Maolin Che
Yimin Wei

Theory and Computation of Complex Tensors and its Applications

Springer

Theory and Computation of Complex Tensors and its Applications

Maolin Che • Yimin Wei

Theory and Computation of Complex Tensors and its Applications

Springer

Maolin Che
School of Economics Mathematics
Southwestern University of Finance
and Economics
Chengdu, Sichuan, China

Yimin Wei
School of Mathematical Sciences
Fudan University
Shanghai, China

ISBN 978-981-15-2058-7 ISBN 978-981-15-2059-4 (eBook)
https://doi.org/10.1007/978-981-15-2059-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

This book can be divided into five categories based on the main purposes: (1) the development of tensor spectral theory; (2) the study of tensor complementarity problems using structured tensors; (3) the development and enrichment of the theory of nonnegative tensors; (4) the presentation of new numerical algorithms for solving the real tensor rank-one approximation and computing the US- or U-eigenpairs of complex tensors; (5) the study of randomized algorithms for the computation of the approximate Tucker and Tensor Train decompositions.

There are eight chapters in this book. In Chap. 1, we give some examples to illustrate that tensors can be abstracted from some real and mathematical objects. Some basic operations and definitions of tensors, for example, tensor-matrix multiplication, the Frobenius norm and inner product of tensors, rank-one and symmetric tensors, are introduced. We also provide a summary of the relevant background for the tensor spectral theory, the Perron–Frobenius theorem of nonnegative tensors, and the tensor rank-one approximation problem.

In Chap. 2, we generalize the pseudo-spectral theory of matrices to tensors. We obtain the fundamental properties of the tensor ϵ-pseudo-spectrum, leading to alternative definitions of the tensor ϵ-pseudo-spectrum. We also consider the stability of homogeneous dynamical systems. Similarly, we derive the fundamental properties for the ϵ-pseudo-spectrum of tensor polynomial eigenvalue problems. Furthermore, we discuss the implications of the ϵ-pseudo-spectrum on computing the backward errors of an approximate eigenpair of a tensor polynomial and the distance from a regular tensor polynomial to its nearest irregular tensor polynomial.

In Chap. 3, we analyze the perturbation of tensor eigenvalue problems. We consider the first-order perturbation results for the algebraically simple Z- and H-eigenvalues of tensors and H-eigenvalues of tensor polynomials with relative Frobenius normwise or componentwise perturbations. Based on the perturbation for the algebraically simple Z-eigenvalue of a symmetric real tensor and mode-symmetric embedding, we obtain the perturbation of the algebraically simple singular value of a real tensor. Specifically, we focus on the perturbation for the smallest eigenvalue of an irreducible and symmetric nonsingular $\mathcal{M}$-tensor for relative componentwise perturbations.

In Chap. 4, we first analyze the first-order necessary conditions for the solution of the tensor complementarity problem. From the properties of copositive tensors, we prove that the problem with copositive tensors has a nonempty and compact solution set. We also consider a special case via structured tensors.

In Chap. 5, we introduce the sign nonsingular tensors and derive the relationship between the combinatorial determinant and the permanent of nonnegative tensors. We generalize the results from doubly stochastic matrices to totally plane stochastic tensors and obtain a probabilistic algorithm for locating a positive diagonal in a nonnegative tensor. We obtain a normalization algorithm to convert some nonnegative tensors to plane stochastic tensors. We obtain a lower bound for the minimum of the axial N-index assignment problem using the set of plane stochastic tensors.

Chapter 6 deals with the local optimal rank-one approximation of a real tensor via neural networks. We prove that the solution of the neural network is locally asymptotically stable in the sense of Lyapunov stability theory. We define the tensor restricted singular pairs and present several numerical algorithms for computing them. Similarly, we use the neural networks for the computation of the local optimal generalized H-eigenpairs of symmetric-definite tensor pairs.

Chapter 7 presents the iterative algorithms (QRCST or QRCT) for computing the US- and U-eigenpairs of complex tensors. Specifically, we derive a higher order power type method for computing a US- or a U-eigenpair, similar to the higher-order power method for computing the best rank-one approximation of a real tensor.

In Chap. 8, we design adaptive randomized algorithms for computing the approximate tensor decompositions. For a low multilinear rank approximation of a real tensor with unknown multilinear rank, we analyze its probabilistic error bound under certain assumptions. Finally, we also consider the tensor train approximations of the tensors. Based on the bounds of the singular values of sub-Gaussian matrices with independent columns or independent rows, we analyze these randomized algorithms. Several illustrated numerical examples are provided.

Chapters 2 and 3 are based closely on [1–3]. The main content in Chap. 4 comes from [4]. Chapters 5, 6, and 7 are adopted from [5–7]. Chapter 8 is from [8].

In this book, the computations are carried out in MATLAB Version 2013a and the MATLAB Tensor Toolbox [9] on a laptop with an Intel Core i5-4200M CPU (2.50 GHz) and a 8.00 GB RAM. All floating point numbers in each example have four digits after the decimal point. For $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, we assume that "all i_n" and "all n" mean $i_n = 1, 2, \ldots, I_n$ and $n = 1, 2, \ldots, N$, respectively; for $\mathcal{A} \in CT_{N,I}$, we assume that "all i_n" means $i_n = 1, 2, \ldots, I$ for all n. We assume that "all l" means "$l = 0, 1, \ldots, L$".

We would like to thank Prof. Andrzej Cichocki for computing the tensor rank-one approximation via neural network models, Prof. Guoyin Li for the tensor pseudo-spectral theory, and Prof. Changjiang Bu for the study of plane stochastic tensors.

We also thank Prof. Liqun Qi of the Hong Kong Polytechnic University, who leads us to the research of tensor spectral theory and always encourages us to explore the research. We would like to thank Prof. Eric King-wah Chu and Prof. Sanzheng Qiao, who read this book carefully and provide feedback during the writing process.

M. Che is supported by the Fundamental Research Funds for the Central Universities under grant JBK1804026 and the National Natural Science Foundation of China under grant 11901471. Y. Wei is supported by the National Natural Science Foundation of China under grant 11771099, the Innovation Program of Shanghai Municipal Education Commission, Key Laboratory of Mathematics for Nonlinear Sciences of Fudan University, and Shanghai Key Laboratory of Contemporary Applied Mathematics.

Chengdu, China — Maolin Che
Shanghai, China — Yimin Wei
July, 2018

References

1. M. Che, Y. Wei, An inequality for the Perron pair of an irreducible and symmetric nonnegative tensor with application. J. Oper. Res. Soc. China **5**(1), 65–82 (2017)
2. M. Che, G. Li, L. Qi, Y. Wei, Pseudo-spectra theory of tensors and tensor polynomial eigenvalue problems. Linear Algebra Appl. **533**(15), 536–572 (2017)
3. M. Che, L. Qi, Y. Wei, Perturbation bounds of tensor eigenvalue and singular value problems with even order. Linear Multilinear Algebra **64**(4), 622–652 (2016)
4. M. Che, L. Qi, Y. Wei, Positive-definite tensors to nonlinear complementarity problems. J. Optim. Theory Appl. **168**(2), 475–487 (2016)
5. M. Che, C. Bu, L. Qi, Y. Wei, Nonnegative tensors revisited: plane stochastic tensors. Linear Multilinear Algebra **67**(7), 1364–1391 (2019)
6. M. Che, A. Cichocki, Y. Wei, Neural networks for computing best rank-one approximations of tensors and its applications. Neurocomputing **267**(3), 114–133 (2017)
7. M. Che, L. Qi, Y. Wei, Iterative algorithms for computing US- and U-eigenpairs of complex tensors. J. Comput. Appl. Math. **317**, 547–564 (2017)
8. M. Che, Y. Wei, Randomized algorithms for the approximations of Tucker and the tensor train decompositions. Adv. Comput. Math. **45**(1), 395–428 (2019)
9. B. Bader, T. Kolda, *Matlab Tensor Toolbox Version 2.6* (2015). http://www.sandia.gov/~tgkolda/TensorToolbox/

Contents

Chapter 1
Introduction

An increasing number of applications in signal processing, data analysis and higher-order statistics, as well as independent component analysis [1–5] involve the manipulation of data whose elements are addressed by more than two indices. In the literature, these higher-order extensions of vectors (first-order) and matrices (second-order) are called *higher-order tensors*, *multi-dimensional matrices*, or *multiway arrays*.

