Eigenvalues of Tensors and Their Applications

by

Liqun Qi

Department of Mathematics
City University of Hong Kong

Outline

- ***Positive Definiteness of Multivariate Forms***
- ***Eigenvalues of Tensors***
- ***Z-Eigenvalue Methods and The Best Rank-One Approximation***
- ***More Study on E-Eigenvalues and Z-eigenvalues***
- ***Lim's Exploration***
- ***Further Discussion***

1. Positive Definiteness of Multivariate Forms

Recently, I defined eigenvalues and eigenvectors of a real supersymmetric tensor, and explored their practical applications in determining positive definiteness of an even degree multivariate form, and finding the best rank-one approximation to a supersymmetric tensor. This work extended the classical concept of eigenvalues of square matrices, and has potential applications in mechanics and physics as well as the classification of hypersurfaces and the study of hypergraphs.

1.1. Quadratic Forms

Consider a quadratic form

$$f(x) \equiv x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where $x = (x_1, \cdots, x_n)^T \in \Re^n$, $A = (a_{ij})$ is a symmetric matrix. If $f(x) > 0$ for all $x \in \Re^n, x \neq 0$, then f and A are called positive definite. A complex number λ is called an eigenvalue of the matrix A if there is a nonzero complex vector x such that

$$Ax = \lambda x.$$

A complex number is an eigenvalue of A if it is a root of the characteristic polynomial of A. An $n \times n$ real symmetric matrix has n real eigenvalues. It is positive definite if and only if all of its eigenvalues are positive. These are some basic knowledge of linear algebra.

Can these be extended to higher orders? Is there a need from practice for such an extension?

1.2. Homogeneous Polynomial Form

An n-dimensional homogeneous polynomial form of degree m, $f(x)$, where $x \in \Re^n$, is equivalent to the tensor product of a **supersymmetric** n-dimensional tensor A of order m, and the rank-one tensor x^m:

$$f(x) \equiv Ax^m := \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1,\cdots,i_m} x_{i_1} \cdots x_{i_m}.$$

The tensor A is called supersymmetric as its entries $a_{i_1,\cdots,i_m}$ are invariant under any permutation of their indices. The tensor A is called positive definite (semidefinite) if $f(x) > 0$ ($f(x) \geq 0$) for all $x \in \Re^n, x \neq 0$. When m is even, the positive definiteness of such a homogeneous polynomial form $f(x)$ plays an important role in the stability study of nonlinear autonomous systems via Liapunov's direct method in automatic control. For $n \geq 3$ and $m \geq 4$, this issue is a hard problem in mathematics.

1.3. Study on the Positive Definiteness

[1]. B.D. Anderson, N.K. Bose and E.I. Jury, "Output feedback stabilization and related problems-solutions via decision methods", *IEEE Trans. Automat. Contr.* **AC20** (1975) 55-66.

[2]. N.K. Bose and P.S. Kamt, "Algorithm for stability test of multidimensional filters", *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-22** (1974) 307-314.

[3]. N.K. Bose and A.R. Modaress, "General procedure for multivariable polynomial positivity with control applications", *IEEE Trans. Automat. Contr.* **AC21** (1976) 596-601.

[4]. N.K. Bose and R.W. Newcomb, "Tellegon's theorem and multivariate realizability theory", *Int. J. Electron.* **36** (1974) 417-425.

[5]. M. Fu, "Comments on 'A procedure for the positive definiteness of forms of even-order' ", *IEEE Trans. Autom. Contr.* **43** (1998) 1430.

[6]. M.A. Hasan and A.A. Hasan, “A procedure for the positive definiteness of forms of even-order”, *IEEE Trans. Autom. Contr.* **41** (1996) 615-617.

[7]. J.C. Hsu and A.U. Meyer, Modern Control Principles and Applications, McGraw-Hill, New York, 1968.

[8]. E.I. Jury and M. Mansour, “Positivity and nonnegativity conditions of a quartic equation and related problems” *IEEE Trans. Automat. Contr.* **AC26** (1981) 444-451.

[9]. W.H. Ku, “Explicit criterion for the positive definiteness of a general quartic form”, *IEEE Trans. Autom. Contr.* **10** (1965) 372-373.

[10]. F. Wang and L. Qi, “Comments on ‘Explicit criterion for the positive definiteness of a general quartic form’ ”, *IEEE Trans. Autom. Contr.* **50** (2005) 416- 418.

