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
To John Zhang, Ellen Zhang and Andrew Wei
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>xvii</td>
</tr>
<tr>
<td>Notation</td>
<td>xxi</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>xxxi</td>
</tr>
<tr>
<td>Algorithms</td>
<td>xxxiv</td>
</tr>
</tbody>
</table>

## PART I  MATRIX ALGEBRA  

1. **Introduction to Matrix Algebra**  
   1.1 Basic Concepts of Vectors and Matrices  
      1.1.1 Vectors and Matrices  
      1.1.2 Basic Vector Calculus  
      1.1.3 Basic Matrix Calculus  
      1.1.4 Linear Independence of Vectors  
      1.1.5 Matrix Functions  
   1.2 Elementary Row Operations and Applications  
      1.2.1 Elementary Row Operations  
      1.2.2 Gauss Elimination Methods  
   1.3 Sets, Vector Subspaces and Linear Mapping  
      1.3.1 Sets  
      1.3.2 Fields and Vector Spaces  
      1.3.3 Linear Mapping  
   1.4 Inner Products and Vector Norms  
      1.4.1 Inner Products of Vectors  
      1.4.2 Norms of Vectors  
      1.4.3 Similarity Comparison Between Vectors  
      1.4.4 Banach Space, Euclidean Space, Hilbert Space  
      1.4.5 Inner Products and Norms of Matrices  
   1.5 Random Vectors  
      1.5.1 Statistical Interpretation of Random Vectors  
      1.5.2 Gaussian Random Vectors
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6 Performance Indexes of Matrices</td>
<td>47</td>
</tr>
<tr>
<td>1.6.1 Quadratic Forms</td>
<td>47</td>
</tr>
<tr>
<td>1.6.2 Determinants</td>
<td>49</td>
</tr>
<tr>
<td>1.6.3 Matrix Eigenvalues</td>
<td>52</td>
</tr>
<tr>
<td>1.6.4 Matrix Trace</td>
<td>54</td>
</tr>
<tr>
<td>1.6.5 Matrix Rank</td>
<td>56</td>
</tr>
<tr>
<td>1.7 Inverse Matrices and Pseudo-Inverse Matrices</td>
<td>59</td>
</tr>
<tr>
<td>1.7.1 Definition and Properties of Inverse Matrices</td>
<td>59</td>
</tr>
<tr>
<td>1.7.2 Matrix Inversion Lemma</td>
<td>60</td>
</tr>
<tr>
<td>1.7.3 Inversion of Hermitian Matrices</td>
<td>61</td>
</tr>
<tr>
<td>1.7.4 Left and Right Pseudo-Inverse Matrices</td>
<td>63</td>
</tr>
<tr>
<td>1.8 Moore–Penrose Inverse Matrices</td>
<td>65</td>
</tr>
<tr>
<td>1.8.1 Definition and Properties</td>
<td>65</td>
</tr>
<tr>
<td>1.8.2 Computation of Moore–Penrose Inverse Matrix</td>
<td>69</td>
</tr>
<tr>
<td>1.9 Direct Sum and Hadamard Product</td>
<td>71</td>
</tr>
<tr>
<td>1.9.1 Direct Sum of Matrices</td>
<td>71</td>
</tr>
<tr>
<td>1.9.2 Hadamard Product</td>
<td>72</td>
</tr>
<tr>
<td>1.10 Kronecker Products and Khatri–Rao Product</td>
<td>75</td>
</tr>
<tr>
<td>1.10.1 Kronecker Products</td>
<td>75</td>
</tr>
<tr>
<td>1.10.2 Generalized Kronecker Products</td>
<td>77</td>
</tr>
<tr>
<td>1.10.3 Khatri–Rao Product</td>
<td>78</td>
</tr>
<tr>
<td>1.11 Vectorization and Matricization</td>
<td>79</td>
</tr>
<tr>
<td>1.11.1 Vectorization and Commutation Matrix</td>
<td>79</td>
</tr>
<tr>
<td>1.11.2 Matricization of a Vector</td>
<td>82</td>
</tr>
<tr>
<td>1.11.3 Properties of Vectorization Operator</td>
<td>83</td>
</tr>
<tr>
<td>1.12 Sparse Representations</td>
<td>84</td>
</tr>
<tr>
<td>1.12.1 Sparse Vectors and Sparse Representations</td>
<td>84</td>
</tr>
<tr>
<td>1.12.2 Sparse Representation of Face Recognition</td>
<td>86</td>
</tr>
<tr>
<td>Exercises</td>
<td>87</td>
</tr>
<tr>
<td>2 Special Matrices</td>
<td>95</td>
</tr>
<tr>
<td>2.1 Hermitian Matrices</td>
<td>95</td>
</tr>
<tr>
<td>2.2 Idempotent Matrix</td>
<td>96</td>
</tr>
<tr>
<td>2.3 Permutation Matrix</td>
<td>98</td>
</tr>
<tr>
<td>2.3.1 Permutation Matrix and Exchange Matrix</td>
<td>98</td>
</tr>
<tr>
<td>2.3.2 Generalized Permutation Matrix</td>
<td>101</td>
</tr>
<tr>
<td>2.4 Orthogonal Matrix and Unitary Matrix</td>
<td>104</td>
</tr>
<tr>
<td>2.5 Band Matrix and Triangular Matrix</td>
<td>107</td>
</tr>
<tr>
<td>2.5.1 Band Matrix</td>
<td>107</td>
</tr>
<tr>
<td>2.5.2 Triangular Matrix</td>
<td>107</td>
</tr>
<tr>
<td>2.6 Summing Vector and Centering Matrix</td>
<td>109</td>
</tr>
<tr>
<td>2.6.1 Summing Vector</td>
<td>110</td>
</tr>
</tbody>
</table>
2.6.2 Centering Matrix

