y}_i^{(m)})$	Ritz pair

**Matrices**

$\mathbf{A} \in \mathbb{R}^{m \times n}$	real $m \times n$ matrix $\mathbf{A}$
$\mathbf{A} \in \mathbb{C}^{m \times n}$	complex $m \times n$ matrix $\mathbf{A}$
$\mathbf{A}[x] \in \mathbb{R}[x]^{m \times n}$	real $m \times n$ polynomial matrix $\mathbf{A}$
$\mathbf{A}[x] \in \mathbb{C}[x]^{m \times n}$	complex $m \times n$ polynomial matrix $\mathbf{A}$
$\mathbf{A}^*$	conjugate of $\mathbf{A}$
$\mathbf{A}^T$	transpose of $\mathbf{A}$

**Matrices** (*continued*)

$\mathbf{A}^H$	conjugate transpose (Hermitian conjugate) of $\mathbf{A}$
$(\mathbf{A}, \mathbf{B})$	matrix pencil
$\det(\mathbf{A}),  \mathbf{A} $	determinant of $\mathbf{A}$
$\text{tr}(\mathbf{A})$	trace of $\mathbf{A}$
$\text{rank}(\mathbf{A})$	rank of $\mathbf{A}$
$\text{eig}(\mathbf{A})$	eigenvalues of the Hermitian matrix $\mathbf{A}$
$\lambda_i(\mathbf{A})$	$i$ th eigenvalue of the Hermitian matrix $\mathbf{A}$
$\lambda_{\max}(\mathbf{A})$	maximum eigenvalue(s) of the Hermitian matrix $\mathbf{A}$
$\lambda_{\min}(\mathbf{A})$	minimum eigenvalue(s) of the Hermitian matrix $\mathbf{A}$
$\lambda(\mathbf{A}, \mathbf{B})$	generalized eigenvalue of the matrix pencil $(\mathbf{A}, \mathbf{B})$
$\sigma_i(\mathbf{A})$	$i$ th singular value of $\mathbf{A}$
$\sigma_{\max}(\mathbf{A})$	maximum singular value(s) of $\mathbf{A}$
$\sigma_{\min}(\mathbf{A})$	minimum singular value(s) of $\mathbf{A}$
$\rho(\mathbf{A})$	spectral radius of $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of the nonsingular matrix $\mathbf{A}$
$\mathbf{A}^\dagger$	Moore–Penrose inverse of $\mathbf{A}$
$\mathbf{A} \succ 0$	positive definite matrix $\mathbf{A}$
$\mathbf{A} \succeq 0$	positive semi-definite matrix $\mathbf{A}$
$\mathbf{A} \prec 0$	negative definite matrix $\mathbf{A}$
$\mathbf{A} \preceq 0$	negative semi-definite matrix $\mathbf{A}$
$\mathbf{A} > \mathbf{O}$	positive (or elementwise positive) matrix $\mathbf{A}$
$\mathbf{A} \geq \mathbf{O}$	nonnegative (or elementwise nonnegative) matrix $\mathbf{A}$
$\mathbf{A} \geq \mathbf{B}$	matrix elementwise inequality $a_{ij} \geq b_{ij}, \forall i, j$
$\ \mathbf{A}\ _1$	maximum absolute column-sum norm of $\mathbf{A}$
$\ \mathbf{A}\ _\infty$	maximum absolute row-sum norm of $\mathbf{A}$
$\ \mathbf{A}\ _{\text{spec}}$	spectrum norm of $\mathbf{A}$ : $\sigma_{\max}(\mathbf{A})$
$\ \mathbf{A}\ _F$	Frobenius norm of $\mathbf{A}$
$\ \mathbf{A}\ _\infty$	max norm of $\mathbf{A}$ : the absolute maximum of all entries of $\mathbf{A}$
$\ \mathbf{A}\ _{\mathbf{G}}$	Mahalanobis norm of $\mathbf{A}$
$\text{vec } \mathbf{A}$	column vectorization of $\mathbf{A}$
$\text{rvec } \mathbf{A}$	row vectorization of $\mathbf{A}$
$\text{off}(\mathbf{A})$	off function of $\mathbf{A} = [a_{ij}]$ : $\sum_{i=1}^m, i \neq j} \sum_{j=1}^n  a_{ij} ^2$
$\text{diag}(\mathbf{A})$	diagonal function of $\mathbf{A} = [a_{ij}]$ : $\sum_{i=1}^n  a_{ii} ^2$

