Available online at www.sciencedirect.com
ScienceDirect
J. Math. Anal. Appl. 325 (2007) 1363–1377
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

Eigenvalues and invariants of tensors

Liqun Qi [1]

Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Received 4 August 2005
Available online 27 March 2006
Submitted by J.A. Filar

Abstract

A tensor is represented by a supermatrix under a co-ordinate system. In this paper, we define E-eigenvalues and E-eigenvectors for tensors and supermatrices. By the resultant theory, we define the E-characteristic polynomial of a tensor. An E-eigenvalue of a tensor is a root of the E-characteristic polynomial. In the regular case, a complex number is an E-eigenvalue if and only if it is a root of the E-characteristic polynomial. We convert the E-characteristic polynomial of a tensor to a monic polynomial and show that the coefficients of that monic polynomial are invariants of that tensor, i.e., they are invariant under co-ordinate system changes. We call them principal invariants of that tensor. The maximum number of principal invariants of mth order n-dimensional tensors is a function of m and n. We denote it by $d(m,n)$ and show that $d(1,n)=1$, $d(2,n)=n$, $d(m,2)=m$ for $m \geqslant 3$ and $d(m,n) \leqslant m^{n-1}+\cdots+m$ for $m,n \geqslant 3$. We also define the rank of a tensor. All real eigenvectors associated with nonzero E-eigenvalues are in a subspace with dimension equal to its rank.

Keywords: Eigenvalue; Tensor; Invariant; Supermatrix; Rank

1. Introduction

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. It was Einstein who applied tensor analysis in his study of general relativity in 1916. This made tensor analysis

E-mail address: maqilq@polyu.edu.hk.
[1] The work of the author was supported by the Research Grant Council of Hong Kong.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2006.02.071

an important tool in theoretical physics, continuum mechanics and many other areas of science and engineering [7,10,16,23].

A tensor is a physical quantity which is independent from co-ordinate system changes. A zero order tensor is a scalar. A first order tensor is a vector. The representation of a second order tensor in a co-ordinate system is a square matrix. The representation matrices of a second order tensor are square matrices similar to each other.

In applications, there are higher order tensors. For example, the elasticity tensor in continuum mechanics is a fourth order tensor [7,10,16,23]. In this paper, in a co-ordinate system, we let an mth order n-dimensional tensor $\mathbf{T}$ be represented by a supermatrix

$$T = \left(T_{i_1}^{\cdot i_2 \cdots i_m}\right),$$

where $i_1, \ldots, i_m = 1, \ldots, n$. Here, i_1 is a subscript, while $i_2, \ldots, i_m$ are superscript. We will explain this in the next section. We use the word "supermatrix" to distinguish it from the tensor.

An important concept in tensor analysis is the invariant. An invariant of a tensor is a scalar associated with that tensor. It does not vary under co-ordinate changes. For example, the magnitude of a vector is an invariant of that vector. For second order tensors, there is a well-developed theory of eigenvalues and invariants. A real second order n-dimensional tensor has n eigenvalues. The product of these eigenvalues is equal to the determinant of the tensor. The sum of these eigenvalues is equal to the trace of the tensor. A complex number is an eigenvalue of a real second order tensor if and only if it is a root of the characteristic polynomial of that second order tensor. The eigenvalues of a real symmetric second order tensor are all real. A real symmetric second order tensor is positive definite (semidefinite) if and only if all of its eigenvalues are positive (nonnegative). The coefficients of the characteristic polynomial of a second order tensor are invariants of that tensor. They are called the principal invariants of that tensor. There are n principal invariants for that tensor.

Can these be extended to the higher order case?

In the recent few years, motivated by the study of signal processing, people [12,14,15,25] studied the best rank-1 approximation of higher order "tensors." Very recently, motivated by the positive definiteness issue of multivariate forms arising in the stability study of nonlinear autonomous systems via Liapunov's direct method in automatic control [1–4,6,9,11,13,24], in [19], Qi introduced the concept of eigenvalues of a supersymmetric "tensor," and developed their theory based upon the theory of resultants [5,8,22].

In [19], two kinds of eigenvalues are defined for real supersymmetric tensors: eigenvalues and E-eigenvalues. For real square symmetric matrices, these two definitions are the same as the classical definition of eigenvalues. An eigenvalue (E-eigenvalue) with a real eigenvector (E-eigenvector) is called an H-eigenvalue (Z-eigenvalue). When the order is even, H-eigenvalues and Z-eigenvalues always exist. An even order supersymmetric tensor is positive definite (semidefinite) if and only if all of its H-eigenvalues or all of its Z-eigenvalues are positive (nonnegative). In [18], an H-eigenvalue method for the positive definiteness identification problem is developed based upon this property. It works for $m = 4$ and $n = 3$. In [20], geometric meanings of Z-eigenvalues are discussed.

Independently, with a variational approach, Lim also defines eigenvalues of tensors in [17] in the real field. The l^2 eigenvalues of tensors defined in [17] are Z-eigenvalues in [19], while the l^k eigenvalues of tensors defined in [17] are H-eigenvalues in [19]. Notably, Lim [17] proposed a multilinear generalization of the Perron–Frobenius theorem based upon the notion of l^k eigenvalues (H-eigenvalues) of tensors.

In general, eigenvalues and E-eigenvalues of tensors are very different. Many properties of eigenvalues of square matrices can be extended to eigenvalues of higher order supersymmetric tensors. In [19], the symmetric hyperdeterminant of a supersymmetric tensor is defined. Based upon this, the characteristic polynomial of a supersymmetric tensor is defined. A complex number is an eigenvalue of a supersymmetric tensor if and only if it is a root of the characteristic polynomial of that tensor. The product of all eigenvalues of a supersymmetric tensor is equal to the symmetric hyperdeterminant of that tensor. The sum of all eigenvalues of an mth order n-dimensional supersymmetric tensor is equal to the trace of that tensor multiplied with $(m-1)^{n-1}$. There are exactly $n(m-1)^{n-1}$ eigenvalues for that tensor. A Gerschgorin-type theorem also holds for eigenvalues of supersymmetric tensors. These properties do not hold for E-eigenvalues of higher order supersymmetric tensors.

