NONNEGATIVE MATRIX AND TENSOR FACTORIZATION
APPLICATIONS TO EXPLORATORY MULTI-WAY DATA ANALYSIS AND BLIND SOURCE SEPARATION

Andrzej Cichocki
Laboratory for Advanced Brain Signal Processing, Riken Brain Science Institute, Japan; and Warsaw University of Technology and Systems Research Institute, PAN, Poland

Rafal Zdunek
Institute of Telecommunications, Teleinformatics and Acoustics, Wroclaw University of Technology, Poland; and Riken Brain Science Institute, Japan

Anh Huy Phan
Laboratory for Advanced Brain Signal Processing, Riken Brain Science Institute, Japan

Shun-ichi Amari
Research Unit for Mathematical Neuroscience, Riken Brain Science Institute, Japan

©WILEY
A John Wiley and Sons, Ltd, Publication
# Contents

Preface xi  
Acknowledgments xv  
Glossary of Symbols and Abbreviations xvii

1 Introduction – Problem Statements and Models 1

1.1 Blind Source Separation and Linear Generalized Component Analysis 2
1.2 Matrix Factorization Models with Nonnegativity and Sparsity Constraints 7
  1.2.1 Why Nonnegativity and Sparsity Constraints? 7
  1.2.2 Basic NMF Model 8
  1.2.3 Symmetric NMF 9
  1.2.4 Semi-Orthogonal NMF 10
  1.2.5 Semi-NMF and Nonnegative Factorization of Arbitrary Matrix 10
  1.2.6 Three-factor NMF 10
  1.2.7 NMF with Offset (Affine NMF) 13
  1.2.8 Multi-layer NMF 14
  1.2.9 Simultaneous NMF 14
  1.2.10 Projective and Convex NMF 15
  1.2.11 Kernel NMF 16
  1.2.12 Convolutive NMF 16
  1.2.13 Overlapping NMF 17
1.3 Basic Approaches to Estimate Parameters of Standard NMF 18
  1.3.1 Large-scale NMF 21
  1.3.2 Non-uniqueness of NMF and Techniques to Alleviate the Ambiguity Problem 22
  1.3.3 Initialization of NMF 24
  1.3.4 Stopping Criteria 25
1.4 Tensor Properties and Basis of Tensor Algebra 26
  1.4.1 Tensors (Multi-way Arrays) – Preliminaries 26
  1.4.2 Subarrays, Tubes and Slices 27
  1.4.3 Unfolding – Matricization 28
  1.4.4 Vectorization 31
  1.4.5 Outer, Kronecker, Khatri-Rao and Hadamard Products 32
  1.4.6 Mode-\(n\) Multiplication of Tensor by Matrix and Tensor by Vector, Contracted Tensor Product 34
  1.4.7 Special Forms of Tensors 38
1.5 Tensor Decompositions and Factorizations 39
  1.5.1 Why Multi-way Array Decompositions and Factorizations? 40
  1.5.2 PARAFAC and Nonnegative Tensor Factorization 42
  1.5.3 NTF1 Model 47
Contents

1.5.4  NTF2 Model 49
1.5.5  Individual Differences in Scaling (INDSCAL) and Implicit Slice Canonical Decomposition Model (IMCAND) 52
1.5.6  Shifted PARAFAC and Convolutive NTF 53
1.5.7  Nonnegative Tucker Decompositions 55
1.5.8  Block Component Decompositions 59
1.5.9  Block-Oriented Decompositions 62
1.5.10 PARATUCK2 and DEDICOM Models 63
1.5.11 Hierarchical Tensor Decomposition 65
1.6 Discussion and Conclusions 66
Appendix 1.A: Uniqueness Conditions for Three-way Tensor Factorizations 66
Appendix 1.B: Singular Value Decomposition (SVD) and Principal Component Analysis (PCA) with Sparsity and/or Nonnegativity Constraints 67
1.B.1 Standard SVD and PCA 68
1.B.2 Sparse PCA 70
1.B.3 Nonnegative PCA 71
Appendix 1.C: Determining a True Number of Components 71
Appendix 1.D: Nonnegative Rank Factorization Using Wedderburn Theorem – Estimation of the Number of Components 74
References 75

2 Similarity Measures and Generalized Divergences 81
2.1 Error-induced Distance and Robust Regression Techniques 82
2.2 Robust Estimation 84
2.3 Csiszár Divergences 90
2.4 Bregman Divergence 96
2.4.1 Bregman Matrix Divergences 103
2.5 Alpha-Divergences 104
2.5.1 Asymmetric Alpha-Divergences 104
2.5.2 Symmetric Alpha-Divergences 110
2.6 Beta-Divergences 112
2.7 Gamma-Divergences 116
2.8 Divergences Derived from Tsallis and Rényi Entropy 118
2.8.1 Concluding Remarks 119
Appendix 2.A: Information Geometry, Canonical Divergence, and Projection 120
2.A.1 Space of Probability Distributions 120
2.A.2 Geometry of Space of Positive Measures 123
Appendix 2.B: Probability Density Functions for Various Distributions 125
References 127