### Matrices (*continued*)

$\text{diag}(\mathbf{A})$	diagonal vector of $\mathbf{A} = [a_{ij}]$ : $[a_{11}, \dots, a_{nn}]^T$
$\mathbf{Diag}(\mathbf{A})$	diagonal matrix of $\mathbf{A} = [a_{ij}]$ : $\mathbf{Diag}(a_{11}, \dots, a_{nn})$
$\langle \mathbf{A}, \mathbf{B} \rangle$	inner product of $\mathbf{A}$ and $\mathbf{B}$ : $(\text{vec } \mathbf{A})^H \text{vec } \mathbf{B}$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \odot \mathbf{B}$	Khatri–Rao product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} * \mathbf{B}$	Hadamard product of matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \oplus \mathbf{B}$	direct sum of matrices $\mathbf{A}$ and $\mathbf{B}$
$\{\mathbf{A}\}_N$	matrix group consisting of matrices $\mathbf{A}_1, \dots, \mathbf{A}_N$
$\{\mathbf{A}\}_N \otimes \mathbf{B}$	generalized Kronecker product of $\{\mathbf{A}\}_N$ and $\mathbf{B}$
$\delta \mathbf{x}, \delta \mathbf{X}$	perturbations of the vector $\mathbf{x}$ and the matrix $\mathbf{X}$
$\text{cond}(\mathbf{A})$	condition number of the matrix $\mathbf{A}$
$\text{In}(\mathbf{A})$	inertia of a symmetric matrix $\mathbf{A}$
$i_+(\mathbf{A})$	number of positive eigenvalues of $\mathbf{A}$
$i_-(\mathbf{A})$	number of negative eigenvalues of $\mathbf{A}$
$i_0(\mathbf{A})$	number of zero eigenvalues of $\mathbf{A}$
$\mathbf{A} \sim \mathbf{B}$	similarity transformation
$\mathbf{A}(\lambda) \cong \mathbf{B}(\lambda)$	balanced transformation
$\mathbf{A} \doteq \mathbf{B}$	essentially equal matrices
$\mathbf{J} = \mathbf{PAP}^{-1}$	Jordan canonical form of the matrix $\mathbf{A}$
$d_k(x)$	$k$ th determinant divisor of a polynomial matrix $\mathbf{A}(x)$
$\sigma_k(x)$	$k$ th invariant factor of a polynomial matrix $\mathbf{A}(x)$
$\mathbf{A}(\lambda)$	$\lambda$ -matrix of the matrix $\mathbf{A}$
$\mathbf{S}(\lambda)$	Smith normal form of the $\lambda$ -matrix $\mathbf{A}(\lambda)$

### Special Vectors and Special Matrices

$\mathbf{P}_S$	projector onto the subspace $S$
$\mathbf{P}_S^\perp$	orthogonal projector onto the subspace $S$
$\mathbf{E}_{H S}$	oblique projector onto the subspace $H$ along the subspace $S$
$\mathbf{E}_{S H}$	oblique projector onto the subspace $S$ along the subspace $H$
$\mathbf{1}$	summing vector with all entries 1
$\mathbf{0}$	null or zero vector with all components 0
$\mathbf{e}_i$	basic vector with $e_i = 1$ and all other entries 0
$\boldsymbol{\pi}$	extracting vector with the last nonzero entry 1