However, on the other hand, the invariant property does not hold for eigenvalues of higher order supersymmetric tensors. As we can see in this paper, the invariant property holds for E-eigenvalues of higher order (not necessarily symmetric) tensors. This motivated us to extend E-eigenvalues defined in [19] to the nonsymmetric case and to study their invariant property.

To do this, we need to distinguish tensor and supermatrices. The "tensors" studied in [12,14, 15,19,25] are fixed in a co-ordinate system. The tensors in theoretical physics and continuum mechanics are physical quantities which are invariant under co-ordinate system changes [7,10, 16,23]. Hence, the "tensors" studied in [12,14,15,19,25] are actually supermatrices in this paper.

In the next section, we will review the notation of tensors and supermatrices, which will be used in this paper. In Section 3, we will define E-eigenvalues and E-eigenvectors for (not necessarily symmetric) tensors and supermatrices. The E-eigenvalues of a tensor are the same as the E-eigenvalues of the representation supermatrix of that tensor in an orthonormal co-ordinate system. We then define E-characteristic polynomial for a tensor in Section 4. An E-eigenvalue of a tensor is a root of the E-characteristic polynomial. In the regular case, a complex number is an E-eigenvalue of a tensor if and only if it is a root of the E-characteristic polynomial. If not all the coefficients of the E-characteristic polynomial of a tensor are zero, then we may convert the E-characteristic polynomial to a monic polynomial. We show that the coefficients of that monic polynomial are invariants of that tensor, i.e., they are invariant under co-ordinate system changes. We call them principal invariants of that tensor. The maximum number of principal invariants of mth order n-dimensional tensors is a function of m and n. We denote it by $d(m,n)$ and show that $d(1,n)=1$, $d(2,n)=n$, $d(m,2)=m$ for $m \geqslant 3$ and $d(m,n) \leqslant m^{n-1}+\cdots+m$ for $m,n \geqslant 3$. In Section 5, we will define recession vectors for an mth order n-dimensional tensor. All the recession vectors of a tensor form a linear subspace. We call it the recession space of that tensor. Let its dimension be d. We call $r=n-d$ the rank of that tensor. We will show that an eigenvector of that tensor associated with a nonzero Z-eigenvalue is always perpendicular to a recession vector. Thus, all Z-eigenvectors associated with nonzero eigenvalues are in a subspace with dimension of r. Some final remarks will be given in Section 6.

2. Tensors and supermatrices

We use the Einstein summation convention and assume that $n \geqslant 2$ throughout this paper.

A part of our notation is similar to the notion used in [10,16,23].

Let $\mathbf{E}$ be a real n-dimensional Euclidean space. We use small bold letters $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$ to denote vectors in $\mathbf{E}$. Let $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ be a basis of $\mathbf{E}$ and $\{\mathbf{g}^1, \ldots, \mathbf{g}^n\}$ be the reciprocal basis. Then we have their scalar products as

$$\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j,$$

where

$$\delta^i_j = \delta^{ij} = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases}$$

is the Kronecker symbol.

For any $\mathbf{x} \in \mathbf{E}$, under the basis $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$, $\mathbf{x}$ has covariant co-ordinates

$$x_i = \mathbf{x} \cdot \mathbf{g}_i$$

and contravariant co-ordinates

$$x^i = \mathbf{x} \cdot \mathbf{g}^i.$$

Then we have

$$\mathbf{x} = x^i \mathbf{g}_i = x_i \mathbf{g}^i.$$

For any $\mathbf{x}, \mathbf{y} \in \mathbf{E}$, their scalar product is

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x} = x^i y_i = x_i y^i.$$

We use $\bar{\mathbf{E}}$ to denote the complex vector space $\mathbf{E} + \sqrt{-1}\mathbf{E}$. For $\mathbf{x} \in \bar{\mathbf{E}}$, there are $\mathbf{u}, \mathbf{v} \in \mathbf{E}$ such that

$$\mathbf{x} = \mathbf{u} + \sqrt{-1}\mathbf{v}.$$

We use capital bold letters, other than $\mathbf{E}$, $\mathbf{V}$, $\mathbf{R}$ and $\mathbf{C}$ to denote second order or higher order tensors defined on $\mathbf{E}$. An mth order tensor $\mathbf{T}$ defined on $\mathbf{E}$ is in fact an m-linear operator on $\mathbf{E}$. If the range of $\mathbf{T}$ is real, then $\mathbf{T}$ is a real tensor. We use $\mathbf{E}^m$ to denote the set of all real mth order tensors defined on $\mathbf{E}$. For $\mathbf{T} \in \mathbf{E}^m$, we regard the supermatrix

$$T = \left(T_{i_1}^{\cdot i_2 \cdots i_m}\right)$$

as its representation supermatrix under the basis $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$. Then

$$\mathbf{T} = T_{i_1}^{\cdot i_2 \cdots i_m} \mathbf{g}^{i_1} \mathbf{g}_{i_2} \cdots \mathbf{g}_{i_m}.$$

The representation matrix of the metric tensor $\mathbf{G}$ is always the unit matrix

$$I = \left(\delta_i^j\right).$$

We use capital ordinary letters, other than E, to denote square matrices and higher order supermatrices. Under an orthonormal co-ordinate system, we do not need to distinguish the covariant and contravariant indices of the entries of a supermatrix. Then we may denote an mth order n-dimensional supermatrix A by

$$A = (A_{i_1 \cdots i_m}).$$

We denote the set of all mth order n-dimensional real supermatrices by E^m. Then E is the set of all n-dimensional real "vectors"

$$x = (x_i) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

"Vectors" in E are different from vectors in $\mathbf{E}$. Vectors in $\mathbf{E}$ are first order tensors. They are invariant under co-ordinate system changes. On the other hand, "vectors" in E are merely vectors in the sense of linear algebra, not tensor analysis [10,16,23]. But we will not distinguish these two kinds of vectors strictly in the later discussion.