2. Eigenvalues of Tensors

L. Qi, "Eigenvalues of a real supersymmetric tensor", *Journal of Symbolic Computation* **40** (2005) 1302-1324,

defined eigenvalues and eigenvectors of a real supersymmetric tensor, and explored their practical applications in determining positive definiteness of an even degree multivariate form.
By the tensor product, Ax^{m-1} for a vector $x \in \Re^n$ denotes a vector in $\Re^n$, whose ith component is

$$\left(Ax^{m-1}\right)_i \equiv \sum_{i_2,\cdots,i_m=1}^{n} a_{i,i_2,\cdots,i_m} x_{i_2} \cdots x_{i_m}.$$

We call a number $\lambda \in C$ an **eigenvalue** of A if it and a nonzero vector $x \in C^n$ are solutions of the following homogeneous polynomial equation:

$$\left(Ax^{m-1}\right)_i = \lambda x_i^{m-1}, \ \ \forall\, i = 1, \cdots, n. \tag{1}$$

and call the solution x an **eigenvector** of A associated with the eigenvalue λ. We call an eigenvalue of A an **H-eigenvalue** of A if it has a real eigenvector x. An eigenvalue which is not an H-eigenvalue is called an **N-eigenvalue**. A real eigenvector associated with an H-eigenvalue is called an **H-eigenvector**.

The **resultant** of (1) is a one-dimensional polynomial of λ. We call it the **characteristic polynomial** of A.

2.1. Theorem on Eigenvalues

Theorem 2.1 *(***Qi 2005***)*
We have the following conclusions on eigenvalues of an mth order n-dimensional supersymmetric tensor A:
(a). A number $\lambda \in C$ is an eigenvalue of A if and only if it is a root of the characteristic polynomial ϕ.
(b). The number of eigenvalues of A is $d = n(m-1)^{n-1}$. Their product is equal to det(A), the resultant of $Ax^{m-1} = 0$.
(c). The sum of all the eigenvalues of A is

$$(m-1)^{n-1}\mathrm{tr}(A),$$

where tr(A) denotes the sum of the diagonal elements of A.
(d). If m is even, then A always has H-eigenvalues. A is positive definite (positive semidefinite) if and only if all of its H-eigenvalues are positive (nonnegative).
(e). The eigenvalues of A lie in the following n disks:

$$|\lambda - a_{i,i,\cdots,i}| \le \sum \{|a_{i,i_2,\cdots,i_m}| : i_2, \cdots, i_m = 1, \cdots, n, \{i_2, \cdots, i_m\} \neq \{i, \cdots, i\}\},$$

for $i = 1, \cdots, n$.

2.2. E-Eigenvalues

In the same paper, I defined another kind of eigenvalues for tensors. Their structure is different from the structure described by Theorem 2.1. Their characteristic polynomial has a lower degree.

Suppose that A is an mth order n-dimensional supersymmetric tensor. We say a complex number λ is an **E-eigenvalue** of A if there exists a complex vector x such that

$$\begin{cases} Ax^{m-1} = \lambda x, \\ x^T x = 1. \end{cases} \tag{2}$$

In this case, we say that x is an E-eigenvector of the tensor A associated with the E-eigenvalue λ. If an E-eigenvalue has a real E-eigenvector, then we call it a **Z-eigenvalue** and call the real E-eigenvector a **Z-eigenvector**.

2.3. The E-Characteristic Polynomial and Orthogonal Similarity

When m is even, the **resultant** of

$$Ax^{m-1} - \lambda(x^T x)^{\frac{m-2}{2}} x = 0$$

is a one dimensional polynomial of λ and is called the **E-characteristic polynomial** of A. We say that A is regular if the following system has no nonzero complex solutions:

$$\begin{cases} Ax^{m-1} = 0, \\ x^T x = 0. \end{cases}$$

Let $P = (p_{ij})$ be an $n \times n$ real matrix. Define $B = P^m A$ as another mth order n-dimensional tensor with entries

$$b_{i_1,i_2,\cdots,i_m} = \sum_{j_1,j_2,\cdots,j_m=1}^{n} p_{i_1 j_1} p_{i_2 j_2} \cdots p_{i_m j_m} a_{j_1,j_2,\cdots,j_m}.$$

If P is an orthogonal matrix, then we say that A and B are **orthogonally similar**.