**Special Vectors and Special Matrices** (*continued*)

- $\mathbf{i}_N$  index vector:  $[\langle 0 \rangle, \langle 1 \rangle, \dots, \langle N-1 \rangle]^T$
- $\mathbf{i}_{N,\text{rev}}$  bit-reversed index vector of  $\mathbf{i}_N$
- $\mathbf{O}$  null or zero matrix, with all components zero
- $\mathbf{I}$  identity matrix
- $\mathbf{K}_{mn}$   $mn \times mn$  commutation matrix
- $\mathbf{J}_n$   $n \times n$  exchange matrix:  $\mathbf{J}_n = [\mathbf{e}_n, \dots, \mathbf{e}_1]$
- $\mathbf{P}$   $n \times n$  permutation matrix:  $\mathbf{P} = [\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}], i_1, \dots, i_n \in \{1, \dots, n\}$
- $\mathbf{G}$  generalized permutation matrix or  $g$ -matrix:  $\mathbf{G} = \mathbf{P}\mathbf{D}$
- $\mathbf{C}_n$   $n \times n$  centering matrix =  $\mathbf{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n^T$
- $\mathbf{F}_N$   $N \times N$  Fourier matrix with entry  $F(i, k) = (e^{-j2\pi/N})^{(i-1)(k-1)}$
- $\mathbf{F}_{N,\text{rev}}$   $N \times N$  bit-reversed Fourier matrix
- $\mathbf{H}_n$  Hadamard matrix:  $\mathbf{H}_n\mathbf{H}_n^T = \mathbf{H}_n^T\mathbf{H}_n = n\mathbf{I}_n$
- $\mathbf{A}$  symmetric Toeplitz matrix:  $[a_{|i-j|}]_{i,j=1}^n$
- $\mathbf{A}$  complex Toeplitz matrix:  $\text{Toep}[a_0, a_1, \dots, a_n]$  with  $a_{-i} = a_i^*$
- $\mathbf{Q}_n$   $n \times n$  real orthogonal matrix:  $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$
- $\mathbf{U}_n$   $n \times n$  unitary matrix:  $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$
- $\mathbf{Q}_{m \times n}$   $m \times n$  semi-orthogonal matrix:  $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$  or  $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}_m$
- $\mathbf{U}_{m \times n}$   $m \times n$  para-unitary matrix:  $\mathbf{U}^H\mathbf{U} = \mathbf{I}_n, m > n$ , or  $\mathbf{U}\mathbf{U}^H = \mathbf{I}_m, m < n$
- $\mathbf{S}_b$  between-class scatter matrix
- $\mathbf{S}_w$  within-class scatter matrix

**Tensors**

- $\mathcal{A} \in \mathbb{K}^{I_1 \times \dots \times I_N}$   $N$ th-order real or complex tensor
- $\mathcal{I}, \mathcal{E}$  identity tensor
- $\mathbf{A}_{i::}$  horizontal slice matrix of  $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$
- $\mathbf{A}_{:j:}$  lateral slice matrix of  $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$
- $\mathbf{A}_{::k}$  frontal slice matrix of  $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$
- $\mathbf{a}_{:jk}, \mathbf{a}_{i:k}, \mathbf{a}_{ij:}$  mode-1, mode-2, mode-3 vectors of  $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$
- $\text{vec } \mathcal{A}$  vectorization of tensor  $\mathcal{A}$
- $\text{unvec } \mathcal{A}$  matricization of tensor  $\mathcal{A}$
- $\mathbf{A}^{(JK \times I)}, \mathbf{A}^{(KI \times J)}, \mathbf{A}^{(IJ \times K)}$  matricization of tensor  $\mathcal{A} \in \mathbb{K}^{I \times J \times K}$