Thus, tensors discussed in [12,14,15,19,21,25] are actually supermatrices in this paper.

A supermatrix is called supersymmetric (completely symmetric) if its entries are invariant under any permutation of their indices. A tensor is called supersymmetric if its representation supermatrix is supersymmetric.

For $A \in E^m$ and $x \in \bar{E}$, we denote

$$Ax^{m-1} = (A_{i_1 i_2 \cdots i_m} x_{i_2} \cdots x_{i_m}),$$
$$xA = (x_{i_1} A_{i_1 i_2 \cdots i_m})$$

and

$$Ax^m \equiv xAx^{m-1} = A_{i_1 i_2 \cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}.$$

Similarly, for $\mathbf{T} \in \mathbf{E}^m$ and $\mathbf{x} \in \bar{\mathbf{E}}$, we denote

$$\mathbf{T} \cdot \mathbf{x}^{m-1} = T_{i_1}^{\cdot i_2 \cdots i_m} x_{i_2} \cdots x_{i_m} \mathbf{g}^{i_1},$$
$$\mathbf{x} \cdot \mathbf{T} = x^{i_1} T_{i_1}^{\cdot i_2 \cdots i_m} \mathbf{g}_{i_2} \cdots \mathbf{g}_{i_m}$$

and

$$\mathbf{T} \cdot \mathbf{x}^m \equiv \mathbf{x} \cdot \mathbf{T} \cdot \mathbf{x}^{m-1} = x^{i_1} T_{i_1}^{\cdot i_2 \cdots i_m} x_{i_2} \cdots x_{i_m}.$$

An even order supersymmetric supermatrix A is called positive definite (semidefinite) if for all $x \in \bar{E}$, $x \neq 0$, we have $Ax^m > 0$ ($Ax^m \geqslant 0$). An even order supersymmetric tensor $\mathbf{T}$ is called positive definite (semidefinite) if for all $\mathbf{x} \in \bar{\mathbf{E}}, \mathbf{x} \neq 0$, we have $\mathbf{T} \cdot \mathbf{x}^m > 0$ ($\mathbf{T} \cdot \mathbf{x}^m \geqslant 0$). Suppose that we have a new basis $\{\mathbf{g}_{1'}, \ldots, \mathbf{g}_{n'}\}$ for $\mathbf{E}$ with the corresponding reciprocal basis $\{\mathbf{g}^{1'}, \ldots, \mathbf{g}^{n'}\}$. Then we have transformation matrices $(\beta_{i'}^{i})$ and $(\beta_{i}^{i'})$ such that

$$\left(\beta_{i'}^{i}\right)^{-1} = \left(\beta_{i}^{i'}\right). \tag{1}$$

For $\mathbf{x} \in \bar{\mathbf{E}}$, we have

$$x_i = \beta_i^{i'} x_{i'} \tag{2}$$

and

$$x^i = \beta_{i'}^{i} x^{i'}. \tag{3}$$

For $\mathbf{T} \in \mathbf{E}^m$, we have

$$T_{i_{1'}}^{\cdot i_{2'} \cdots i_{m'}} = T_{i_1}^{\cdot i_2 \cdots i_m} \beta_{i_1'}^{i_1} \beta_{i_2}^{i_2'} \cdots \beta_{i_m}^{i_m'}. \tag{4}$$

For the derivation of (3) and (4), please see [10,23].

It is not difficult to show that the scalar product $\mathbf{x} \cdot \mathbf{x}$ for $\mathbf{x} \in \bar{\mathbf{E}}$ is invariant under co-ordinate system changes.

Note that in the discussion involving tensors, superscripts of small ordinary letters are used for contravariant indices, while in the other occasions, superscripts are used for powers.

3. E-eigenvalues of tensors and supermatrices

We use small Greek letters to denote real and complex numbers. We use $\mathbf{R}$ and $\mathbf{C}$ to denote the set of all real numbers and the set of all complex numbers, respectively.

Suppose that $\mathbf{T} \in \mathbf{E}^m$ and $m \geqslant 1$. We say a complex number λ is an E-eigenvalue of the tensor $\mathbf{T}$ if there exists $\mathbf{x} \in \bar{\mathbf{E}}$ such that

$$\begin{cases} \mathbf{T} \cdot \mathbf{x}^{m-1} = \lambda \mathbf{x}, \\ \mathbf{x} \cdot \mathbf{x} = 1. \end{cases} \tag{5}$$

In this case, we say that $\mathbf{x}$ is an E-eigenvector of the tensor $\mathbf{T}$ associated with the E-eigenvalue λ. If an eigenvalue has a real eigenvector, then we say that it is a Z-eigenvalue.

Theorem 1.

(i) *The E-eigenvalues of a tensor are invariants of that tensor.*
(ii) *If $\mathbf{x}$ is an E-eigenvector of $\mathbf{T}$ associated with an eigenvalue λ, then*

$$\lambda = \mathbf{T} \cdot \mathbf{x}^m.$$

Proof. (i) As we said before, the second equation of (5) is invariant under co-ordinate system changes. Hence, we only need to consider the first equation of (5).

Under the basis $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$, the first equation of (5) has the form

$$T_{i_1}^{\cdot i_2 \cdots i_m} x_{i_2} \cdots x_{i_m} = \lambda x_{i_1}. \tag{6}$$

By (2), under the new basis $\{\mathbf{g}_{1'}, \ldots, \mathbf{g}_{n'}\}$, (6) becomes

$$T_{i_1}^{\cdot i_2 \cdots i_m} \beta_{i_2}^{i_2'} \cdots \beta_{i_m}^{i_m'} x_{i_{2'}} \cdots x_{i_{m'}} = \lambda \beta_{i_1}^{i_1'} x_{i_{1'}}.$$

By (1), we have

$$T_{i_1}^{\cdot i_2 \cdots i_m} \beta_{i_1'}^{i_1} \beta_{i_2}^{i_2'} \cdots \beta_{i_m}^{i_m'} x_{i_{2'}} \cdots x_{i_{m'}} = \lambda x_{i_{1'}}.$$

By (4), we have

$$T_{i_{1'}}^{\cdot i_{2'} \cdots i_{m'}} x_{i_{2'}} \cdots x_{i_{m'}} = \lambda x_{i_{1'}}.$$

This shows that the first equation of (5) is also invariant under co-ordinate system changes. This completes the proof of (i).