2.4. Theorem on E-eigenvalues

Theorem 2.2 (**Qi 2005**)
We have the following conclusions on E-eigenvalues of an mth order n-dimensional supersymmetric tensor A:
(a). When A is regular, a complex number is an E-eigenvalue of A if and only if it is a root of its E-characteristic polynomial.
(b). Z-eigenvalues always exist. An even order supersymmetric tensor is positive definite if and only if all of its Z-eigenvalues are positive.
(c). If A and B are orthogonally similar, then they have the same E-eigenvalues and Z-eigenvalues.
(d). If λ is the Z-eigenvalue of A with the largest absolute value and x is a Z-eigenvector associated with it, then λx^m is the best rank-one approximation of A, i.e.,

$$\|A - \lambda x^m\|_F = \sqrt{\|A\|_F^2 - \lambda^2} = \min\{\|A - \alpha u^m\|_F : \alpha \in \Re, u \in \Re^n, \|u\|_2 = 1\},$$

where $\|\cdot\|_F$ is the Frobenius norm.

3. Z-eigenvalue Methods and Best Rank-One Approximation

If we may find the smallest Z-eigenvalue and the Z-eigenvector, then we have a way to solve the best rank-one approximation problem and the positive definiteness problem. Z-eigenvalue methods have been developed in the following paper:

L. Qi, F. Wang and Y. Wang, “Z-eigenvalue methods for a global polynomial optimization problem”, June 2006.

3.1. The Best Rank-One Approximation

The best rank-one approximation of higher order tensors has extensive engineering and statistical applications. The following are some papers on this topic:

[1]. L. De Lathauwer, B. De Moor and J. Vandwalle, "On the best rank-1 and rank-$(R_1, R_2, \cdots, R_n)$ approximation of higher order tensors", *SIAM J. Matrix Anal. Appl*, **21** (2000) 1324-1342.

[2]. E. Kofidies and Ph.A. Regalia, "On the best rank-1 approximation of higher order supersymmetric tensors", *SIAM J. Matrix Anal. Appl*, **23** (2002) 863-884.

[3]. T. Zhang and G.H. Golub, "Rank-1 approximation to higher order tensors", *SIAM J. Matrix Anal. Appl*, **23** (2001) 534-550.

3.2. The Pseudo-Canonical Form

The problem for finding the smallest Z-eigenvalue of a supersymmetric tensor A is equivalent to solving the following global polynomial optimization problem

$$\begin{aligned} \min f(x) &= Ax^m \\ \text{subject to } \; x^T x &= 1. \end{aligned} \tag{3}$$

When $n = 2$, we may use definition (2) to find all the Z-eigenvalues. When $n = 3$, we may use a pseudo-canonical form of A for help to find all the Z-eigenvalues.

An m-th order n-dimensional supersymmetric tensor B is said to be a pseudo-canonical form of another mth order n-dimensional supersymmetric tensor A if A and B are orthogonally similar and

$$b_{i,\cdots,i,j} = 0 \tag{4}$$

for all $1 \leq i < j \leq n$. In this case, we say that B is a pseudo-canonical form. If $m = 2$, then a pseudo-canonical form is a diagonal matrix.

3.3. Find A Pseudo-Canonical Form

We may use a conventional optimization method to find a local minimizer of (3). Then we find a Z-eigenvalue and a Z-eigenvector of A. Suppose that λ is a Z-eigenvalue of A with a Z-eigenvector x. Let P be an orthogonal matrix with x as its first row vector. Let $B = P^m A$. Then we see that $y = Px = (1, 0, \cdots, 0)^T$. By (2), we see that

$$b_{1,\cdots,1} = \lambda$$

and

$$b_{1,\cdots,1,i} = 0, \tag{5}$$

for $i = 2, \cdots, n$. Repeating this procedure to some principal subtensors of A, we may find a pseudo-canonical form of A.

3.4. The Third Order Pseudo-Canonical Form

When $n \geq 4$, we developed two heuristic Z-eigenvalue methods for finding an approximate global minimizer of (3).
The first heuristic Z-eigenvalue method is based upon finding a third order pseudo-canonical form of A.

An mth order n-dimensional supersymmetric tensor B is called a first order pseudo-canonical form of another mth order n-dimensional supersymmetric tensor A if it is a pseudo-canonical form of A and $b_{i,\cdots,i} \leq b_{j,\cdots,j}$ for $1 \leq i < j \leq n$.