Tensor problems have wide applications in chemometrics, signal processing and high order statistics [4]. For the theory and applications of tensors, we refer to Comon et al. [6], Kolda and Bader [7], Cichocki et al. [8], Yang and Yang [9], Qi and Luo [10], Wei and Ding [11], and Qi et al. [12].

We use $\mathbb{C}$ or $\mathbb{R}$ to denote the complex or real field. A tensor is an Nth-order array of numbers denoted by script notation $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ with entries given by $a_{i_1 i_2 \ldots i_N} \in \mathbb{C}$ for all i_n and n. When all the I_n are the same, i.e., $I_n = I$ for all n, $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is called an Nth-order I-dimensional complex tensor. We use $RT_{N,I}$ (or $CT_{N,I}$) to denote the set of all Nth-order I-dimensional real (or complex) tensors.

1.1 Examples for Tensors

Tensors can be abstracted from some real and mathematical objects.

Example 1.1.1 (Homogeneous Polynomials) Given $\mathbf{c} \in \mathbb{R}^I$, the first-degree homogeneous polynomial with respect to $\mathbf{x} \in \mathbb{R}^I$ can be represented as $\mathbf{c}^\top \mathbf{x} = c_1 x_1 + c_2 x_2 + \cdots + c_I x_I$. Given $\mathbf{A} \in \mathbb{R}^{I \times J}$, the second-degree homogeneous polynomial with respect to $\mathbf{x} \in \mathbb{R}^I$ and $\mathbf{y} \in \mathbb{R}^J$ can be represented as $\mathbf{x}^\top \mathbf{A} \mathbf{y} = \sum_{i=1}^{I} \sum_{j=1}^{J} a_{ij} x_i y_j$.

M. Che, Y. Wei, *Theory and Computation of Complex Tensors and its Applications*,
https://doi.org/10.1007/978-981-15-2059-4_1

In general, the Nth-degree homogeneous polynomial, with respect to $\mathbf{x}_1 \in \mathbb{R}^{I_1}$, $\mathbf{x}_2 \in \mathbb{R}^{I_2}, \ldots, \mathbf{x}_N \in \mathbb{R}^{I_N}$, can be denoted by

$$p_N(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} a_{i_1 i_2 \ldots i_N} x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N},$$

where $a_{i_1 i_2 \ldots i_N}$ is the coefficient of the term $x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N}$ in $p_N(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ and x_{n,i_n} is the i_nth entry of $\mathbf{x}_n$ for all i_n and n. If we set a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ such that its $(i_1, i_2, \ldots, i_N)$-entry is $a_{i_1 i_2 \ldots i_N}$, then the polynomial can be expressed as

$$p_N(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \mathcal{A} \times_1 \mathbf{x}_1^\top \times_2 \mathbf{x}_2^\top \cdots \times_N \mathbf{x}_N^\top.$$

The tensor-vector multiplication will be introduced in the next section.

Example 1.1.2 (The Discretization of Multivariate Functions) Suppose that

$$f(x, y, z) = \frac{1}{x + y + z}, \quad g(x, y, z) = \cos(x + y + z),$$

$$h(x, y, z) = \exp((-0.01 + 4\pi\iota)(x + y + z - 2)) + \exp((-0.02 + 4.2\pi\iota)(x + y + z - 2)),$$

where $1 < x, y, z < 2$ and $\iota = \sqrt{-1}$.

Let $\{x_1, x_2, \ldots, x_I\}$ be any monotonically increasing sequence in the open interval $(1, 2)$. When the values of x, y and z are chosen from $\{x_1, x_2, \ldots, x_I\}$, we define three tensors $\mathcal{A}, \mathcal{B} \in RT_{3,I}$ and $\mathcal{C} \in CT_{3,I}$ such that

$$a_{ijk} := f(x_i, x_j, x_k), \quad b_{ijk} = g(x_i, x_j, x_k), \quad c_{ijk} = h(x_i, x_j, x_k).$$

From the definitions of f, g and h, it is clear that $\mathcal{A} \in RT_{3,I}$ is symmetric and positive, $\mathcal{B} \in RT_{3,I}$ is symmetric but not nonnegative and $\mathcal{C} \in CT_{3,I}$ is complex symmetric.

Example 1.1.3 (The Associated Tensors of Uniform Hypergraphs) Analogous to spectral graph theory [13], adjacency tensors and Laplacian tensors have been introduced in spectral hypergraph theory. The notations related to the hypergraph can be referred to [14].

A hypergraph $\mathbb{H}$ is a pair $(\mathbb{V}, \mathbb{E})$, where $\mathbb{E} \subseteq \mathfrak{P}(\mathbb{V})$. The elements of $\mathbb{V} = \mathbb{V}(\mathbb{H})$ are referred to as vertices and the elements of $\mathbb{E} = \mathbb{E}(\mathbb{H})$ are called hyperedges. A hypergraph $\mathbb{H}$ is said to be N-uniform for an integer $N \geq 2$, if, for all $\mathbf{e} \in \mathbb{E}(\mathbb{H})$, the cardinality of $\mathbf{e}$ is N. Such an N-uniform hypergraph is also called an N-graph.

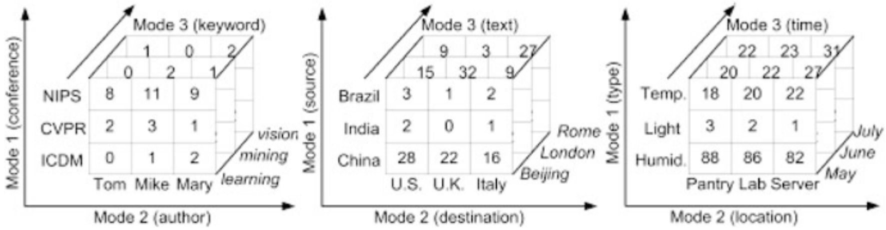

Fig. 1.1 Examples of the third-order tensor data

For a given N-uniform hypergraph $\mathbb{H} = (\mathbb{V}, \mathbb{E})$, the adjacency tensor $\mathcal{A} \in RT_{N,I}$, with respect to $\mathbb{H}$, is defined as [15–18]:

$$a_{i_1 i_2 \ldots i_N} = \frac{1}{(N-1)!} \begin{cases} 1, & \text{if } \{i_1, i_2, \ldots, i_N\} \in \mathbb{E}; \\ 0, & \text{otherwise.} \end{cases}$$

Obviously, $a_{i_1 i_2 \ldots i_N} = 0$ if at least two indices are the same. Note that $\mathcal{A}$ is symmetric and nonnegative. The degree of $i \in \mathbb{V}$ is defined as $d(i) = |\{e_{\mathbf{p}} : i \in \mathbf{p} \in \mathbb{V}\}|$. We assume that every vertex has at least one edge. Thus, $d(i) > 0$ for $i \in \mathbb{V}$. The degree tensor $\mathcal{D}(\mathbb{H})$ is an Nth-order I-dimensional diagonal tensor, with its main diagonal elements as $d(i)$. The Laplacian tensor $\mathcal{L}$ of $\mathbb{H}$ is defined by $\mathcal{D}(\mathbb{H}) - \mathcal{A}$. The signless Laplacian tensor $\mathcal{L}^+$ of $\mathbb{H}$ is defined by $\mathcal{D}(\mathbb{H}) + \mathcal{A}$ [19–22]. Note that $\mathcal{L}$ is symmetric and nonnegative and $\mathcal{L}^+$ is symmetric.

Example 1.1.4 (Description of Complex Social Networks [23]) Third-order tensors are three-dimensional (3D), with some examples shown in Fig. 1.1. The left part of Fig. 1.1 illustrates social network analysis data organized in three modes of conference, author and keyword. The web graph mining data organized in three modes of source, destination and text is demonstrated in the middle part of Fig. 1.1. Lastly, the environmental sensor monitoring data organized in the three modes of type, location and time is demonstrated in the right part of Fig. 1.1.