(ii) This follows from (5) directly. □

Suppose that $A \in E^m$ and $m \geqslant 1$. We say a complex number λ is an E-eigenvalue of the supermatrix A if there exists $x \in \bar{E}$ such that

$$\begin{cases} Ax^{m-1} = \lambda x, \\ x^T x = 1. \end{cases} \tag{7}$$

In this case, we say that x is an E-eigenvector of the supermatrix A associated with the E-eigenvalue λ. If an E-eigenvalue has a real E-eigenvector, then we say that it is a Z-eigenvalue. These generalize the definitions of E-eigenvalues and Z-eigenvalues for supersymmetric supermatrices in [19] to the nonsymmetric case.

We may compare the definitions of eigenvalues and E-eigenvalues for a supermatrix. We may first generalize the definition of eigenvalues of a supersymmetric supermatrix in [19] to the nonsymmetric case. Let $x \in \bar{E}$, we denote $x^{[m-1]}$ as a vector in $\bar{E}$ with its components as x_i^{m-1}. We say a complex number λ is an eigenvalue of the supermatrix A if there exists $x \in \bar{E}$ such that

$$Ax^{m-1} = \lambda x^{[m-1]}. \tag{8}$$

In this case, we say that x is an eigenvector of the supermatrix A associated with the eigenvalue λ. If an eigenvalue has a real eigenvector, then we say that it is an H-eigenvalue.

Note that (8) is homogeneous, but (7) is not. In fact, in a certain sense, eigenvalues of supermatrices defined by (8) are genuine extensions of eigenvalues of square matrices, but (7) is not. When $m = 2$ and A is nonsymmetric, the definition (7) misses such eigenvalues whose eigenvectors satisfies $x^T x = 0$ but $x \neq 0$. For example,

$$x = \begin{pmatrix} 1 \\ \sqrt{-1} \end{pmatrix}.$$

Also, by [19], the number of eigenvalues of an mth order n-dimensional supersymmetric supermatrix only depends upon m and n. It is $n(m-1)^{n-1}$. We may easily extend that result to the nonsymmetric case. But the number of E-eigenvalues of an mth order n-dimensional supersymmetric supermatrix varies for different supermatrices. As we see above, the number of E-eigenvalues of an n-dimensional square matrix is $n - 2k \geqslant 0$, where k is an integer. When $m \geqslant 3$, it is possible that all complex numbers are E-eigenvalues for a certain supermatrix. For example, let $m = 3$ and $n = 2$. Let $A_{111} = A_{221} = 1$ and other $A_{ijk} \equiv 0$. Then (7) becomes

$$\begin{cases} x_1^2 = \lambda x_1, \\ x_1 x_2 = \lambda x_2, \\ x_1^2 + x_2^2 = 1. \end{cases} \tag{9}$$

For any complex number λ,

$$(x_1, x_2) = \left(\lambda, \sqrt{1-\lambda^2}\right)$$

is a nonzero solution of (9). Hence, all complex numbers are E-eigenvalues. We call such a case the *singular* case.

Though E-eigenvalues are somewhat irregular, but they are invariants under co-ordinate system changes, in the sense of Theorem 1. This is why we need to study them in this paper.

We call the basis $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ orthonormal if

$$\mathbf{g}_i \cdot \mathbf{g}_j = \delta_{ij}.$$

Comparing (5) and (7), we have the following theorem.

Theorem 2. *The E-eigenvalues of a tensor are the same as the E-eigenvalues of its representation supermatrix under an orthonormal basis.*

By Theorem 2 and [19], we have the following theorem.

Theorem 3. *A supersymmetric tensor/supermatrix always has Z-eigenvalues. An even order supersymmetric tensor/supermatrix is positive definite* (*semidefinite*) *if and only if all of its Z-eigenvalues are positive* (*nonnegative*).

4. The E-characteristic polynomial

Suppose that $A \in E^m$ and $m \geqslant 3$ throughout this section.

Suppose now that m is even and $m = 2h$. Let

$$F(x) = Ax^{m-1} - \lambda\left(x^T x\right)^{h-1} x.$$

According to the resultant theory [5,8,22], the resultant of $F(x)$ is an irreducible polynomial of the coefficient of $F(x)$, i.e., a one-dimensional polynomial of λ, which vanishes as long as $F(x) = 0$ has a nonzero solution. We call this polynomial of λ as the E-characteristic polynomial of A, and denote it by $\psi(\lambda)$. This definition generalizes the definition of the E-characteristic polynomial of an even order supersymmetric supermatrix, given in [19] to the nonsymmetric case.

Suppose now that m is odd. Let

$$G(x_0, x) = \begin{cases} Ax^{m-1} - \lambda x_0^{m-2} x, \\ x_0^2 - x^T x. \end{cases} \tag{10}$$

Again, according to the resultant theory [5,8,22], the resultant of $G(x_0, x)$ is an irreducible polynomial of the coefficient of $G(x_0, x)$, i.e., a one-dimensional polynomial of λ, which vanishes as long as $G(x_0, x) = 0$ has a nonzero solution. We call this polynomial of λ as the E-characteristic polynomial of A, and denote it by $\psi(\lambda)$. This definition further generalizes the definition of the E-characteristic polynomial of an even order supermatrix to the odd order case.

We say that A is *irregular* if there is an $x \in \bar{E}$ such that $x \neq 0$, $Ax^{m-1} = 0$ and $x^T x = 0$. Otherwise we say that A is *regular*.