**Tensors** (*continued*)

$\langle \mathcal{A}, \mathcal{B} \rangle$	inner product of tensors: $(\text{vec } \mathcal{A})^H \text{vec } \mathcal{B}$
$\ \mathcal{A}\ _F$	Frobenius norm of tensor $\mathcal{A}$
$\mathcal{A} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}$	outer product of three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$
$\mathcal{X} \times_n \mathbf{A}$	Tucker mode- $n$ product of $\mathcal{X}$ and $\mathbf{A}$
$\text{rank}(\mathcal{A})$	rank of tensor $\mathcal{A}$
$[\mathcal{G}; \mathbf{U}^{(1)}, \dots, \mathbf{U}^{(N)}]$	Tucker operator of tensor $\mathcal{G}$
$\mathcal{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}$	Tucker decomposition (third-order SVD)
$\mathcal{G} \times_1 \mathbf{U}^{(1)} \dots \times_N \mathbf{U}^{(N)}$	higher-order SVD of $N$ th-order tensor
$x_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr}$	CP decomposition of the third-order tensor $\mathcal{X}$
$\mathcal{A}\mathbf{x}^m, \mathcal{A}\mathbf{x}^{m-1}$	tensor–vector product of $\mathcal{A} \in \mathbb{S}^{[m,n]}$ , $\mathbf{x} \in \mathbb{C}^{n \times 1}$
$\det(\mathcal{A})$	determinant of tensor $\mathcal{A}$
$\lambda_i(\mathcal{A})$	$i$ th eigenvalue of tensor $\mathcal{A}$
$\sigma(\mathcal{A})$	spectrum of tensor $\mathcal{A}$

**Functions and Derivatives**

$\stackrel{\text{def}}{=}$	defined to be equal
$\sim$	asymptotically equal (in scaling sense)
$\approx$	approximately equal (in numerical value)
$f: \mathbb{R}^m \rightarrow \mathbb{R}$	real function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^m, f \in \mathbb{R}$
$f: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$	real function $f(\mathbf{X}), \mathbf{X} \in \mathbb{R}^{m \times n}, f \in \mathbb{R}$
$f: \mathbb{C}^m \times \mathbb{C}^m \rightarrow \mathbb{R}$	real function $f(\mathbf{z}, \mathbf{z}^*), \mathbf{z} \in \mathbb{C}^m, f \in \mathbb{R}$
$f: \mathbb{C}^{m \times n} \times \mathbb{C}^{m \times n} \rightarrow \mathbb{R}$	real function $f(\mathbf{Z}, \mathbf{Z}^*), \mathbf{Z} \in \mathbb{C}^{m \times n}, f \in \mathbb{R}$
$\text{dom } f, \mathcal{D}$	definition domain of function $f$
$\mathcal{E}$	domain of equality constraint function
$\mathcal{I}$	domain of inequality constraint function
$\mathcal{F}$	feasible set
$B_c(\mathbf{c}; r), B_o(\mathbf{c}; r)$	closed, open neighborhoods of $\mathbf{c}$ with radius $r$
$B_c(\mathbf{C}; r), B_o(\mathbf{C}; r)$	closed, open neighborhoods of $\mathbf{C}$ with radius $r$
$f(\mathbf{z}, \mathbf{z}^*), f(\mathbf{Z}, \mathbf{Z}^*)$	function of complex variables $\mathbf{z}$ , or $\mathbf{Z}$
$df(\mathbf{z}, \mathbf{z}^*), df(\mathbf{Z}, \mathbf{Z}^*)$	complex differentials
$D(\mathbf{p} \parallel \mathbf{g})$	distance between vectors $\mathbf{p}$ and $\mathbf{g}$