2.7 Vandermonde Matrix and Fourier Matrix
2.7.1 Vandermonde Matrix
2.7.2 Fourier Matrix
2.7.3 Index Vectors
2.7.4 FFT Algorithm

2.8 Hadamard Matrix

2.9 Toeplitz Matrix
2.9.1 Symmetric Toeplitz Matrix
2.9.2 Discrete Cosine Transform of Toeplitz Matrix

Exercises

3 Matrix Differential
3.1 Jacobian Matrix and Gradient Matrix
3.1.1 Jacobian Matrix
3.1.2 Gradient Matrix
3.1.3 Calculation of Partial Derivative and Gradient

3.2 Real Matrix Differential
3.2.1 Calculation of Real Matrix Differential
3.2.2 Jacobian Matrix Identification
3.2.3 Jacobian Matrix of Real Matrix Functions

3.3 Real Hessian Matrix and Identification
3.3.1 Real Hessian Matrix
3.3.2 Real Hessian Matrix Identification

3.4 Complex Gradient Matrices
3.4.1 Holomorphic Function and Complex Partial Derivative
3.4.2 Complex Matrix Differential
3.4.3 Complex Gradient Matrix Identification

3.5 Complex Hessian Matrices and Identification
3.5.1 Complex Hessian Matrices
3.5.2 Complex Hessian Matrix Identification

Exercises

PART II MATRIX ANALYSIS

4 Gradient Analysis and Optimization
4.1 Real Gradient Analysis
4.1.1 Stationary Points and Extreme Points
4.1.2 Real Gradient Analysis of \( f(x) \)
4.1.3 Real Gradient Analysis of \( f(X) \)

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

6 Solving Matrix Equations 315
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

7 Eigenanalysis 413
   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
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

8 Subspace Analysis and Tracking 511
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
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

9 Projection Analysis 551
  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

PART III HIGHER-ORDER MATRIX ANALYSIS 587

10 Tensor Analysis 589
  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

References 681

Index 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.

[Diagram]

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
Preface

technology (including signal and image processing, communications, automation
control, system engineering and pattern recognition), artificial intelligence, bioin-
formatics, biomedical engineering, to name just a selection. This book is dedicated
to providing individuals in those disciplines with a solid foundation of the funda-
mental 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 cal-
culus, 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, feed-
back 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, inform-
ation 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 Man-
ger, Esther Miguélez, 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

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

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

Sets (continued)

\( S_n \times n \) symmetric \( n \times n \) matrices
\( S_{n+} \times n \) symmetric positive semi-definite \( n \times n \) matrices
\( S_{n+}^n \) symmetric positive definite \( n \times n \) matrices
\( S_{m,n} \) symmetric \( m \)-th order \( n \)-dimensional tensors \( A^{I_1 \times \cdots \times I_m} I_1 = \cdots = I_n \)
\( S_{m,n}^+ \) 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 \( V \)
\( U \to W \) \( U \) transforms to \( W \)
\( \exists \) 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 \( X_1, \ldots, X_n \)
\( \mathcal{L} \) 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 \( S \)
\( \mathcal{K}^m(A, f) \) order-\( m \) Krylov subspace generated by \( A \) and \( f \)
\( \text{Col}(A) \) column space of the matrix \( A \)
\( \text{Ker}(A) \) kernel space of the matrix \( A \)
\( \text{Null}(A) \) null space of the matrix \( A \)
\( \text{nullity}(A) \) nullity of the matrix \( A \)
\( \text{Range}(A) \) range space of the matrix \( A \)
**Sets (continued)**