Generally speaking, there are two kinds of tensors: a data structure, which admits different dimensions according to the complexity of the data; or an operator, where it possesses different meanings in different situations. All tensors mentioned in this book can be viewed as a data structure.

1.2 Basics of Tensors

Throughout this book, I, J, and N are reserved to denote the index upper bounds, unless stated otherwise. We use small letters $x, u, v, \ldots$ for scalars, small bold letters $\mathbf{x}, \mathbf{u}, \mathbf{v}, \ldots$ for vectors, bold capital letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ for matrices, and calligraphic letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ for higher-order tensors. This notation is consistently used for lower-order parts of a given structure. For example, the entry with row index i and column index j in a matrix $\mathbf{A}$, i.e., $(\mathbf{A})_{ij}$, is written by a_{ij} (also $(\mathbf{x})_i = x_i$ and $(\mathcal{A})_{i_1 i_2 \ldots i_N} = a_{i_1 i_2 \ldots i_N}$).

We use $\cdot^\top$, $\bar{\cdot}$, $|\cdot|$ and $\cdot^*$ to denote the transpose, complex conjugate, element-wise modulus and complex conjugated transpose, respectively. We use $\|\cdot\|_2$ and $\|\cdot\|_F$ to denote the 2-norm and the Frobenius norm, respectively. We use $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$ to denote the real and imaginary parts of a vector $\mathbf{z} \in \mathbb{C}^I$, respectively. The argument of $z \in \mathbb{C}$ is denoted by $\arg(z) \in (-\pi, \pi]$. Also, $\mathbf{0}_I \in \mathbb{C}^I$ is the zero vector. We use $\mathbf{I}_I \in \mathbb{C}^{I \times I}$ and $\mathbf{0}_{I \times J} \in \mathbb{C}^{I \times J}$ to denote the identity and the zero matrices, respectively. Lastly, $\iota = \sqrt{-1}$ and $\mathbb{S}_I$ are the imaginary unit and the symmetric group on the set $\{1, 2, \ldots, I\}$, respectively.

We use parentheses to denote the concatenation of two or more vectors, e.g., $(\mathbf{a}, \mathbf{b})$ is equivalent to $(\mathbf{a}^\top, \mathbf{b}^\top)^\top$, where $\mathbf{a} \in \mathbb{C}^I$ and $\mathbf{b} \in \mathbb{C}^J$ are two column vectors. Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I$, $\mathbf{x} > \mathbf{y}$ and $\mathbf{x} \geq \mathbf{y}$ mean $x_i > y_i$ and $x_i \geq y_i$, respectively, for all i. Similarly, we can also define $\mathbf{x} < \mathbf{y}$ and $\mathbf{x} \leq \mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I$. Finally, we introduce the following notations:

$$\mathbb{R}_+^{I_1 \times I_2 \times \cdots \times I_N} := \{\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} : a_{i_1 i_2 \ldots i_N} \geq 0, i_n = 1, 2, \ldots, I, n = 1, 2, \ldots, N\};$$

$$\mathbb{R}_{++}^{I_1 \times I_2 \times \cdots \times I_N} := \{\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} : a_{i_1 i_2 \ldots i_N} > 0, i_n = 1, 2, \ldots, I, n = 1, 2, \ldots, N\}.$$

$\mathbb{R}_+^{I_1 \times I_2 \times \cdots \times I_N}$ is called the set of all nonnegative tensors and $\mathbb{R}_{++}^{I_1 \times I_2 \times \cdots \times I_N}$ is called the set of all positive tensors in $\mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. In particular, when $N = 1$, we have $\mathbb{R}_+^I := \{\mathbf{x} \in \mathbb{R}^I : x_i \geq 0, i = 1, 2, \ldots, I\}$ and $\mathbb{R}_{++}^I := \{\mathbf{x} \in \mathbb{R}^I : x_i > 0, i = 1, 2, \ldots, I\}$. For a given $\mathbf{x} \in \mathbb{C}^I$, $\operatorname{diag}(\mathbf{x})$ denotes the diagonal matrix whose main diagonal entries are the entries of $\mathbf{x}$.

1.2.1 Basic Operations

The mode-n product [7] of a complex tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ by a matrix $\mathbf{B} \in \mathbb{C}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{B}$, is a tensor $\mathcal{C} \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times J_n \times I_{n+1} \times \cdots \times I_N}$, with entries

$$c_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 \ldots i_N} b_{j i_n},$$

for all i_n and n.

In particular, the mode-n multiplication of a complex tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ by a vector $\mathbf{x} \in \mathbb{C}^{I_n}$ is denoted by $\mathcal{A} \bar{\times}_n \mathbf{x}$. If we set $\mathcal{C} = \mathcal{A} \times_n \mathbf{x}^\top \in$

$\mathbb{C}^{I_1\times\cdots\times I_{n-1}\times I_{n+1}\times\cdots\times I_N}$, then we have element-wise [7],

$$c_{i_1\ldots i_{n-1}i_{n+1}\ldots i_N}=\sum_{i_n=1}^{I_n}a_{i_1\ldots i_{n-1}i_n i_{n+1}\ldots i_N}x_{i_n}.$$

For any given tensor $\mathcal{A}\in\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ and the matrices $\mathbf{F}\in\mathbb{C}^{J_n\times I_n}$ and $\mathbf{G}\in\mathbb{C}^{J_m\times I_m}$, one has [7]

$$\begin{cases}(\mathcal{A}\times_n\mathbf{F})\times_m\mathbf{G}=(\mathcal{A}\times_m\mathbf{G})\times_n\mathbf{F}=\mathcal{A}\times_n\mathbf{F}\times_m\mathbf{G};\\ (\mathcal{A}\times_n\mathbf{F})\times_n\mathbf{G}=\mathcal{A}\times_n(\mathbf{G}\cdot\mathbf{F}),\quad\text{with}\quad J_n=I_m,\end{cases}$$

with different integers m and n, where '$\cdot$' represents the multiplication of two matrices.

For a given $\mathbf{x}\in\mathbb{C}^I$ and $\mathcal{A}\in CT_{N,I}$, we introduce the following two notations [24]:

$$\mathcal{A}\mathbf{x}^{N-1}:=\mathcal{A}\times_2\mathbf{x}^\top\cdots\times_N\mathbf{x}^\top,\ \mathcal{A}\mathbf{x}^N:=\mathcal{A}\times_1\mathbf{x}^\top\times_2\mathbf{x}^\top\cdots\times_N\mathbf{x}^\top.$$

Scalar products and the Frobenius norm of a tensor are extensions of the well-known definitions, from matrices to tensors of arbitrary order [7, 25]. Suppose that $\mathcal{A},\mathcal{B}\in\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$, the scalar product $\langle\mathcal{A},\mathcal{B}\rangle$ is defined as [25]

$$\langle\mathcal{A},\mathcal{B}\rangle=\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\cdots\sum_{i_N=1}^{I_N}b_{i_1i_2\ldots i_N}\overline{a_{i_1i_2\ldots i_N}},$$

and the Frobenius norm of a tensor $\mathcal{A}$ is given by $\|\mathcal{A}\|_F=\sqrt{\langle\mathcal{A},\mathcal{A}\rangle}$.

1.2.2 Structured Tensors

We recommend [10] for a thorough survey of structured tensors. For given N vectors $\mathbf{x}_n\in\mathbb{C}^{I_n}$ with all n, if the entries of $\mathcal{A}\in\mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ can be represented as $a_{i_1i_2\ldots i_N}=x_{1,i_1}x_{2,i_2}\ldots x_{N,i_N}$, where x_{n,i_n} is the i_nth element of $\mathbf{x}_n$, then $\mathcal{A}$ is a complex rank-one tensor [26, 27] given by

$$\mathcal{A}=\mathbf{x}_1\otimes\mathbf{x}_2\otimes\cdots\otimes\mathbf{x}_N,$$

where "$\otimes$" denotes the outer (tensor) product. If $\mathbf{x}_n\in\mathbb{R}^{I_n}$ with all n, then $\mathcal{A}$ is a real rank-one tensor.