**Functions and Derivatives** (*continued*)

$D(\mathbf{x}, \mathbf{y})$	dissimilarity between vectors $\mathbf{x}$ and $\mathbf{y}$
$D_E(\mathbf{x}, \mathbf{y})$	Euclidean distance between vectors $\mathbf{x}$ and $\mathbf{y}$
$D_M(\mathbf{x}, \mathbf{y})$	Mahalanobis distance between vectors $\mathbf{x}$ and $\mathbf{y}$
$D_J^g(\mathbf{x}, \mathbf{y})$	Bregman distance between vectors $\mathbf{x}$ and $\mathbf{y}$
$D_{\mathbf{x}}, D_{\text{vec } \mathbf{X}}$	row partial derivative operators
$D_{\mathbf{x}}f(\mathbf{x})$	row partial derivative vectors of $f(\mathbf{x})$
$D_{\text{vec } \mathbf{X}}f(\mathbf{X})$	row partial derivative vectors of $f(\mathbf{X})$
$D_{\mathbf{X}}$	Jacobian operator
$D_{\mathbf{X}}f(\mathbf{X})$	Jacobian matrix of the function $f(\mathbf{X})$
$D_{\mathbf{z}}, D_{\text{vec } \mathbf{Z}}$	complex cogradient operator
$D_{\mathbf{z}^*}, D_{\text{vec } \mathbf{Z}^*}$	complex conjugate cogradient operator
$D_{\mathbf{z}}f(\mathbf{z}, \mathbf{z}^*)$	cogradient vector of complex function $f(\mathbf{z}, \mathbf{z}^*)$
$D_{\text{vec } \mathbf{Z}}f(\mathbf{Z}, \mathbf{Z}^*)$	cogradient vector of complex function $f(\mathbf{Z}, \mathbf{Z}^*)$
$D_{\mathbf{z}^*}f(\mathbf{z}, \mathbf{z}^*)$	conjugate cogradient vector of $f(\mathbf{z}, \mathbf{z}^*)$
$D_{\text{vec } \mathbf{Z}^*}f(\mathbf{Z}, \mathbf{Z}^*)$	conjugate cogradient vector of $f(\mathbf{Z}, \mathbf{Z}^*)$
$D_{\mathbf{Z}}, \nabla_{\mathbf{Z}^*}$	Jacobian, gradient matrix operator
$D_{\mathbf{Z}}f(\mathbf{Z}, \mathbf{Z}^*)$	Jacobian matrices of $f(\mathbf{Z}, \mathbf{Z}^*)$
$\nabla_{\mathbf{Z}}f(\mathbf{Z}, \mathbf{Z}^*)$	gradient matrices of $f(\mathbf{Z}, \mathbf{Z}^*)$
$D_{\mathbf{Z}^*}f(\mathbf{Z}, \mathbf{Z}^*)$	conjugate Jacobian matrices of $f(\mathbf{Z}, \mathbf{Z}^*)$
$\nabla_{\mathbf{Z}^*}f(\mathbf{Z}, \mathbf{Z}^*)$	conjugate gradient matrices of $f(\mathbf{Z}, \mathbf{Z}^*)$
$\nabla_{\mathbf{x}}, \nabla_{\text{vec } \mathbf{X}}$	gradient vector operator
$\nabla_{\mathbf{x}}f(\mathbf{x})$	gradient vector of function $f(\mathbf{x})$
$\nabla_{\text{vec } \mathbf{X}}f(\mathbf{X})$	gradient vector of function $f(\mathbf{X})$
$\nabla f(\mathbf{X})$	gradient matrix of function $f$
$\nabla^2 f$	Hessian matrix of function $f$
$\mathbf{H}_{\mathbf{x}}f(\mathbf{x})$	Hessian matrix of function $f$
$\mathbf{H}f(\mathbf{z}, \mathbf{z}^*)$	full Hessian matrix of $f(\mathbf{z}, \mathbf{z}^*)$
$\mathbf{H}_{\mathbf{z}, \mathbf{z}}, \mathbf{H}_{\mathbf{z}, \mathbf{z}^*}, \mathbf{H}_{\mathbf{z}^*, \mathbf{z}^*}$	part Hessian matrices of function $f(\mathbf{z}, \mathbf{z}^*)$
$df, \partial f$	differential or subdifferential of function $f$
$\mathbf{g} \in \partial f$	subgradient of function $f$
$\Delta \mathbf{x}$	descent direction of function $f(\mathbf{x})$
$\Delta \mathbf{x}_{\text{nt}}$	Newton step of function $f(\mathbf{x})$
$\max f, \min f$	maximize, minimize function $f$