3 Multiplicative Iterative Algorithms for NMF with Sparsity Constraints 131
3.1 Extended ISRA and EMMI Algorithms: Regularization and Sparsity 132
3.1.1 Multiplicative NMF Algorithms Based on the Squared Euclidean Distance 132
3.1.2 Multiplicative NMF Algorithms Based on Kullback-Leibler 1-Divergence 139
3.2 Multiplicative Algorithms Based on Alpha-Divergence 143
3.2.1 Multiplicative Alpha NMF Algorithm 143
3.2.2 Generalized Multiplicative Alpha NMF Algorithms 147
3.3 Alternating SMART: Simultaneous Multiplicative Algebraic Reconstruction Technique 148
3.3.1 Alpha SMART Algorithm 148
3.3.2 Generalized SMART Algorithms 150
6.6 Numerical Results 319
6.7 Discussions 323
Appendix 6.A: Gradient and Hessian of Cost Functions 324
Appendix 6.B: MATLAB Source Codes 325
References 333

7 Multi-Way Array (Tensor) Factorizations and Decompositions 337

7.1 Learning Rules for the Extended Three-way NTF1 Problem 337
  7.1.1 Basic Approaches for the Extended NTF1 Model 338
  7.1.2 ALS Algorithms for NTF1 340
  7.1.3 Multiplicative Alpha and Beta Algorithms for the NTF1 Model 341
  7.1.4 Multi-layer NTF1 Strategy 343

7.2 Algorithms for Three-way Standard and Super Symmetric Nonnegative Tensor Factorization 344
  7.2.1 Multiplicative NTF Algorithms Based on Alpha- and Beta-Divergences 345
  7.2.2 Simple Alternative Approaches for NTF and SSNTF 350

7.3 Nonnegative Tensor Factorizations for Higher-Order Arrays 351
  7.3.1 Alpha NTF Algorithm 353
  7.3.2 Beta NTF Algorithm 355
  7.3.3 Fast HALS NTF Algorithm Using Squared Euclidean Distance 355
  7.3.4 Generalized HALS NTF Algorithms Using Alpha- and Beta-Divergences 358
  7.3.5 Tensor Factorization with Additional Constraints 360

7.4 Algorithms for Nonnegative and Semi-Nonnegative Tucker Decompositions 361
  7.4.1 Higher Order SVD (HOSVD) and Higher Order Orthogonal Iteration (HOOI) Algorithms 362
  7.4.2 ALS Algorithm for Nonnegative Tucker Decomposition 365
  7.4.3 HOSVD, HOOI and ALS Algorithms as Initialization Tools for Nonnegative Tensor Decomposition 366
  7.4.4 Multiplicative Alpha Algorithms for Nonnegative Tucker Decomposition 366
  7.4.5 Beta NTD Algorithm 370
  7.4.6 Local ALS Algorithms for Nonnegative TUCKER Decompositions 370
  7.4.7 Semi-Nonnegative Tucker Decomposition 374

7.5 Nonnegative Block-Oriented Decomposition 375
  7.5.1 Multiplicative Algorithms for NBOD 376

7.6 Multi-level Nonnegative Tensor Decomposition - High Accuracy Compression and Approximation 377

7.7 Simulations and Illustrative Examples 378
  7.7.1 Experiments for Nonnegative Tensor Factorizations 378
  7.7.2 Experiments for Nonnegative Tucker Decomposition 384
  7.7.3 Experiments for Nonnegative Block-Oriented Decomposition 392
  7.7.4 Multi-Way Analysis of High Density Array EEG – Classification of Event Related Potentials 395
  7.7.5 Application of Tensor Decompositions in Brain Computer Interface – Classification of Motor Imagery Tasks 404
  7.7.6 Image and Video Applications 409

7.8 Discussion and Conclusions 412
Appendix 7.A: Evaluation of Interactions and Relationships Among Hidden Components for NTD model 415
Appendix 7.B: Computation of a Reference Tensor 416
Appendix 7.C: Trilinear and Direct Trilinear Decompositions for Efficient Initialization 418
Appendix 7.D: MATLAB Source Code for Alpha NTD Algorithm
Appendix 7.E: MATLAB Source Code for Beta NTD Algorithm
Appendix 7.F: MATLAB Source Code for HALS NTD Algorithm
Appendix 7.G: MATLAB Source Code for ALS NTF1 Algorithm
Appendix 7.H: MATLAB Source Code for ISRA BOD Algorithm
Appendix 7.I: Additional MATLAB functions
References

8 Selected Applications
8.1 Clustering
  8.1.1 Semi-Binary NMF
  8.1.2 NMF vs. Spectral Clustering
  8.1.3 Clustering with Convex NMF
  8.1.4 Application of NMF to Text Mining
  8.1.5 Email Surveillance
8.2 Classification
  8.2.1 Musical Instrument Classification
  8.2.2 Image Classification
8.3 Spectroscopy
  8.3.1 Raman Spectroscopy
  8.3.2 Fluorescence Spectroscopy
  8.3.3 Hyperspectral Imaging
  8.3.4 Chemical Shift Imaging
8.4 Application of NMF for Analyzing Microarray Data
  8.4.1 Gene Expression Classification
  8.4.2 Analysis of Time Course Microarray Data
References

Index