For any $\mathcal{A}\in CT_{N,I}$, $\mathcal{A}$ is *complex symmetric* [28], if $a_{i_1i_2\ldots i_N}$ is invariant by any permutation $\pi\in\mathbb{S}_I$, that is, $a_{i_1i_2\ldots i_N}=a_{i_{\pi(1)}i_{\pi(2)}\ldots i_{\pi(N)}}$ for all i_n and n. For any

$\mathcal{A} \in RT_{N,I}$, $\mathcal{A}$ is *real symmetric* [24, 29], if $a_{i_1 i_2 \ldots i_N}$ is invariant by any permutation π, that is, $a_{i_1 i_2 \ldots i_N} = a_{i_{\pi(1)} i_{\pi(2)} \ldots i_{\pi(N)}}$, for all i_n and n. For a given vector $\mathbf{x} \in \mathbb{C}^I$, if the entries of $\mathcal{A} \in CT_{N,I}$ can be represented as

$$a_{i_1 i_2 \ldots i_N} = x_{i_1} x_{i_2} \ldots x_{i_N},$$

where x_{i_n} is the i_nth element of $\mathbf{x}$, then $\mathcal{A}$ is a complex symmetric rank-one tensor given by

$$\mathcal{A} = \underbrace{\mathbf{x} \otimes \mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{N} := \mathbf{x}^{\otimes N}.$$

If $\mathbf{x} \in \mathbb{R}^I$, then $\mathbf{x}^{\otimes N}$ is a real symmetric rank-one tensor.

For any given symmetric tensor $\mathcal{A} \in RT_{N,I}$ with an even N, if $\mathcal{A}\mathbf{x}^N > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^I$, then $\mathcal{A}$ is positive definite [24]; if $\mathcal{A}\mathbf{x}^N \geq 0$ for all $\mathbf{x} \in \mathbb{R}^I$, then $\mathcal{A}$ is positive semi-definite [24]; if $\mathcal{A}\mathbf{x}^N \geq 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^I_+$, then $\mathcal{A}$ is strictly copositive [30]; if $\mathcal{A}\mathbf{x}^N \geq 0$ for all $\mathbf{x} \in \mathbb{R}^I_+$, then $\mathcal{A}$ is copositive [30].

When we consider positive definite tensors and copositive tensors, symmetry is unnecessary. According to [31], we can relate a tensor $\mathcal{A} \in RT_{N,I}$ to a symmetric tensor $\mathcal{B} \in RT_{N,I}$ as follows. For any $\mathcal{A} \in RT_{N,I}$, there is a unique symmetric tensor $\mathcal{B} \in RT_{N,I}$ such that $\mathcal{A}\mathbf{x}^N = \mathcal{B}\mathbf{x}^N$ for all $\mathbf{x} \in \mathbb{R}^I$.

We call an index set $\{i_1, i_2, \ldots, i_N\}$ a permutation of another index set $\{k_1, k_2, \ldots, k_N\}$ if the former is a rearrangement of the latter, denoting this operation by π, that is $\pi(k_1, k_2, \ldots, k_N) = \{i_1, i_2, \ldots, i_N\}$. Denote the set of all distinct permutations of an index set $\{k_1, k_2, \ldots, k_N\}$ by $\Sigma(k_1, k_2, \ldots, k_N)$.

Note that $|\Sigma(k_1, k_2, \ldots, k_N)|$, the cardinality of $\Sigma(k_1, k_2, \ldots, k_N)$, is variant for different index sets. Then the entries of $\mathcal{B}$ are given as

$$b_{j_1 j_2 \ldots j_N} = \frac{\sum_{\pi \in \Sigma(k_1, k_2, \ldots, k_N)} a_{\pi(k_1, k_2, \ldots, k_N)}}{|\Sigma(k_1, k_2, \ldots, k_N)|}$$

for all i_n, j_n and n. Here, we call $\mathcal{B}$ a symmetrization of $\mathcal{A}$.

Suppose that $\mathcal{A} \in CT_{N,I}$, the I-tuple $\{a_{1\pi_2(1)\ldots\pi_N(1)}, a_{2\pi_2(2)\ldots\pi_N(2)}, \ldots, a_{I\pi_2(I)\ldots\pi_N(I)}\}$ is a *diagonal* [32] of the tensor $\mathcal{A}$ associated with $\pi_n \in \mathbb{S}_I$ and $n = 2, 3, \ldots, N$. In particular, $\{a_{11\ldots1}, a_{22\ldots2}, \ldots, a_{II\ldots I}\}$ is the main diagonal of the tensor $\mathcal{A}$. A diagonal is *positive*, if its elements are positive.

The product $\prod_{i=1}^{I} a_{i\pi_2(i)\ldots\pi_N(i)}$ is the *diagonal product* [32] of the tensor $\mathcal{A}$ associated with $\pi_n \in \mathbb{S}_I$ with $n = 2, 3, \ldots, N$. Meanwhile, the sum $\sum_{i=1}^{I} a_{i\pi_2(i)\ldots\pi_N(i)}$ is the *diagonal sum* of $\mathcal{A}$ associated with $\pi_n \in \mathbb{S}_I$ with $n = 2, 3, \ldots, N$.

A tensor $\mathcal{A} \in RT_{N,I}$ is *nonnegative* [33], if the elements are nonnegative, and we denote the set of all nonnegative tensors by $NT_{N,I}$; a tensor $\mathcal{D} \in RT_{N,I}$ is *diagonal* [24], if the entries not in the main diagonal are zero. In particular, if the entries on the main diagonal of any diagonal tensor $\mathcal{A} \in CT_{N,I}$ are 1, then $\mathcal{A}$ is the unit tensor or the identity tensor, denoted by $\mathcal{I}$.

Analogous to the reducible matrices [34, Chapter 2], $\mathcal{A} \in RT_{N,I}$ is *reducible* [33], if there exists a nonempty proper index subset $\mathbb{I} \subset \{1, 2, \ldots, I\}$ such that

$$a_{i_1 \ldots i_N} = 0, \text{ for all } i_1 \in \mathbb{I} \text{ and } i_2, \ldots, i_N \notin \mathbb{I}.$$

Otherwise, $\mathcal{A}$ is *irreducible*. Similarly, we can define the irreducibility of any $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$.

Qi [24] introduces the principal symmetric sub-tensors of any symmetric tensor $\mathcal{A} \in RT_{N,I}$ and proves that if $\mathcal{A}$ is positive definite, then all of its principal symmetric sub-tensors are also positive definite. Now, we introduce the definition of a sub-tensor of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$.

Suppose that $L_n \leq I_n$ are positive integers for all n. We use $\mathbb{Q}_{L_n, I_n}$ to denote the set of $C_{I_n}^{L_n}$ increasing sequences $\omega = (\omega_1, \omega_2, \ldots, \omega_{L_n})$ such that $1 \leq \omega_1 < \omega_2 < \ldots < \omega_{L_n} \leq I_n$, where $C_{I_n}^{L_n} = \frac{I_n!}{L_n!(I_n - L_n)!}$. If $\alpha_n \in \mathbb{Q}_{P_n, I_n}$, where $P_n \leq I_n$ is a positive integer with $n = 1, 2, \ldots, N$, then $\mathcal{A}[\alpha_1|\alpha_2|\ldots|\alpha_N] \in \mathbb{C}^{P_1 \times P_2 \times \cdots \times P_N}$ is a sub-tensor of $\mathcal{A}$, whose $(i_1, i_2, \ldots, i_N)$-entry is $a_{\alpha_{1,i_1}, \alpha_{2,i_2}, \ldots, \alpha_{N,i_N}}$.