### Functions and Derivatives (*continued*)

$\max\{x, y\}$	maximum of $x$ and $y$
$\min\{x, y\}$	minimum of $x$ and $y$
$\inf$	infimum
$\sup$	supremum
Re, Im	real part, imaginary part of complex number
arg	argument of objective function or complex number
$\mathcal{P}_C(\mathbf{y}), \mathbf{P}_{CY}$	projection operator of the vector $\mathbf{y}$ onto the subspace $C$
$\mathbf{x}^+$	nonnegative vector with entry $[\mathbf{x}^+]_i = \max\{x_i, 0\}$
$\text{prox}_h(\mathbf{u})$	proximal operator of function $h(\mathbf{x})$ to point $\mathbf{u}$
$\text{prox}_h(\mathbf{U})$	proximal operator of function $h(\mathbf{X})$ to point $\mathbf{U}$
$\text{soft}(x, \tau), \mathcal{S}_\tau[x]$	soft thresholding operator of real variable $x$

### Probability

$\text{soft}(\mathbf{x}, \tau), \text{soft}(\mathbf{X}, \tau)$	soft thresholding operator of real variables $\mathbf{x}, \mathbf{X}$
$\mathcal{D}_\mu(\Sigma)$	singular value (matrix) thresholding (operation)
$I_C(\mathbf{x})$	indicator function
$\mathbf{x}_{\text{LS}}, \mathbf{X}_{\text{LS}}$	least squares solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{DLS}}, \mathbf{X}_{\text{DLS}}$	data least squares solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{WLS}}, \mathbf{X}_{\text{WLS}}$	weighted least squares solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{opt}}, \mathbf{X}_{\text{opt}}$	optimal solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{Tik}}, \mathbf{X}_{\text{Tik}}$	Tikhonov solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{TLS}}, \mathbf{X}_{\text{TLS}}$	total least squares (TLS) solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{GTLS}}, \mathbf{X}_{\text{GTLS}}$	generalized TLS solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$\mathbf{x}_{\text{ML}}, \mathbf{X}_{\text{ML}}$	maximum likelihood solutions to $\mathbf{Ax} = \mathbf{b}, \mathbf{AX} = \mathbf{B}$
$D_{AB}^{(\alpha, \beta)}(\mathbf{P} \parallel \mathbf{G})$	alpha–beta (AB) divergence of matrices $\mathbf{P}$ and $\mathbf{G}$
$D_\alpha(\mathbf{P} \parallel \mathbf{G})$	alpha-divergence of matrices $\mathbf{P}$ and $\mathbf{G}$
$D_\beta(\mathbf{P} \parallel \mathbf{G})$	beta-divergence of matrices $\mathbf{P}$ and $\mathbf{G}$
$D_{\text{KL}}(\mathbf{P} \parallel \mathbf{G})$	Kullback–Leibler divergence of matrices $\mathbf{P}$ and $\mathbf{G}$
$\ln_q(x)$	Tsallis logarithm
$\exp_q(x)$	$q$ -exponential
$\ln_{1-\alpha}(x)$	deformed logarithm
$\exp_{1-\alpha}(x)$	deformed exponential

xxx

Notation

**Probability** (*continued*)

$\text{sign}(x)$	signum function of real valued variable $x$
$\text{SGN}(x)$	signum multifunction of real valued variable $x$
$\text{shrink}(y, \alpha)$	shrink operator
$R(\mathbf{x})$	Rayleigh quotient, generalized Rayleigh quotient
$\mathbf{M}_{\text{off}}$	off-diagonal matrix corresponding to matrix $\mathbf{M}$
$z^{-j} \mathbf{x}(n)$	time-shifting operation on vector $\mathbf{x}(n)$
$E\{\mathbf{x}\} = \bar{\mathbf{x}}$	expectation (mean) of random vector $\mathbf{x}$
$E\{\mathbf{x}\mathbf{x}^H\}$	autocorrelation matrix of random vector $\mathbf{x}$
$E\{\mathbf{x}\mathbf{y}^H\}$	cross-correlation matrix of random vectors $\mathbf{x}$ and $\mathbf{y}$
$\rho_{xy}$	correlation coefficient of random vectors $\mathbf{x}$ and $\mathbf{y}$
$N(\mathbf{c}, \mathbf{\Sigma})$	Gaussian random vector with mean (vector) $\mathbf{c}$ and covariance (matrix) $\mathbf{\Sigma}$
$CN(\mathbf{c}, \mathbf{\Sigma})$	complex Gaussian random vector with mean (vector) $\mathbf{c}$ and covariance (matrix) $\mathbf{\Sigma}$
$f(x_1, \dots, x_m)$	joint probability density function of random vector $\mathbf{x} = [x_1, \dots, x_m]^T$
$\Phi_{\mathbf{x}}(\boldsymbol{\omega})$	characteristic function of random vector $\mathbf{x}$