Suppose that $\mathbf{T} \in \mathbf{E}^m$. We call the E-characteristic polynomial of a representation supermatrix of $\mathbf{T}$ as the E-characteristic polynomial of $\mathbf{T}$. We say that $\mathbf{T}$ is irregular if there is an $\mathbf{x} \in \bar{\mathbf{E}}$ such that $\mathbf{x} \neq \mathbf{0}$, $\mathbf{T} \cdot \mathbf{x}^{m-1} = \mathbf{0}$ and $\mathbf{x} \cdot \mathbf{x} = 0$. Otherwise we say that $\mathbf{T}$ is regular.

It is not difficult to prove that the regularity is also an invariant property of a tensor. A tensor is regular if and only if its supermatrix is regular.

Theorem 4. *Suppose that $m \geqslant 3$. Then an E-eigenvalue of an mth order tensor/supermatrix is a root of the E-characteristic polynomial of that tensor/supermatrix. If that tensor/supermatrix is regular, then a complex number is an E-eigenvalue of that tensor/supermatrix if and only if it is root of the E-characteristic polynomial of that tensor/supermatrix.*

Proof. Suppose that $A \in E^m$ and $m \geqslant 3$. Let λ be an E-eigenvalue of A and x is an E-eigenvector of A associated with λ. By (7), when m is even, x is a nonzero solution of $F(x) = 0$ for that λ; when m is odd, $(x_0 = \sqrt{x^T x}, x)$ is a nonzero solution of $G(x_0, x) = 0$ for that λ. According to the resultant theory and our definition, we have $\psi(\lambda) = 0$.

On the other hand, suppose that A is regular. Let λ be a root of $\psi(\lambda) = 0$. By the resultant theory and our definition, when m is even, there is an $x \in \bar{E}$, $x \neq 0$ such that $F(x) = 0$; when m is odd, there is an $x \in \bar{E}$ and an $x_0 \in C$, $x \neq 0$ such that $G(x_0, x) = 0$. If $x^T x = 0$, then by $F(x) = 0$ or $G(x_0, x) = 0$, we have $Ax^{m-1} = 0$. This contradicts the assumption that A is regular. Hence, $x^T x \neq 0$. Let

$$\hat{x} = \frac{x}{\sqrt{x^T x}}.$$

We see that (7) is satisfied with λ and $\hat{x}$. This implies that λ is an E-eigenvalue of A.

The case for a tensor follows this and the relation between tensors and supermatrices. □

Note that when $m = 2$, the E-characteristic polynomial is exactly the characteristic polynomial in the classical sense. In that case, a root of $\psi(\lambda) = 0$ is an eigenvalue of the second order tensor or the square matrix in the classical sense, though it may not satisfy (7) as its eigenvector x may not satisfy $x^T x \neq 0$. Thus, in the regular case, the E-characteristic polynomial may inherit properties of the characteristic polynomial in the classical sense.

By the definition, when m is odd, if λ is an E-eigenvalue of a tensor (supermatrix) with an eigenvector $\mathbf{x}$ (x), then $-\lambda$ is an E-eigenvalue of that tensor (supermatrix) with an eigenvector $-\mathbf{x}$ ($-x$). Thus, when m is odd, the E-characteristic polynomial only contains even power terms of λ.

Then we may denote the E-characteristic polynomial of a tensor $\mathbf{T}$ by

$$\psi(\lambda) = \begin{cases} \sum_{j=0}^{h} a_j(\mathbf{T})\lambda^j, & \text{if } m \text{ is even}, \\ \sum_{j=0}^{h} a_j(\mathbf{T})\lambda^{2j}, & \text{if } m \text{ is odd}, \end{cases} \tag{11}$$

where $a_h(\mathbf{T}) \neq 0$, $h = h(\mathbf{T})$ depends upon $\mathbf{T}$. For example, let $m = 3$ and $n = 2$. Let $A_{111} = A_{222} = 1$ and other $A_{ijk} \equiv 0$. Then (10) becomes

$$G(x_0, x) = \begin{cases} x_1^2 - \lambda x_0 x_1, \\ x_2^2 - \lambda x_0 x_2, \\ x_0^2 - x_1^2 - x_2^2. \end{cases}$$

It has nonzero roots if and only if $\lambda^2 = 1$ or $2\lambda^2 = 1$. Hence, the E-characteristic polynomial of the tensor $\mathbf{T}$, whose representation supermatrix under an orthonormal basis is A, is

$$\psi(\lambda) = (\lambda^2 - 1)(2\lambda^2 - 1).$$

It is possible that for some $\mathbf{T}$, it has a *zero E-characteristic polynomial*. In this case, we denote $h(\mathbf{T}) = -1$. Define

$$d(m, n) := \max\{h(\mathbf{T}):\ \mathbf{T} \in \mathbf{E}^m\}. \tag{12}$$

When $h(\mathbf{T}) = -1$, by Theorem 4, if $\mathbf{T}$ is regular, then all complex numbers are E-eigenvalues of $\mathbf{T}$, i.e., $\mathbf{T}$ is singular. Hence, in the regular case, if $\mathbf{T}$ does not have a zero E-characteristic polynomial, the number of E-eigenvalues is not greater than $d(m, n)$ when m is even, and $2d(m, n)$ when m is odd. Thus, we have the following corollary.

Corollary 1. *In the regular case, a tensor/supermatrix has either only a finite number of E-eigenvalues, or all complex numbers are its E-eigenvalues, i.e., it is singular.*

Note that the singular example given in (9) is regular.

In the irregular case, when $m \geqslant 3$, for all complex values of λ, $F(x) = 0$ or $G(x_0, x) = 0$ has nonzero solutions. Thus, in the irregular case, we always have a zero E-characteristic polynomial. But it is not necessarily singular in this case.

We then have two questions.

Question 1. Is Corollary 1 also true in the irregular case?

Question 2. Is a regular completely symmetric supermatrix always nonsingular?

By Theorem 1 and the relations between roots and coefficients of a one-dimensional polynomial, we now have the following theorem.