The $(i_1, \ldots, i_{m-1}, i_{m+1}, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N)$th mode-$(m, n)$ slice of $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as a matrix in $\mathbb{C}^{I_m \times I_n}$, denoted by $\mathcal{A}_{m,n}^{(i)}$, equaling $\mathcal{A} \times_1 \mathbf{e}_{1,i_1}^\top \cdots \times_{m-1} \mathbf{e}_{m-1,i_{m-1}}^\top \times_{m+1} \mathbf{e}_{m+1,i_{m+1}}^\top \cdots \times_{n-1} \mathbf{e}_{n-1,i_{n-1}}^\top \times_{n+1} \mathbf{e}_{n+1,i_{n+1}}^\top \cdots \times_N \mathbf{e}_{N,i_N}^\top$, where $\mathbf{e}_{n,i_n}$ is the i_nth column of $\mathbf{I}_{I_n} \in \mathbb{R}^{I_n \times I_n}$ for all i_n and $m < n$. Note that for a given i, all the $(i, i, \ldots, i)$th mode-(m, n) slices, or the ith mode-(m, n) slices, of any complex symmetric tensor $\mathcal{A} \in CT_{N,I}$ are the same complex symmetric matrix with all $m < n$.

1.3 Basic Results

1.3.1 Tensor Spectral Theory

The eigenvalue problem of tensors can be regarded as the generalizations of the eigenvalue problem of matrices (i.e., matrix standard eigenvalue problems, matrix generalized eigenvalue problems and matrix polynomial eigenvalue problems).The eigenvalue problem of tensors are widely used in polynomial optimization [35], spectral hypergraph theory [16, 36], higher-order Markov chain [37], image science [38] and other fields. Very recently, the eigenvalue problem of tensors, positive semi-definite tensors and copositive tensors have been used to study some physical problems, such as the quantum spin state, the quantum field theory and liquid crystals [39–41].

In 2005, Qi [24] defines two kinds of eigenvalues and investigates relative results similar to the matrix eigenvalues. Independently, Lim [29] proposes another definition of eigenvalues, eigenvectors, singular values, and singular vectors for tensors based on a constrained variational approach, in the flavor of the Rayleigh quotient for symmetric matrix eigenvalues [42, Chapter 8].

Definition 1.3.1 ([24]) Suppose that $\mathcal{A} \in RT_{N,I}$ is symmetric. If there exist a nonzero vector $\mathbf{x} \in \mathbb{C}^I$ and $\lambda \in \mathbb{C}$ such that $\mathcal{A}\mathbf{x}^{N-1} = \lambda \mathbf{x}^{[N-1]}$, where $\mathbf{x}^{[N-1]} = (x_1^{N-1}, x_2^{N-1}, \dots, x_I^{N-1})^\top$, then $(\lambda; \mathbf{x})$ is called an eigenpair of $\mathcal{A}$. The spectrum $\Lambda(\mathcal{A})$, and the spectral radius $\rho(\mathcal{A})$ of $\mathcal{A}$ are defined as

$$\Lambda(\mathcal{A}) = \{\lambda : \lambda \text{ is an eigenvalue of } \mathcal{A}\}, \quad \rho(\mathcal{A}) = \max\{|\lambda| : \lambda \in \Lambda(\mathcal{A})\}.$$

Moreover, if $\mathbf{x} \in \mathbb{R}^I$ and $\lambda \in \mathbb{R}$, then $(\lambda; \mathbf{x})$ is called an H-eigenpair of $\mathcal{A}$.

Definition 1.3.2 Suppose that $\mathcal{A} \in RT_{N,I}$ is symmetric. If there exist a nonzero vector $\mathbf{x} \in \mathbb{C}^I$ and $\lambda \in \mathbb{C}$ such that $\mathcal{A}\mathbf{x}^{N-1} = \lambda \mathbf{x}$ and $\mathbf{x}^*\mathbf{x} = 1$, then $(\lambda; \mathbf{x})$ is called an E-eigenpair of $\mathcal{A}$ [43, Definition 5.1.1].

Moreover, if $\mathbf{x} \in \mathbb{R}^I$ is unit and $\lambda \in \mathbb{R}$, then $(\lambda; \mathbf{x})$ is called a Z-eigenpair of $\mathcal{A}$.

Note that the E-eigenpair of any real symmetric tensor, defined by Definition 1.3.2, is different from the E-eigenpair in [24].

In 2008, Qi et al. [44] introduce the D-eigenvalues for a diffusion kurtosis tensor and indicate that the largest, the smallest and the average D-eigenvalues correspond with the largest, the smallest and the average apparent kurtosis coefficients of a water molecule in the space, respectively. The strong ellipticity condition plays an important role in nonlinear elasticity and in materials. In 2009, Qi et al. [45] define the M-eigenvalues for an elasticity tensor and prove that the strong ellipticity condition holds if and only if the smallest M-eigenvalue is positive. Hu et al. [35] investigate properties of the determinants of tensors, and their applications. It is pointed out in [33, 46–48] that the generalized eigenvalue framework unifies several definitions of eigenvalues of tensors, such as eigenvalues and H-, E-, Z- and D-eigenvalues. Ding and Wei [11, 49] focus on the properties and perturbations of the spectra of regular tensor pairs and extend several classical results from matrices or matrix pairs to tensor pairs.

Kolda and Mayo [43] derive a shifted symmetric higher-order power method (SS-HOPM) for computing the Z-eigenpairs of real symmetric tensors and indicate that SS-HOPM can be viewed as a generalization of the power iteration method for matrices or the symmetric higher-order power method. Kolda and Mayo [48] present the adaptive power method for solving the tensor generalized eigenvalue problem associated with symmetric positive tensor pairs, which is an extension of SS-HOPM for finding the Z-eigenpairs. Cui et al. [47] propose a new approach for computing all real eigenvalues (that is, Z- and H-eigenvalues) of real symmetric tensors sequentially, from the largest to the smallest. Chen et al. [50] derive an upper bound for the number of equivalence classes of generalized tensor eigenpairs using mixed volumes. Based on this bound and the structures of tensor eigenvalue problems, they propose two homotopy algorithms for the tensor eigenvalue problem. Using the state-of-the-art L-BFGS approach, Chang et al. [51] develop a first-order optimization algorithm for computing the H- and Z-eigenvalues of large scale sparse real symmetric tensors. Batselier and Wong [52] derive the QR algorithm for computing the Z-eigenpairs of a real symmetric tensor, based on the symmetric QR algorithm for the real symmetric matrix eigenvalue problem.

Chang et al. [33] generalize the Perron-Frobenius theorem for nonnegative matrices to the class of nonnegative tensors. We state the Perron-Frobenius theorem for nonnegative tensors as follows:

Theorem 1.3.1 *If* $\mathcal{A} \in RT_{N,I}$ *is nonnegative, then there exist* $\lambda_0 \geq 0$ *and a nonzero* $\mathbf{x}_0 \in \mathbb{R}^I_+$ *such that*

$$\mathcal{A}\mathbf{x}_0^{N-1} = \lambda_0 \mathbf{x}_0^{[N-1]}. \tag{1.3.1}$$

Theorem 1.3.2 *If* $\mathcal{A} \in RT_{N,I}$ *is irreducible nonnegative, then* $(\lambda_0; \mathbf{x}_0)$ *in* (1.3.1) *satisfies:* (1) λ_0 *is an* H*-eigenvalue;* (2) *all components of* $\mathbf{x}_0$ *are positive;* (3) *if* λ *is an eigenvalue with nonnegative eigenvector, then* $\lambda = \lambda_0$*; moreover, the nonnegative eigenvector is unique up to a multiplicative constant;* (4) *if* λ *is an eigenvalue of* $\mathcal{A}$*, then* $|\lambda| \leq \lambda_0$.