Let $1 \leq j < k < l \leq n$. We use $B(j,k,l)$ to denote the m-th order three dimensional supersymmetric tensor whose entries are $b_{i_1,i_2,\cdots,i_m}$, where $i_1, i_2, \cdots, i_m = j,k,l$, and use $[B(j,k,l)]_{\min}$ to denote the smallest Z-eigenvalue of $B(j,k,l)$. An mth order n-dimensional supersymmetric tensor B is called a third order pseudo-canonical form of another mth order n-dimensional supersymmetric tensor A if it is a first order pseudo-canonical form of A and

$$b_{1,\cdots,1} = \min_{1 \leq j<k<l \leq n} [B(j,k,l)]_{\min}. \tag{6}$$

3.5. A Heuristic Cross-Hill Z-Eigenvalue Method

We may describe this method for $m = 3$. Certainly, the global minimum of (3) is negative unless A is a zero tensor. Suppose that we have found a local minimizer $y = y^{(1)}$ of (3) with a negative objective function value. We may find $n-1$ unit vectors $y^{(i)}$ of $\Re^n$ for $i = 2, \cdots, n$ such that $\{y^{(i)} : i = 1, \cdots, n\}$ constitute an orthogonal basis of $\Re^n$. Now, for $i = 2, \cdots, n$, restrict problem (3) on the plane spanned by y and $y^{(i)}$. Certainly, this is a two-dimensional form of problem (3) and the current local minimizer y of (3) is also a local minimizer of the restricted problem. This two dimensional problem has at most two local minimizers with negative objective function values. Then we may use the low dimensional method mentioned earlier to find $z^{(i)}$, the other local minimizer of the two-dimensional problem with a negative objective function value for $i = 2, \cdots, n$. Since $z^{(i)}$ may not be a local minimizer of (3), and y and $z^{(i)}$ are separated by a "hill" of the objective function value in two dimensional case, if we use a conventional descent optimization method with $z^{(i)}$ as the starting point, we will find a local minimizer $w^{(i)}$ of (3), which has a negative objective function value and is different from y. We may continue this process until no new local minimizers can be found. Comparing the objective function values of these local minimizers of (3), we may have a better solution.

3.6. Numerical Results

We did numerical results for $m = 3$. They show that our heuristic Z-eigenvalue methods are efficient and promising.

To construct the testing problems in our numerical experiments, we take two sets of three order supersymmetric tensors of different dimensions. The elements of tensors in one set are generated by a uniform distribution in the interval (-1,1); while the elements of tensors in another set are generated by a standard normal distribution. In the following, these two sets of testing problems are labeled as TPI and TPII, respectively.
To analyze the numerical results of our algorithms, we compute the global minimizers of the testing problems by the uniform grid method combining with the descent solution method.

Tables 1 and 2 show the performance of Algorithm 1 (The Third-Order Pseudo-Canonical Form Z-eigenvalue Method) and Algorithm 2 (The Cross-Hill Z-eigenvalue Method) for problems TPI and TPII, where **Dim** denotes the dimension of the tensor A, **Num** denotes the number of tests for each dimension, **Alg1** and **Alg2** denote Algorithms 1 and 2 respectively, **RS** denotes the success ratio of finding the global minimizers, **AT** denotes the average time for each sample, and **AN** denotes the average number of local minimizers found by Algorithm 2 for each sample.
Numerical results show that Algorithm 2 has higher success ratios, while Algorithm 1 uses less computational time.

Table 1: **Results of TPI**

Dim	Num	Alg1		Alg2		
		RS(%)	AT(s)	RS(%)	AT(s)	AN
3	1000	100.0	0.013	99.4	0.038	2.9
4	1000	96.6	0.031	99.5	0.103	3.6
6	1000	91.6	0.089	99.7	0.335	4.6
8	100	88	0.196	99.0	0.728	5.6
10	100	89	0.379	99.0	1.359	6.4

Table 2: **Results of TPII**

Dim	Num	Alg1		Alg2		
		RS(%)	AT(s)	RS(%)	AT(s)	AN
3	1000	100.0	0.014	99.2	0.042	2.9
4	1000	96.5	0.031	99.3	0.097	3.6
6	1000	92.2	0.085	99.8	0.265	4.5
8	100	88	0.178	99.0	0.521	5.3
10	100	87	0.331	99.0	0.933	6.2

4. More Study on E-eigenvalues and Z-eigenvalues

Tensors are practical physical quantities in relativity theory, fluid dynamics, solid mechanics and electromagnetism, etc. The concept of tensors was introduced by Gauss, Riemann and Christoffel, etc., in the 19th century in the study of differential geometry. In the very beginning of the 20th century, Ricci, Levi-Civita, etc., further developed tensor analysis as a mathematical discipline. But it was **Einstein** who applied tensor analysis in his study of general relativity in 1916. This made tensor analysis an important tool in theoretical physics, continuum mechanics and many other areas of science and engineering.