Theorem 5. *Let $\mathbf{T} \in \mathbf{E}^m$ and $m \geqslant 1$. Suppose that $h(\mathbf{T}) \geqslant 1$, where $h(\mathbf{T})$ is defined in* (11). *Then*

$$\frac{(-1)^{h-j} a_j(\mathbf{T})}{a_h(\mathbf{T})}$$

for $j = 0, \ldots, h-1$, are invariants of that tensor $\mathbf{T}$.

We call them principal invariants of that tensor $\mathbf{T}$. It is also not difficult to show that $h(\mathbf{T})$ is also an invariant of $\mathbf{T}$. When $h(\mathbf{T}) = 0$ or -1, we regard that $\mathbf{T}$ has no principal invariants. Then, the maximum number of principal variants of mth order n-dimensional tensors is $d(m,n)$. We have the following proposition.

Proposition 1. *When m is even, $d(m,n)$ is the maximum value of numbers of E-eigenvalues of regular mth order n-dimensional tensors/supermatrices which have a finite number of E-eigenvalues. When m is odd, $2d(m,n)$ is the maximum value of numbers of E-eigenvalues of regular mth order n-dimensional tensors/supermatrices which have a finite number of E-eigenvalues.*

Proof. In the irregular case we always have zero E-characteristic polynomials. In the regular case, by Corollary 1, if a tensor/supermatrix has infinitely many E-eigenvalues, then it also has a zero E-characteristic polynomial. Hence, to find the value of $d(m,n)$, we only need to consider regular mth order n-dimensional tensors/supermatrices which have a finite number of E-eigenvalues. The conclusions now follow from the structure of the E-characteristic polynomial. □

We then have the following theorem.

Theorem 6.

(i) $d(1,n) = 1$. *This is the vector case. Each vector has exactly two E-eigenvalues, which are its magnitude and its magnitudes times -1.*
(ii) $d(2,n) = n$. *This is the case of second order tensors. An E-eigenvalue of a second order tensor is an eigenvalue of that tensor in the classical sense. A second order tensor has n eigenvalues in the classical sense. If the second order tensor is symmetric, then all of these n eigenvalues are Z-eigenvalues.*
(iii) *If $m \geqslant 3$, then $d(m,2) = m$. Let A be the representation supermatrix of an mth order two-dimensional tensor $\mathbf{T}$ in an orthonormal co-ordinate system. If*

$$\sum_{i_2,\ldots,i_m=1}^{2} A_{2i_2\cdots i_m} x_1 x_{i_2} \cdots x_{i_m} \equiv \sum_{i_2,\ldots,i_m=1}^{2} A_{1i_2\cdots i_m} x_2 x_{i_2} \cdots x_{i_m} \tag{13}$$

does not hold, then the number of E-eigenvalues of an mth order two-dimensional tensor is $m-2k$ when m is even, and $2(m-2k)$ when m is odd. This implies that the singular case occurs only if (13) *holds.*

Furthermore, when A is completely symmetric, (13) *holds only when m is even and*

$$Ax^m = c_0\left(x_1^2+x_2^2\right)^{\frac{m}{2}}, \tag{14}$$

where c_0 is a constant. In this case, A has only one E-eigenvalue c_0. Thus, a completely symmetric two-dimensional tensor/supermatrix has only a finite number of E-eigenvalues.

(iv) *Suppose that $n, m \geqslant 3$. Then*

$$d(m,n) \leqslant m^{n-1} + \cdots + m.$$

Proof. (i) This follows from the definition directly.

(ii) This is common knowledge of tensor analysis [10,16,23].

(iii) When $n=2$, (7) can be written as

$$\begin{cases} \sum_{i_2,\ldots,i_m=1}^2 A_{1i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = \lambda x_1, \\ \sum_{i_2,\ldots,i_m=1}^2 A_{2i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = \lambda x_2, \\ x_1^2+x_2^2 = 1, \end{cases} \tag{15}$$

which is equivalent to

$$\begin{cases} \sum_{i_2,\ldots,i_m=1}^2 A_{2i_2\cdots i_m}x_1x_{i_2}\cdots x_{i_m} - \sum_{i_2,\ldots,i_m=1}^2 A_{1i_2\cdots i_m}x_2x_{i_2}\cdots x_{i_m} = 0, \\ x_1^2+x_2^2 = 1, \\ \lambda = Ax^m. \end{cases} \tag{16}$$

We now suppose that (13) does not hold. The first equation of (16) is a homogeneous polynomial of x_1 and x_2. Its degree is m, as we have assumed that (13) does not hold. It can be decomposed to m linear factors. Each linear factor corresponds to a solution of (16). If such a solution satisfies

$$x_1^2+x_2^2 \neq 0,$$

then it corresponds to one solution of (7) when m is even, and two solutions of (7) when m is odd. Then, the number of E-eigenvalues of an mth order two-dimensional tensor is $m-2k$ when m is even, and $2(m-2k)$ when m is odd. This implies that the singular case occurs only if (13) holds.

Suppose now that (13) holds. Then there is a real homogeneous polynomial $b(x_1,x_2)$ such that

$$\begin{cases} \sum_{i_2,\ldots,i_m=1}^2 A_{1i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = x_1 b(x_1,x_2), \\ \sum_{i_2,\ldots,i_m=1}^2 A_{2i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = x_2 b(x_1,x_2). \end{cases} \tag{17}$$

If $b(x_1,x_2)$ has a factor $(x_1^2+x_2^2)$, then by (17),

$$\begin{cases} \sum_{i_2,\ldots,i_m=1}^2 A_{1i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = 0, \\ \sum_{i_2,\ldots,i_m=1}^2 A_{2i_2\cdots i_m}x_{i_2}\cdots x_{i_m} = 0, \\ x_1^2+x_2^2 = 0, \end{cases} \tag{18}$$

has nonzero solutions. This means that A is irregular. Then $h(\mathbf{T}) = -1$.