More similar results can be found in [53, 54]. Ng et al. [37] derive an iterative method (denoted by NQZ) for computing the spectral radius of an irreducible nonnegative tensor. Zhang and Qi [55] establish an explicit linear convergence rate of the NQZ method for nonnegative tensors under certain conditions. Liu et al. [56] propose an inverse iterative method for computing the Perron pair of an irreducible nonnegative third-order tensor and prove that this method converges quadratically and is positivity preserving in the sense that the vectors approximating the Perron vector are strictly positive in each iteration. By combining the idea of Newton's method with the idea of the Noda iteration, Liu et al. [57] present a Newton-Noda iteration (NNI) for computing the Perron pair of a weakly irreducible nonnegative tensor. A survey on eigenvalues of nonnegative tensors can be found in [58]. Li and Ng [59, 60] extend the well-known column sum bound of the spectral radius for nonnegative matrices to the tensor case, and also derive an upper bound of the spectral radius for a nonnegative tensor via the largest eigenvalue of a symmetric tensor. Chen et al. [61] introduce three new classes of symmetric nonnegative tensors and discuss their properties and applications in the context of polynomial and tensor optimization.

1.3.2 Real Tensor Rank-One Approximations

The rank-one approximation of a real tensor is a special case of tensor low-rank approximations. The common tensor low-rank approximations consist of approximated canonical polyadic (CP) decompositions, approximated Tucker decompositions and approximated tensor train (TT) decompositions. We recommend

[7, 62, 63] and their references for thorough surveys of these three types of tensor decompositions. Note that when the order of tensors is two, then these three tensor decompositions reduce to the singular value decompositions (SVD). However, when the order of the tensors is larger than 2, these three kinds of tensor low-rank approximations have essential differences. The low CP-rank approximation is generally ill-posed [64] for the case of CP-rank larger than 1, contrary to the low Tucker-rank approximation.

The problem of the best rank-one approximation of $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is to find a real scalar $\sigma \in \mathbb{R}$ and N unit vectors $\mathbf{x}_n \in \mathbb{R}^{I_n}$ with all n to minimize

$$\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} (a_{i_1 i_2 \ldots i_N} - \sigma \cdot (x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N}))^2,$$

where x_{n,i_n} is the i_nth element of $\mathbf{x}_n \in \mathbb{R}^{I_n}$ for all i_n and n, and $\sigma \in \mathbb{R}$. Note that the best tensor rank-1 approximation is in fact always well-posed. The Kurash-Kuhn-Tucker (KKT) conditions for the rank-one approximation of real tensors are given below.

Definition 1.3.3 ([29]) Suppose that $\mathbf{x}_n \in \mathbb{R}^{I_n}$ satisfies $\|\mathbf{x}_n\|_2 = 1$ with all n and $\sigma \in \mathbb{R}$. For a given $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, if $(\sigma; \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ solves the following system of nonlinear equations $F(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)_{-n} = \sigma \mathbf{x}_n$ with $F(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)_{-n} = \mathcal{A} \times_1 \mathbf{x}_1^\top \cdots \times_{n-1} \mathbf{x}_{n-1}^\top \times_{n+1} \mathbf{x}_{n+1}^\top \cdots \times_N \mathbf{x}_N^\top$, then σ and the unit vectors $\mathbf{x}_n$ are a singular value of $\mathcal{A}$ and the mode-n singular vector associated to σ, respectively.

Another analogue of the Perron-Frobenius theorem is proved for nonnegative normalized singular pairs in [64], $(\sigma; \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, defined in Definition 1.3.3, is called the normalized singular pair of $\mathcal{A}$. For all n, when the entries of $\mathbf{u}_n$ are nonnegative, $(\sigma; \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ is called the nonnegative normalized singular pair of $\mathcal{A}$. The singular value and singular value inclusion sets for tensors are investigated in [65].

There are several numerical methods for computing a tensor rank-one approximation, such as the alternating least squares (ALS) or higher-order power method (HOPM) [26, 27], the truncated higher-order singular value decomposition, optimization methods based on the product of several Grassmannian manifolds [66–68], semi-definite relaxation methods [69], and sequential rank-one approximation and projection [70]. Recently, Jiang and Kong [71] study the uniqueness of the best rank-one approximation of a tensor under the Frobenius norm. Espig and Khachatryan [72] analyze the convergence of the alternating least squares algorithm. For applications of best rank-one and low multilinear rank approximations, we refer to Cichocki et al. [2], Yang et al. [73], Konakli and Sudret [74], Shah et al. [75], da Silva et al. [76] and the references therein. Applications in machine learning can be found in [77–79].

Many scholars have researched the computation of the symmetric rank-one approximations of real symmetric tensors; see Friedland [80], Kofidis and Regalia

[81], Qi [82], Qi et al. [83], Hu et al. [84], Ni and Wang [85], Wang and Qi [86], and Jiang et al. [87] and the references therein.

For a given $\mathcal{A} \in \mathbb{R}^{I_1 \times I_1 \times I_2 \times I_2}$, if the entries of $\mathcal{A}$ satisfy the symmetric property: $a_{i_1 i_2 i_3 i_4} = a_{i_2 i_1 i_3 i_4} = a_{i_1 i_2 i_4 i_3}$ with $i_1, i_2 = 1, 2, \ldots, I_1$ and $i_3, i_4 = 1, 2, \ldots, I_2$, then we call $\mathcal{A}$ partially symmetric.

The fourth-order partially symmetry have received much attention [45, 88–93]. Zhang et al. [94] prove that the best rank-one approximation of a symmetric tensor is its best symmetric rank-one approximation. Similarly, we can prove that the best rank-one approximation of a fourth order partially symmetric tensor is its best partially symmetric rank-one approximation. The rank-one approximation problem of a partially symmetric tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_1 \times I_2 \times I_2}$ is to find a real scalar $\sigma \in \mathbb{R}$ and two unit vectors $\mathbf{x}_n \in \mathbb{R}^{I_n}$ with $n = 1, 2$ to minimize

$$\sum_{i_1, i_2 = 1}^{I_1} \sum_{i_3, i_4 = 1}^{I_2} (a_{i_1 i_2 i_3 i_4} - \sigma \cdot (x_{1,i_1} x_{1,i_2} x_{2,i_3} x_{2,i_4}))^2. \tag{1.3.2}$$

The minimization problem (1.3.2) is equivalent to finding two unit vectors $\mathbf{x}_n \in \mathbb{R}^{I_n}$ ($\|\mathbf{x}_n\|_2 = 1; n = 1, 2$) to maximize

$$\max |\mathcal{A}\mathbf{x}_1^2 \mathbf{x}_2^2|, \tag{1.3.3}$$

with $\mathcal{A}\mathbf{x}_1^2\mathbf{x}_2^2 := \mathcal{A} \times_1 \mathbf{x}_1^\top \times_2 \mathbf{x}_1^\top \times_3 \mathbf{x}_2^\top \times_4 \mathbf{x}_2^\top$. The biquadratic optimization problems arise from the strong ellipticity condition problem in solid mechanics [95–98] and the entanglement problem in quantum physics [88, 99]. Before considering the KKT conditions for the minimization problem (1.3.2), we introduce two notations:

$$\mathcal{A}\mathbf{x}_1^2\mathbf{x}_2 := \mathcal{A} \times_1 \mathbf{x}_1^\top \times_2 \mathbf{x}_1^\top \times_4 \mathbf{x}_2^\top, \ \mathcal{A}\mathbf{x}_1\mathbf{x}_2^2 := \mathcal{A} \times_2 \mathbf{x}_1^\top \times_3 \mathbf{x}_2^\top \times_4 \mathbf{x}_2^\top.$$

For any maximizer $(\mathbf{x}_1, \mathbf{x}_2)$ of (1.3.3), by the optimality theory [100], there exist $\lambda, \sigma \in \mathbb{R}$ such that

$$\mathcal{A}\mathbf{x}_1^2\mathbf{x}_2 = \sigma \mathbf{x}_2, \ \mathcal{A}\mathbf{x}_1\mathbf{x}_2^2 = \lambda \mathbf{x}_1, \ \|\mathbf{x}_1\|_2 = \|\mathbf{x}_2\|_2 = 1. \tag{1.3.4}$$

The optimal conditions can further be simplified with (1.3.4) and $\lambda = \sigma$. If σ, $\mathbf{x}_1$ and $\mathbf{x}_2$ are real solutions of (1.3.4), then σ is said to be an M-eigenvalue of $\mathcal{A}$, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are said to be the first and the second (left and right in [45]) M-eigenvector of $\mathcal{A}$, associated with σ, respectively.