Row(\(A\)) \(\) row space of the matrix \(A\)

Span(\(a_1, \ldots, a_m\)) \(\) span of vectors \(a_1, \ldots, a_m\)

**Vectors**

\(x^*\) \(\) conjugate of the vector \(x\)

\(x^T\) \(\) transpose of the vector \(x\)

\(x^H\) \(\) conjugate transpose (Hermitian conjugate) of the vector \(x\)

\(\mathcal{L}(u)\) \(\) linear transform of the vector \(u\)

\(\|x\|_0\) \(\) \(\ell_0\)-norm: the number of nonzero entries in the vector \(x\)

\(\|x\|_1\) \(\) \(\ell_1\)-norm of the vector \(x\)

\(\|x\|_2\) \(\) Euclidean norm of the vector \(x\)

\(\|x\|_p\) \(\) \(\ell_p\)-norm or Hölder norm of the vector \(x\)

\(\|x\|_*\) \(\) nuclear norm of the vector \(x\)

\(\|x\|_\infty\) \(\) \(\ell_\infty\)-norm of the vector \(x\)

\((x, y) = x^H y\) \(\) inner product of vectors \(x\) and \(y\)

\(x \circ y = xy^H\) \(\) outer product of vectors \(x\) and \(y\)

\(x \perp y\) \(\) orthogonality of vectors \(x\) and \(y\)

\(x > 0\) \(\) positive vector, with components \(x_i > 0, \forall i\)

\(x \geq 0\) \(\) nonnegative vector, with components \(x_i \geq 0, \forall i\)

\(x \geq y\) \(\) vector elementwise inequality \(x_i \geq y_i, \forall i\)

\(\text{unvec}(x)\) \(\) matricization of the column vector \(x\)

\(\text{unrvec}(x)\) \(\) row matricization of the column vector \(x\)

\(\theta_i^{(m)}, y_i^{(m)}\) \(\) Rayleigh–Ritz (RR) values, RR vectors

\((\theta_i^{(m)}, y_i^{(m)})\) \(\) Ritz pair

**Matrices**

\(A \in \mathbb{R}^{m \times n}\) \(\) real \(m \times n\) matrix \(A\)

\(A \in \mathbb{C}^{m \times n}\) \(\) complex \(m \times n\) matrix \(A\)

\(A[x] \in \mathbb{R}[x]^{m \times n}\) \(\) real \(m \times n\) polynomial matrix \(A\)

\(A[x] \in \mathbb{C}[x]^{m \times n}\) \(\) complex \(m \times n\) polynomial matrix \(A\)

\(A^*\) \(\) conjugate of \(A\)

\(A^T\) \(\) transpose of \(A\)
Matrices (continued)