If $b(x_1, x_2)$ has no factor $(x_1^2 + x_2^2)$, then (18), has no nonzero solutions. This means that A is regular. By Corollary 1, A either has a finite number of E-eigenvalues, or it is singular. By (15), we have

$$\begin{cases} \lambda = b(x_1, x_2), \\ x_1^2 + x_2^2 = 1. \end{cases} \tag{19}$$

Let $x_1 = \cos\theta, x_2 = \sin\theta$. Then for any $\theta \in C$,

$$\lambda = b(\cos\theta, \sin\theta)$$

is a solution of (19). This says that λ is a continuous function of θ. Then λ cannot take more than one isolated values. It is not difficult to see that λ takes one single value only when m is even and

$$b(x_1, x_2) = c_1\left(x_1^2 + x_2^2\right)^{\frac{m-2}{2}},$$

where c_1 is a nonzero constant. Since we assume $b(x_1, x_2)$ has no factor $(x_1^2 + x_2^2)$, this case does not occur when $m \geqslant 3$. Then it must be singular. This implies that $h(\mathbf{T}) = -1$, whenever (13) holds.

Combining the two cases that (13) holds and does not hold, by Proposition 1, we have $d(m, 2) = m$ when $m \geqslant 3$.

Suppose now that A is completely symmetric and (13) holds. Then by (17),

$$\begin{cases} Ax^{m-1} = b(x_1, x_2)x, \\ Ax^m = b(x_1, x_2)(x_1^2 + x_2^2). \end{cases} \tag{20}$$

Let $y = Ax^m$. Then $\frac{dy}{dx} = mAx^{m-1}$. By (20), we have

$$\frac{1}{y}\frac{dy}{dx} = \frac{mx}{x_1^2 + x_2^2}.$$

This implies that (14) holds and m is even. Thus, by (16), the E-eigenvalues of A only takes one value. The last conclusion of (iii) follows now.

(iv) Let A be the representation supermatrix of an mth order n-dimensional tensor $\mathbf{T}$ in an orthonormal co-ordinate system.

If $x_1 \neq 0$, then the solutions of (7) may be found by solving the system of polynomial equations

$$\begin{cases} \sum_{i_2 \cdots i_m = 1}^n A_{i i_2 \cdots i_m} x_1 x_{i_2} \cdots x_{i_m} = \sum_{i_2 \cdots i_m = 1}^n A_{1 i_2 \cdots i_m} x_i x_{i_2} \cdots x_{i_m}, \quad i = 2, \ldots, n, \\ \sum_{i=1}^n x_i^2 = 1, \end{cases} \tag{21}$$

and calculating

$$\lambda = Ax^m. \tag{22}$$

By the Bezout Theorem [5,8], (21) either has infinitely many solutions, or has at most $2m^{n-1}$ solutions.

If $x_1 = 0$ but $x_2 \neq 0$, then the solutions of (7) may be found by solving the system of polynomial equations

$$\begin{cases} \sum_{i_2 \cdots i_m = 2}^n A_{i i_2 \cdots i_m} x_2 x_{i_2} \cdots x_{i_m} = \sum_{i_2 \cdots i_m = 2}^n A_{2 i_2 \cdots i_m} x_i x_{i_2} \cdots x_{i_m}, \quad i = 3, \ldots, n, \\ \sum_{i=2}^n x_i^2 = 1, \end{cases} \tag{23}$$

and calculating (22). By the Bezout Theorem, (23) either has infinitely many solutions, or has at most $2m^{n-2}$ solutions.

Continue this process. The final case is that $x_1 = x_2 = \cdots = x_{n-2} = 0$. Then the solutions of (7) may be found by solving

$$\begin{cases} \sum_{i_2\cdots i_m=n-1}^n A_{ni_2\cdots i_m} x_{n-1} x_{i_2} \cdots x_{i_m} = \sum_{i_2\cdots i_m=2}^n A_{n-1,i_2\cdots i_m} x_n x_{i_2} \cdots x_{i_m}, \\ x_{n-1}^2 + x_n^2 = 1, \end{cases} \tag{24}$$

and calculating (22). By the Bezout Theorem, (24) either has infinitely many solutions, or has at most $2m$ solutions.

Thus, (7) either has infinitely many solutions, or has at most $2(m^{n-1} + m^{n-2} + \cdots + m)$ solutions. Because of the structures of (21), (23), ..., (24), if (7) has a finite number of solutions then such solutions appear in pairs, i.e., if x is a solution, then $-x$ is also a solution. By (22), if m is even, this implies that A either has infinitely many E-eigenvalues, or has at most $m^{n-1} + m^{n-2} + \cdots + m$ E-eigenvalues; if m is even, this implies that A either has infinitely many E-eigenvalues, or has at most $2(m^{n-1} + m^{n-2} + \cdots + m)$ E-eigenvalues.

Now, by Proposition 1 and (11), we have the conclusion of (iv). $\square$

We have one more question.

Question 3. Is the last conclusion of Theorem 6(iii) also true in general? That is, does a completely symmetric tensor/supermatrix has only a finite number of E-eigenvalues?

5. Recession vectors and the rank of a tensor

Suppose that $\mathbf{T} \in \mathbf{E}^m$. We call $\mathbf{y} \in \mathbf{E}$ a recession vector of $\mathbf{T}$ if

$$\mathbf{y} \cdot \mathbf{T} = \mathbf{0}.$$

Clearly, all recession vectors of $\mathbf{T}$ form a linear subspace $\mathbf{V}_R$ of $\mathbf{E}$. We call it the recession space of $\mathbf{T}$. Let

$$\mathbf{V}_B = \{\mathbf{x} \in \mathbf{E}\colon\ \mathbf{y} \cdot \mathbf{x} = \mathbf{0}\ \forall \mathbf{y} \in \mathbf{V}_R\}.$$

Let the dimension of $\mathbf{V}_B$ be r. We say that the rank of $\mathbf{T}$ is r. Clearly, $0 \leqslant r \leqslant n$.