The tensors in theoretical physics and continuum mechanics are physical quantities which are invariant under co-ordinate system changes.

4.1. Tensors in Physics and Mechanics

In the study of tensors in physics and mechanics, an important concept is **invariant**. A scalar associated with a tensor is an invariant of that tensor. It does not change under co-ordinate system changes.

A first order tensor, i.e., a vector, has one invariant, which is its magnitude. A second order n-dimensional tensor has n eigenvalues, which are roots of its characteristic polynomial. Then, the coefficients of the characteristic polynomial are principal invariants of that tensor. They have important physical meanings.

L. Qi, "Eigenvalues and invariants of tensors", to appear in: *Journal of Mathematical Analysis and Applications*,

considered E-eigenvalues of general (not necessarily symmetric) tensors and showed that E-eigenvalues and the coefficients of the E-characteristic polynomial are invariants. Thus, they may have potential applications in mechanics and physics, and deserve further exploration.

4.2. The Classification of Hypersurfaces

The classification of algebraic hypersurfaces is important and hard. **Newton's classification of cubic curves** in the late 1600s was the first great success of analytic geometry apart from its role in calculus.

L. Qi, "Rank and eigenvalues of a supersymmetric tensor, the multivariate homogeneous polynomial and the algebraic hypersurface it defines", to appear in: *Journal of Symbolic Computation*,

deals with the orthogonal classification problem of real hypersurfaces given by an equation of the form

$$S = \{x \in \Re^n : f(x) = Ax^m = c\},$$

where A is a supersymmetric tensor, through the rank, the Z-eigenvalues and the asymptotic directions.

4.3. The Rank of a Supersymmetric Tensor

Let A be an mth order n-dimensional real supersymmetric tensor. We call a real n-dimensional vector x a **recession vector** of A if

$$Ax = 0.$$

All the recession vectors of A form a linear subspace of $\Re^n$. We call it the **recession space** of A and denote it by V_R. We call

$$r = n - \dim V_R$$

the **rank** of A. Clearly, $0 \leq r \leq n$. If $r < n$, then we may use a linear transformation $y = Lx$ to transform $f(x) = Ax^m$ to $g(y) = f(Lx)$ such that $y \in \Re^r$ and g is an r-dimensional homogeneous polynomial form of degree m.

5. Lim's Exploration

Independently, Lek-Heng Lim also defined eigenvalues for tensors in his paper:

L-H. Lim, "Singular values and eigenvalues of tensors: A variational approach", Proceedings of the 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), December 13-15, 2005, pp. 129-132.

Lim (2005) defined eigenvalues for general real tensors in the real field. The l^2 eigenvalues of tensors defined by Lim (2005) are Z-eigenvalues of Qi (2005), while the l^k eigenvalues of tensors defined by Lim (2005) are H-eigenvalues in Qi (2005). Notably, Lim (2005) proposed a multilinear generalization of the Perron-Frobenius theorem based upon the notion of l^k eigenvalues (H-eigenvalues) of tensors. extended his definition in his report and cited my paper. In particular, Lim mentioned a **multilinear generalization of the Perron-Frobenius theorem** based upon the notion of H-eigenvalues of tensors.

5.1. Multi-Linear Data Analysis and Hypergraphs

Lek-Heng visited me in January 2006. He pointed out that another potential application of eigenvalues of tensors is on **hypergraphs**. A graph can be described by its adjacency matrix. The properties of eigenvalues of the adjacency matrix of a graph are related to the properties of the graph. This is the topic of spectral graph theory. A hypergraph can be described by a (0, 1)- supersymmetric tensor, which is called its adjacency tensor. Are the properties of the eigenvalues of the adjacency tensor of a hypergraph related to the properties of the hypergraph? This can be investigated.

A more important motivation of Lek-Heng is the **statistical analysis of multi-way data**. This is the main theme of this workshop.