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

\begin{itemize}
    \item \textbf{diag}(\mathbf{A}) \quad \text{diagonal vector of } \mathbf{A} = [a_{ij}]: [a_{11}, \ldots, a_{nn}]^T
    \item \textbf{Diag}(\mathbf{A}) \quad \text{diagonal matrix of } \mathbf{A} = [a_{ij}]: \text{Diag}(a_{11}, \ldots, a_{nn})
    \item \langle \mathbf{A}, \mathbf{B} \rangle \quad \text{inner product of } \mathbf{A} \text{ and } \mathbf{B}: (\text{vec } \mathbf{A})^H \text{vec } \mathbf{B}
    \item \mathbf{A} \otimes \mathbf{B} \quad \text{Kronecker product of matrices } \mathbf{A} \text{ and } \mathbf{B}
    \item \mathbf{A} \odot \mathbf{B} \quad \text{Khatri–Rao product of matrices } \mathbf{A} \text{ and } \mathbf{B}
    \item \mathbf{A} \oplus \mathbf{B} \quad \text{Hadamard product of matrices } \mathbf{A} \text{ and } \mathbf{B}
    \item \{\mathbf{A}\}_N \quad \text{direct sum of matrices } \mathbf{A} \text{ and } \mathbf{B}
    \item \{\mathbf{A}\}_N \otimes \mathbf{B} \quad \text{generalized Kronecker product of } \{\mathbf{A}\}_N \text{ and } \mathbf{B}
    \item \delta \mathbf{x}, \delta \mathbf{X} \quad \text{perturbations of the vector } \mathbf{x} \text{ and the matrix } \mathbf{X}
    \item \text{cond}(\mathbf{A}) \quad \text{condition number of the matrix } \mathbf{A}
    \item \text{In}(\mathbf{A}) \quad \text{inertia of a symmetric matrix } \mathbf{A}
    \item \text{i}_+ (\mathbf{A}) \quad \text{number of positive eigenvalues of } \mathbf{A}
    \item \text{i}_- (\mathbf{A}) \quad \text{number of negative eigenvalues of } \mathbf{A}
    \item \text{i}_0 (\mathbf{A}) \quad \text{number of zero eigenvalues of } \mathbf{A}
    \item \mathbf{A} \sim \mathbf{B} \quad \text{similarity transformation}
    \item \mathbf{A}(\lambda) \equiv \mathbf{B}(\lambda) \quad \text{balanced transformation}
    \item \mathbf{A} \doteq \mathbf{B} \quad \text{essentially equal matrices}
    \item \mathbf{J} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1} \quad \text{Jordan canonical form of the matrix } \mathbf{A}
    \item \text{d}_k (\mathbf{x}) \quad \text{kth determinant divisor of a polynomial matrix } \mathbf{A}(\mathbf{x})
    \item \text{\sigma}_k (\mathbf{x}) \quad \text{kth invariant factor of a polynomial matrix } \mathbf{A}(\mathbf{x})
    \item \mathbf{A}(\lambda) \quad \lambda\text{-matrix of the matrix } \mathbf{A}
    \item \mathbf{S}(\lambda) \quad \text{Smith normal form of the } \lambda\text{-matrix } \mathbf{A}(\lambda)
\end{itemize}

\textbf{Special Vectors and Special Matrices}

\begin{itemize}
    \item \mathbf{P}_S \quad \text{projector onto the subspace } S
    \item \mathbf{P}_S^\perp \quad \text{orthogonal projector onto the subspace } S
    \item \mathbf{E}_{H|S} \quad \text{oblique projector onto the subspace } H \text{ along the subspace } S
    \item \mathbf{E}_{S|H} \quad \text{oblique projector onto the subspace } S \text{ along the subspace } H
    \item \mathbf{1} \quad \text{summing vector with all entries 1}
    \item \mathbf{0} \quad \text{null or zero vector with all components 0}
    \item \mathbf{e}_i \quad \text{basic vector with } e_i = 1 \text{ and all other entries 0}
    \item \pi \quad \text{extracting vector with the last nonzero entry 1}
\end{itemize}
Special Vectors and Special Matrices (continued)

\( i_N \) index vector: \([\langle 0 \rangle, \langle 1 \rangle, \ldots, \langle N-1 \rangle]^T \)

\( i_{N,\text{rev}} \) bit-reversed index vector of \( i_N \)

\( \mathbf{O} \) 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: \( \mathbf{J}_n = [\mathbf{e}_n, \ldots, \mathbf{e}_1] \)

\( \mathbf{P} \) \( n \times n \) permutation matrix: \( \mathbf{P} = [\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}], i_1, \ldots, i_n \in \{1, \ldots, n\} \)

\( \mathbf{G} \) generalized permutation matrix or \( g \)-matrix: \( \mathbf{G} = \mathbf{PD} \)

\( \mathbf{C}_n \) \( n \times n \) 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: \( [a_{i-j}]_{i,j=1}^n \)

\( \mathbf{Q}_n \) \( n \times n \) real orthogonal matrix: \( \mathbf{QQ}^T = \mathbf{Q}^T \mathbf{Q} = \mathbf{I} \)

\( \mathbf{U}_n \) \( n \times n \) unitary matrix: \( \mathbf{UU}^H = \mathbf{U}^H \mathbf{U} = \mathbf{I} \)

\( \mathbf{Q}_{m \times n} \) \( m \times n \) semi-orthogonal matrix: \( \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_n \) or \( \mathbf{QQ}^T = \mathbf{I}_m \)

\( \mathbf{U}_{m \times n} \) \( m \times n \) para-unitary matrix: \( \mathbf{U}^H \mathbf{U} = \mathbf{I}_n, m > n \), or \( \mathbf{UU}^H = \mathbf{I}_m, m < n \)

\( \mathbf{S}_b \) between-class scatter matrix

\( \mathbf{S}_w \) within-class scatter matrix

Tensors

\( \mathbf{A} \in \mathbb{K}^{I_1 \times \cdots \times I_N} \) \( N \)th-order real or complex tensor