Theorem 7. *If λ is a nonzero Z-eigenvalue of $\mathbf{T}$ and $\mathbf{x} \in \mathbf{E}$ is an eigenvector of $\mathbf{T}$, associated with λ, then $\mathbf{x} \in \mathbf{V}_B$.*

Proof. For any $\mathbf{y} \in V_B$, we have

$$\lambda \mathbf{y} \cdot \mathbf{x} = \mathbf{y} \cdot (\lambda \mathbf{x}) = \mathbf{y} \cdot \mathbf{T} \cdot \mathbf{x}^{m-1} = (\mathbf{y} \cdot \mathbf{T}) \cdot \mathbf{x}^{m-1} = \mathbf{0} \cdot \mathbf{x}^{m-1} = 0.$$

The conclusion follows. $\square$

The rank of a supersymmetric supermatrix and its meaning were discussed in [20].

6. Final remarks

1. One may think to eliminate the irregularity of E-eigenvalues by using

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

to replace (7). But such a definition has an indefinite factor, i.e., if λ is an eigenvalue of A defined in this way, then any complex number which has the same modulus as λ will be an eigenvalue of A in such a definition.

2. It is interesting to find answers to the three questions raised in Section 4.

3. Theorem 6(iv) states that when $n, m \geqslant 3$,

$$d(m,n) \leqslant m^{n-1} + \cdots + m.$$

This bound seems a little big. In fact, the proof of Theorem 6(iv) also works for $m = 2$. When $m = 2$, this bound is

$$2^{n-1} + \cdots + 2 = 2^n - 2.$$

This is much greater than the actually value n. So we believe that this bound can be improved. This also implies that it is possible to use Z-eigenvalues to develop a better method for the positive definiteness identification problem than the H-eigenvalue method developed in [18].

4. A further research topic is to investigate the meanings of eigenvalues of higher order tensors in nonlinear continuum mechanics and physics [10,16,23].

Acknowledgments

The author is thankful to Colin Rogers, Wolfgang Schief, Lek-Heng Lim, Guyan Ni, Peter Vassiliou and Yiju Wang for discussion and references, and to two referees for their comments.

References

[1] B.D. Anderson, N.K. Bose, E.I. Jury, Output feedback stabilization and related problems-solutions via decision methods, IEEE Trans. Automat. Control. AC20 (1975) 55–66.
[2] N.K. Bose, 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, A.R. Modaress, General procedure for multivariable polynomial positivity with control applications, IEEE Trans. Automat. Control. AC21 (1976) 596–601.
[4] N.K. Bose, R.W. Newcomb, Tellegon's theorem and multivariate realizability theory, Int. J. Electron. 36 (1974) 417–425.
[5] D. Cox, J. Little, D. O'Shea, Using Algebraic Geometry, Springer-Verlag, New York, 1998, pp. 59–86.
[6] M. Fu, Comments on 'A procedure for the positive definiteness of forms of even-order', IEEE Trans. Automat. Control. 43 (1998) 1430.
[7] D. Gao, Duality Principles in Nonconvex Systems, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
[8] I.M. Gelfand, M.M. Kapranov, A.V. Zelevinsky, Discriminants, Resultants and Multidimensional Determinants, Birkhäuser, Boston, 1994.
[9] M.A. Hasan, A.A. Hasan, A procedure for the positive definiteness of forms of even-order, IEEE Trans. Automat. Control. 41 (1996) 615–617.
[10] K.C. Huang, M.D. Xue, M.W. Lu, Tensor Analysis, second ed., Tsinghua University Publisher, Beijing, 2003 (in Chinese).
[11] E.I. Jury, M. Mansour, Positivity and nonnegativity conditions of a quartic equation and related problems, IEEE Trans. Automat. Control. AC26 (1981) 444–451.
[12] E. Kofidis, Ph. Regalia, On the best rank-1 approximation of higher-order supersymmetric tensors, SIAM J. Matrix Anal. Appl. 23 (2002) 863–884.
[13] W.H. Ku, Explicit criterion for the positive definiteness of a general quartic form, IEEE Trans. Automat. Control. 10 (1965) 372–373.
[14] L. De Lathauwer, B. De Moor, From matrix to tensor: Multilinear algebra and signal processing, in: J. McWhirter (Ed.), Mathematics in Signal Processing IV, Selected papers presented at 4th IMA Int. Conf. on Mathematics in Signal Processing, Oxford University Press, Oxford, United Kingdom, 1998, pp. 1–15.

[15] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-$(R_1, R_2, \ldots, R_N)$ approximation of higher-order tensor, SIAM J. Matrix Anal. Appl. 21 (2000) 1324–1342.
[16] K.T. Li, A.X. Huang, Tensor Analysis and Their Applications, Scientific Publisher, Beijing, 2004 (in Chinese).
[17] 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.
[18] Q. Ni, L. Qi, F. Wang, An Eigenvalue Method for the Positive Definiteness Identification Problem, Department of Applied Mathematics, The Hong Kong Polytechnic University, June, 2005.
[19] L. Qi, Eigenvalues of a real supersymmetric tensor, J. Symbolic Comput. 40 (2005) 1302–1324.
[20] L. Qi, Rank and eigenvalues of a supersymmetric tensor, the multivariate homogeneous polynomial and the algebraic hypersurface it defines, J. Symbolic Comput., in press.
[21] L. Qi, K.L. Teo, Multivariate polynomial minimization and its application in signal processing, J. Global Optim. 26 (2003) 419–433.
[22] B. Sturmfels, Solving Systems of Polynomial Equations, Amer. Math. Soc., Providence, RI, 2002.
[23] Y.R. Talpaert, Tensor Analysis and Continuum Mechanics, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
[24] F. Wang, L. Qi, Comments on ‘Explicit criterion for the positive definiteness of a general quartic form’, IEEE Trans. Automat. Control. 50 (2005) 416–418.
[25] T. Zhang, G.H. Golub, Rank-1 approximation of higher-order tensors, SIAM J. Matrix Anal. Appl. 23 (2001) 534–550.