## Abbreviations

AB	alpha–beta
ADMM	alternating direction method of multipliers
ALS	alternating least squares
ANLS	alternating nonnegative least squares
APGL	accelerated proximal gradient line
ARNLS	alternating regularization nonnegative least squares
BBGP	Barzilai–Borwein gradient projection
BCQP	bound-constrained quadratic program
BFGS	Broyden–Fletcher–Goldfarb–Shanno
BP	basis pursuit
BPDN	basis pursuit denoising
BSS	blind source separation
CANDECOMP	canonical factor decomposition
CNMF	constrained nonnegative matrix factorization
CoSaMP	compression sampling matching pursuit
CP	CANDECOMP/PARAFAC
DCT	discrete cosine transform
DFT	discrete Fourier transform
DLS	data least squares
DOA	direction of arrival
EEG	electroencephalography
EM	expectation-maximization
EMML	expectation-maximization maximum likelihood
ESPRIT	estimating signal parameters via rotational invariance technique

EVD	eigenvalue decomposition
FAJD	fast approximate joint diagonalization
FAL	first-order augmented Lagrangian
FFT	fast Fourier transform
FISTA	fast iterative soft thresholding algorithm
GEAP	generalized eigenproblem adaptive power
GEVD	generalized eigenvalue decomposition
GSVD	generalized singular value decomposition
GTLS	generalized total least squares
HBM	heavy ball method
HOOI	higher-order orthogonal iteration
HOSVD	higher-order singular value decomposition
ICA	independent component analysis
IDFT	inverse discrete Fourier transform
iid	independent and identically distributed
inf	infimum
KKT	Karush–Kuhn–Tucker
KL	Kullback–Leibler
LARS	least angle regressive
Lasso	least absolute shrinkage and selection operator
LDA	linear discriminant analysis
LMV	Lathauwer–Moor–Vanderwalle
LP	linear programming
LS	least squares
LSI	latent semantic indexing
max	maximize, maximum
MCA	minor component analysis
MIMO	multiple-input–multiple-output
min	minimize, minimum
ML	maximum likelihood
MP	matching pursuit
MPCA	multilinear principal component analysis
MUSIC	multiple signal classification

*Abbreviations*

xxxiii

NeNMF	Nesterov nonnegative matrix factorization
NMF	nonnegative matrix factorization
NTD	nonnegative tensor decomposition
OGM	optimal gradient method
OMP	orthogonal matching pursuit
PARAFAC	parallel factor decomposition
PAST	projection approximation subspace tracking
PASTd	projection approximation subspace tracking via deflation
PCA	principal component analysis
PCG	preconditioned conjugate gradient
PCP	principal component pursuit
pdf	positive definite
PMF	positive matrix factorization
psdf	positive semi-definite
PSF	point-spread function
PSVD	product singular value decomposition
RIC	restricted isometry constant
RIP	restricted isometry property
ROMP	regularization orthogonal matching pursuit
RR	Rayleigh–Ritz
QCLP	quadratically constrained linear programming
QEP	quadratic eigenvalue problem
QP	quadratic programming
QV	quadratic variation
sign	signum
SPC	smooth PARAFAC tensor completion
StOMP	stagewise orthogonal matching pursuit
sup	supremum
SVD	singular value decomposition
SVT	singular value thresholding
TLS	total least squares
TV	total variation
UPCA	unfold principal component analysis
VQ	vector quantization