There are several numerical methods for solving the minimization problem (1.3.2), such as ALS or HOPM [92] and semi-definite programming (SDP) relaxations [93].

1.3.3 Complex Tensor Rank-One Approximations

Entanglement has been identified as a resource central to quantum information processing and we are motivated to investigate its quantification in the bipartite and multipartite pure states [101]. A useful tool for quantifying the amount of entanglement of a state is given by the so-called *entanglement measures*. The geometric measure of entanglement is one of most natural entanglement measures for pure states in bipartite and multipartite systems [102, 103]. This measure is the injective tensor norm [104], which appears in the theory of operator algebra [105]. It also has applications in many-body physics [106, 107], entanglement witnesses [108, 109] and the study of quantum channel capacities [110–112].

The best rank-one approximation of complex tensors is the core problem for computing the geometric measure for pure states. For $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, we need to find a real scalar $\sigma \in \mathbb{R}$ and N unitary vectors $\mathbf{x}_n \in \mathbb{C}^{I_n}$ with all n to minimize

$$\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \left| a_{i_1 i_2 \ldots i_N} - \sigma \cdot (x_{1,i_1} x_{2,i_2} \ldots x_{N,i_N}) \right|^2,$$

where x_{n,i_n} is the i_nth element of $\mathbf{x}_n \in \mathbb{C}^{I_n}$ for all i_n and n, and $\sigma \in \mathbb{R}$. By means of the Wirtinger calculus of complex functions [113–116], the corresponding KKT conditions are given below.

Definition 1.3.4 For a given $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$, let $\mathbf{v}_n \in \mathbb{C}^{I_n}$ be nonzero vectors with $\|\mathbf{v}_n\|_2 = 1$ $(n = 1, 2, \ldots, N)$ and let $\sigma \in \mathbb{R}$. If $(\sigma; \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$ solves the following system of nonlinear equations

$$F(\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_N)_{-n} = \sigma \mathbf{v}_n, \quad \overline{F}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)_{-n} = \sigma \overline{\mathbf{v}}_n,$$

where

$$\begin{aligned} F(\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_N)_{-n} &= \mathcal{A} \times_1 \overline{\mathbf{v}}_1^\top \cdots \times_{n-1} \overline{\mathbf{v}}_{n-1}^\top \times_{n+1} \overline{\mathbf{v}}_{n+1}^\top \cdots \times_N \overline{\mathbf{v}}_N^\top \\ &= \mathcal{A} \times_1 \mathbf{v}_1^* \cdots \times_{n-1} \mathbf{v}_{n-1}^* \times_{n+1} \mathbf{v}_{n+1}^* \cdots \times_N \mathbf{v}_N^*, \\ \overline{F}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)_{-n} &= \overline{\mathcal{A}} \times_1 \mathbf{v}_1^\top \cdots \times_{n-1} \mathbf{v}_{n-1}^\top \times_{n+1} \mathbf{v}_{n+1}^\top \cdots \times_N \mathbf{v}_N^\top, \end{aligned}$$

then $\mathbf{v}_n$ and σ are called the mode-n unitary eigenvector (the mode-n U-eigenvector) and unitary eigenvalue (U-eigenvalue) of $\mathcal{A}$, respectively. We call $(\sigma; \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$ a U-eigenpair of $\mathcal{A}$.

In particular, if $\mathcal{A}$ is complex symmetric, then all $\mathbf{v}_n$ are identical, denoted by $\mathbf{v}$, and $(\sigma; \mathbf{v})$ is a US-eigenpair of $\mathcal{A}$. Hu et al. [117] consider how to use the spectral theory of nonnegative tensors for computing the geometric measure of entanglement in multipartite pure states. Ni et al. [28] define the concept of the U-eigenvalue of a complex tensor, the US-eigenvalue of a complex symmetric tensor and the best complex rank-one approximation. They also derive an upper bound on the number of distinct US-eigenvalues of a complex symmetric tensor. Based on the theory of the spherical optimization problem with complex variables, Ni and Bai [118] design an iterative algorithm for computing the US-eigenpairs of complex symmetric tensors. Che et al. [119] present the complex-valued neural networks for solving the quantum eigenvalue problem for multipartite pure states. Wang et al. [120] derive the partial orthogonal rank-one decomposition of complex symmetric tensors based on the Takagi factorization.

References

1. J. Cardoso, High-order contrasts for independent component analysis. Neural Comput. **11**(1), 157–192 (1999)
2. A. Cichocki, N. Lee, I. Oseledets, A. Phan, Q. Zhao, D. Mandic, Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions. Found. Trends Mach. Learn. **9**(4-5), 249–429 (2016)
3. P. Comon, Independent component analysis, a new concept? Signal Process. **36**(3), 287–314 (1994)
4. P. Comon, Tensor decompositions: state of the art and applications, in *Mathematics in Signal Processing, V (Coventry, 2000)*. The Institute of Mathematics and its Applications Conference Series, vol. 71 (Oxford University, Oxford, 2002), pp. 1–24
5. C. Nikias, J. Mendel, Signal processing with higher-order spectra. IEEE Signal Process. Mag. **10**(3), 10–37 (1993)
6. P. Comon, G. Golub, L. Lim, B. Mourrain, Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. **30**(3), 1254–1279 (2008)
7. T. Kolda, B. Bader, Tensor decompositions and applications. SIAM Rev. **51**(3), 455–500 (2009)
8. A. Cichocki, R. Zdunek, A. Phan, S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation* (Wiley, New York, 2009)
9. Y. Yang, Q. Yang, *A Study on Eigenvalues of Higher-Order Tensors and Related Polynomial Optimization Problems* (Science, Beijing, 2015)
10. L. Qi, Z. Luo, *Tensor Analysis: Spectral Theory and Special Tensors* (Society of Industrial and Applied Mathematics, Philadelphia, 2017)
11. Y. Wei, W. Ding, *Theory and Computation of Tensors: Multi-Dimensional Arrays* (Academic, Amsterdam, 2016)
12. L. Qi, H. Chen, Y. Chen, *Tensor Eigenvalues and their Applications.*, vol. 39 (Springer: Singapore, 2018)
13. F. Chung, *Spectral Graph Theory* (American Mathematical Society, New York, 1997)
14. C. Berge, Graphs and Hypergraphs, in *North-Holland Mathematical Library*, vol. 45 (North-Holland, Amsterdam, 1976)
15. S. Bulo, M. Pelillo, New bounds on the clique number of graphs based on spectral hypergraph theory, in *Learning and Intelligent Optimization* (Springer, New York, 2009), pp. 45–58
16. J. Cooper, A. Dutle, Spectra of uniform hypergraphs. Linear Algebra Appl. **436**(9), 3268–3292 (2012)
17. K. Pearson, T. Zhang, On spectral hypergraph theory of the adjacency tensor. Graphs and Combinatorics **30**(5), 1233–1248 (2014)
18. J. Xie, A. Chang, On the Z-eigenvalues of the adjacency tensors for uniform hypergraphs. Linear Algebra Appl. **439**(8), 2195–2204 (2013)