\( \mathbf{I}, \mathbf{E} \) identity tensor

\( \mathbf{A}_{i:} \) horizontal slice matrix of \( \mathbf{A} \in \mathbb{K}^{I \times J \times K} \)

\( \mathbf{A}_{j:} \) lateral slice matrix of \( \mathbf{A} \in \mathbb{K}^{I \times J \times K} \)

\( \mathbf{A}_{k:} \) frontal slice matrix of \( \mathbf{A} \in \mathbb{K}^{I \times J \times K} \)

\( a_{jk}, a_{ik}, a_{ij} \) mode-1, mode-2, mode-3 vectors of \( \mathbf{A} \in \mathbb{K}^{I \times J \times K} \)

\( \text{vec} \mathbf{A} \) vectorization of tensor \( \mathbf{A} \)

\( \text{unvec} \mathbf{A} \) matricization of tensor \( \mathbf{A} \)

\( \mathbf{A}^{(JK \times I)}, \mathbf{A}^{(KI \times J)}, \mathbf{A}^{(IJ \times K)} \) matricization of tensor \( \mathbf{A} \in \mathbb{K}^{I \times J \times K} \)
Tensors (continued)

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

Functions and Derivatives

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

- **D**(x, y) dissimilarity between vectors x and y
- **D_E**(x, y) Euclidean distance between vectors x and y
- **D_M**(x, y) Mahalanobis distance between vectors x and y
- **D_g**(x, y) Bregman distance between vectors x and y
- **D_x, D_{vec X}** row partial derivative operators
- **D_x f(x)** row partial derivative vectors of **f**(x)
- **D_{vec X} f(X)** row partial derivative vectors of **f**(X)
- **D_X** Jacobian operator
- **D_X f(X)** Jacobian matrix of the function **f**(X)
- **D_{x, D_{vec Z}}** complex conjugate cogradient operator
- **D_z^* f(z, z^*)** conjugate cogradient vector of complex function **f**(z, z^*)
- **D_{vec Z} f(Z, Z^*)** conjugate cogradient vector of complex function **f**(Z, Z^*)
- **D_Z, D_{vec Z}^*** Jacobian, gradient matrix operator
- **D_Z f(Z, Z^*)** Jacobian matrices of **f**(Z, Z^*)
- **∇_Z f(Z, Z^*)** gradient matrices of **f**(Z, Z^*)
- **D_{Z, Z^*} f(Z, Z^*)** conjugate Jacobian matrices of **f**(Z, Z^*)
- **∇_{Z, Z^*} f(Z, Z^*)** conjugate gradient matrices of **f**(Z, Z^*)
- **∇_x, ∇_{vec X}** gradient vector operator
- **∇_x f(x)** gradient vector of function **f**(x)
- **∇_{vec X} f(X)** gradient vector of function **f**(X)
- **∇ f(X)** gradient matrix of function **f**
- **∇^2 f** Hessian matrix of function **f**
- **H_x f(x)** Hessian matrix of function **f**
- **H_f(z, z^*)** full Hessian matrix of **f**(z, z^*)
- **H_{x, x^*}, H_{x, z^*}, H_{z, x^*}** part Hessian matrices of function **f**(z, z^*)
- **df, ∂f** differential or subdifferential of function **f**
- **g ∈ ∂f** subgradient of function **f**
- **Δx** descent direction of function **f**(x)
- **Δ_{x_{nt}}** Newton step of function **f**(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
\(\mathcal{P}_C(y), P_C y\) projection operator of the vector y onto the subspace C
\(x^+\) nonnegative vector with entry \([x^+]_i = \max\{x_i, 0\}\)
prox\(_\mu(u)\) proximal operator of function \(h(x)\) to point u
prox\(_\mu(U)\) proximal operator of function \(h(X)\) to point U
soft\((x, \tau), S_\tau[x]\) soft thresholding operator of real variable x

Probability

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

Signum multifunction of real valued variable $x$

Shrink operator

Rayleigh quotient, generalized Rayleigh quotient

Off-diagonal matrix corresponding to matrix $M$

time-shifting operation on vector $x(n)$

Expectation (mean) of random vector $x$

Autocorrelation matrix of random vector $x$

Cross-correlation matrix of random vectors $x$ and $y$

Correlation coefficient of random vectors $x$ and $y$

Gaussian random vector with mean (vector) $c$ and covariance (matrix) $\Sigma$

Complex Gaussian random vector with mean (vector) $c$ and covariance (matrix) $\Sigma$

Joint probability density function of random vector

$x = [x_1, \ldots, x_m]^T$

Characteristic function of random vector $x$
## Abbreviations

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

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

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