19. S. Hu, L. Qi, J. Shao, Cored hypergraphs, power hypergraphs and their Laplacian H-eigenvalues. Linear Algebra Appl. **439**(10), 2980–2998 (2013)
20. S. Hu, L. Qi, J. Xie, The largest Laplacian and signless Laplacian H-eigenvalues of a uniform hypergraph. Linear Algebra Appl. **469**, 1–27 (2015)
21. V. Nikiforov, Analytic methods for uniform hypergraphs. Linear Algebra Appl. **457**, 455–535 (2014)
22. L. Qi, J. Shao, Q. Wang, Regular uniform hypergraphs, s-cycles, s-paths and their largest Laplacian H-eigenvalues. Linear Algebra Appl. **443**, 215–227 (2014)
23. H. Lu, K. Plataniotis, A. Venetsanopoulos, *Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data* (CRC, Boca Raton, 2013)
24. L. Qi, Eigenvalues of a real supersymmetric tensor. J. Symb. Comput. **40**(6), 1302–1324 (2005)
25. L. De Lathauwer, B. De Moor, J. Vandewalle, A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. **21**(4), 1253–1278 (2000)
26. L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-(r_1, r_2, ..., r_n) approximation of higher-order tensors. SIAM J. Matrix Anal. Appl. **21**(4), 1324–1342 (2000)
27. T. Zhang, G. Golub, Rank-one approximation to high order tensors. SIAM J. Matrix Anal. Appl. **23**(2), 534–550 (2001)
28. G. Ni, L. Qi, M. Bai, Geometric measure of entanglement and U-eigenvalues of tensors. SIAM J. Matrix Anal. Appl. **35**(1), 73–87 (2014)
29. L. Lim, Singular values and eigenvalues of tensors: a variational approach, in *IEEE CAMSAP 2005: First International Workshop on Computational Advances in Multi-Sensor Adaptive Processing* (IEEE, Piscataway, 2005), pp. 129–132
30. L. Qi, Symmetric nonnegative tensors and copositive tensors. Linear Algebra Appl. **439**(1), 228–238 (2013)
31. Z. Chen, L. Qi, Circulant tensors with applications to spectral hypergraph theory and stochastic process. J. Ind. Manag. Optim. **12**(4), 1227–1247 (2013)
32. S. Dow, P. Gibson, Permanents of d-dimensional matrices. Linear Algebra Appl. **90**, 133–145 (1987)
33. K. Chang, K. Pearson, T. Zhang, Perron-Frobenius theorem for nonnegative tensors. Commun. Math. Sci. **6**(2), 507–520 (2008)
34. A. Berman, R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* (Society for Industrial and Applied Mathematics, Philadelphia, 1994)
35. S. Hu, G. Li, L. Qi, Y. Song, Finding the maximum eigenvalue of essentially nonnegative symmetric tensors via sum of squares programming. J. Optim. Theory Appl. **158**(3), 713–738 (2013)
36. G. Li, L. Qi, G. Yu, The Z-eigenvalues of a symmetric tensor and its application to spectral hypergraph theory. Numer. Linear Algebra Appl. **20**(6), 1001–1029 (2013)
37. M. Ng, L. Qi, G. Zhou, Finding the largest eigenvalue of a nonnegative tensor. SIAM J. Matrix Anal. Appl. **31**(3), 1090–1099 (2009)
38. L. Qi, G. Yu, E. Wu, Higher order positive semi-definite diffusion tensor imaging. SIAM J. Imag. Sci. **3**(3), 416–433 (2010)
39. F. Bohnetwaldraff, D. Braun, O. Giraud, Tensor eigenvalues and entanglement of symmetric states. Phys. Rev. A **94**(4), 042324 (2016)
40. G. Gaeta, E. Virga, Octupolar order in three dimensions. Eur. Phys. J. E **39**(11), 113 (2016)
41. K. Kannike, Vacuum stability of a general scalar potential of a few fields. Eur. Phys. J. C **76**(6), 324 (2016)
42. G. Golub, C. Van Loan, *Matrix Computations*, 4th edn. (Johns Hopkins University Press, Baltimore, 2013)
43. T. Kolda, J. Mayo, Shifted power method for computing tensor eigenpairs. SIAM J. Matrix Anal. Appl. **32**(4), 1095–1124 (2011)
44. L. Qi, Y. Wang, E. Wu, D-eigenvalues of diffusion kurtosis tensors. J. Comput. Appl. Math. **221**(1), 150–157 (2008)

45. L. Qi, H. Dai, D. Han, Conditions for strong ellipticity and M-eigenvalues. Front. Math. China **4**(2), 349–364 (2009)
46. K. Chang, K. Pearson, T. Zhang, On eigenvalue problems of real symmetric tensors. J. Math. Anal. Appl. **350**(1), 416–422 (2009)
47. C. Cui, Y. Dai, J. Nie, All real eigenvalues of symmetric tensors. SIAM J. Matrix Anal. Appl. **35**(4), 1582–1601 (2014)
48. T. Kolda, J. Mayo, An adaptive shifted power method for computing generalized tensor eigenpairs. SIAM J. Matrix Anal. Appl. **35**(4), 1563–1581 (2014)
49. W. Ding, Y. Wei, Generalized tensor eigenvalue problems. SIAM J. Matrix Anal. Appl. **36**(3), 1073–1099 (2015)
50. L. Chen, L. Han, L. Zhou, Computing tensor eigenvalues via homotopy methods. SIAM J. Matrix Anal. Appl. **37**(1), 290–319 (2016)
51. J. Chang, Y. Chen, L. Qi, Computing eigenvalues of large scale sparse tensors arising from a hypergraph. SIAM J. Sci. Comput. **38**(6), A3618–A3643 (2016)
52. K. Batselier, N. Wong, *A QR Algorithm for Symmetric Tensors* (2014). ArXiv preprint:1411.1926v1
53. Q. Yang, Y. Yang, Further results for Perron-Frobenius theorem for nonnegative tensors II. SIAM J. Matrix Anal. Appl. **32**(4), 1236–1250 (2011)
54. Y. Yang, Q. Yang, Further results for Perron-Frobenius theorem for nonnegative tensors. SIAM J. Matrix Anal. Appl. **31**(5), 2517–2530 (2010)
55. L. Zhang, L. Qi, Linear convergence of an algorithm for computing the largest eigenvalue of a nonnegative tensor. Numer. Linear Algebra Appl. **19**(5), 830–841 (2012)
56. C. Liu, C. Guo, W. Lin, A positivity preserving inverse iteration for finding the perron pair of an irreducible nonnegative third order tensor. SIAM J. Matrix Anal. Appl. **37**(3), 911–932 (2016)
57. C. Liu, C. Guo, W. Lin, Newton-Noda iteration for finding the Perron pair of a weakly irreducible nonnegative tensor. Numer. Math. **137**(1), 63–90 (2017)
58. K. Chang, L. Qi, T. Zhang, A survey on the spectral theory of nonnegative tensors. Numer. Linear Algebra Appl. **20**(6), 891–912 (2013)
59. W. Li, M. Ng, The perturbation bound for the spectral radius of a nonnegative tensor. Adv. Numer. Anal., Article ID 109525, 10 (2014)
60. W. Li, M. Ng, Some bounds for the spectral radius of nonnegative tensors. Numer. Math. **130**(2), 315–335 (2015)
61. B. Chen, S. He, Z. Li, S. Zhang, On new classes of nonnegative symmetric tensors. SIAM J. Optim. **27**, 292–318 (2017)
62. A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, H. Phan, Tensor decompositions for signal processing applications: From two-way to multiway component analysis. IEEE Signal Process. Mag. **32**(2), 145–163 (2015)
63. L. Grasedyck, D. Kressner, C. Tobler, A literature survey of low-rank tensor approximation techniques. GAMM-Mitt. **36**(1), 53–78 (2013)
64. Y. Qi, P. Comon, L. Lim, Uniqueness of nonnegative tensor approximations. IEEE Trans. Inf. Theory **62**(4), 2170–2183 (2016)
65. Y. Miao, C. Li, Y. Wei, Z-singular value and Z-singular value inclusion sets for tensors. Jpn. J. Ind. Appl. Math. **36**, 1055–1087 (2019)
66. L. Eldén, B. Savas, A Newton-Grassmann method for computing the best multilinear rank-(r_1, r_2, r_3) approximation of a tensor. SIAM J. Matrix Anal. Appl. **31**(2), 248–271 (2009)
67. M. Ishteva, P. Absil, S. Van Huffel, L. De Lathauwer, Best low multilinear rank approximation of higher-order tensors, based on the Riemannian trust-region scheme. SIAM J. Matrix Anal. Appl. **32**(1), 115–135 (2011)
68. B. Savas, L. Lim, Quasi-Newton methods on Grassmannians and multilinear approximations of tensors. SIAM J. Sci. Comput. **32**(6), 3352–3393 (2010)
69. J. Nie, L. Wang, Semidefinite relaxations for best rank-1 tensor approximations. SIAM J. Matrix Anal. Appl. **35**(3), 1155–1179 (2014)