



Contents lists available at SciVerse ScienceDirect

# Linear Algebra and its Applications

journal homepage: [www.elsevier.com/locate/laa](http://www.elsevier.com/locate/laa)



## The number of eigenvalues of a tensor

Dustin Cartwright, Bernd Sturmfels<sup>\*</sup>

Department of Mathematics, University of California, Berkeley, CA 94720, USA

### ARTICLE INFO

#### Article history:

Received 20 April 2010

Accepted 31 May 2011

Available online 18 July 2011

Submitted by V. Mehrmann

#### Keywords:

Eigenvector

Symmetric tensor

Discriminant

Toric geometry

### ABSTRACT

Eigenvectors of tensors, as studied recently in numerical multilinear algebra, correspond to fixed points of self-maps of a projective space. We determine the number of eigenvectors and eigenvalues of a generic tensor, and we show that the number of normalized eigenvalues of a symmetric tensor is always finite. We also examine the characteristic polynomial and how its coefficients are related to discriminants and resultants.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

In the current numerical analysis literature, considerable interest has arisen in extending concepts that are familiar from linear algebra to the setting of multilinear algebra. One such familiar concept is that of an eigenvalue of a square matrix. Several authors have explored definitions of eigenvalues and eigenvectors for higher-order tensors, and they have argued that these notions are useful for a wide range of applications [9]. Our approach in this paper is based on the definition of *E-eigenvalues* of tensors introduced by Qi in [10,12]. Throughout this paper, eigenvalue will mean E-eigenvalue, as defined in Definition 1.1.

We fix two positive integers  $m$  and  $n$ , and we consider order- $m$  tensors  $A = (a_{i_1 i_2 \dots i_m})$  of format  $\underbrace{n \times n \times \dots \times n}_{m \text{ times}}$  with entries in the field of complex numbers  $\mathbb{C}$ .

**Definition 1.1.** Let  $x$  be in  $\mathbb{C}^n$  and  $A$  a tensor as above. Adopting the notation introduced in [10,12], we define  $Ax^{m-1}$  to be the vector in  $\mathbb{C}^n$  whose  $j$ th coordinate is the scalar

$$(Ax^{m-1})_j = \sum_{i_2=1}^n \dots \sum_{i_m=1}^n a_{ji_2 \dots i_m} x_{i_2} \dots x_{i_m}. \quad (1.1)$$

<sup>\*</sup> Corresponding author.

E-mail addresses: [dustin@math.berkeley.edu](mailto:dustin@math.berkeley.edu) (D. Cartwright), [bernd@math.berkeley.edu](mailto:bernd@math.berkeley.edu) (B. Sturmfels).

If  $\lambda$  is a complex number and  $x \in \mathbb{C}^n$  a non-zero vector such that  $Ax^{m-1} = \lambda x$ , then  $\lambda$  is an *eigenvalue* of  $A$  and  $x$  is an *eigenvector* of  $A$ . We will refer to the pair of  $\lambda$  and  $x$  as an *eigenpair*. Two eigenpairs  $(\lambda, x)$  and  $(\lambda', x')$  of the same tensor  $A$  are considered to be *equivalent* if there exists a complex number  $t \neq 0$  such that  $t^{m-2}\lambda = \lambda'$  and  $tx = x'$ .

If  $m = 2$  then  $Ax^1$  is the ordinary matrix-vector product, and Definition 1.1 recovers the familiar eigenvalues and eigenvectors of a square matrix  $A$ . In that case, our notion of equivalence amounts to rescaling the eigenvector, but the eigenvalue is uniquely determined.

For  $m \geq 3$ , Qi [12] normalizes the eigenvectors  $x$  of  $A$  by additionally requiring  $x \cdot x = 1$ . When  $x \cdot x$  is non-zero, this has the effect of choosing two distinguished representatives, related by  $\lambda' = (-1)^m \lambda$  and  $x' = -x$ , from each equivalence class. In particular, when  $m$  is even, the eigenvalue is uniquely determined, and when  $m$  is odd, it is determined up to sign. However, since equivalence classes with  $x \cdot x = 0$  are not allowed by Qi's normalization, his definition does not strictly generalize the classical eigenvalues of a matrix. We will call eigenvalues  $\lambda$  with an eigenvector satisfying  $x \cdot x = 1$  *normalized eigenvalues* of the tensor  $A$ .

In Section 6 of [12], Qi considers an alternative normalization by requiring  $x \cdot \bar{x} = 1$ , where  $\bar{x}$  is the complex conjugate. This reduces the equivalence classes from two real dimensions to one real dimension. One can still get an equivalent eigenpair  $(t^{m-2}\lambda, tx)$  for any complex number  $t$  with unit modulus. An alternative definition of eigenvectors and eigenvalues, based on the  $p$ -norm over the real numbers  $\mathbb{R}$ , was introduced by Lim in his variational approach [9].

For most of this paper, we prefer not to choose any normalization whatsoever. Instead, we depend on the notion of equivalence in Definition 1.1 in order to have a finite number of equivalence classes of eigenpairs in the generic case. This equivalence is a generalization of the usual ambiguity of eigenvectors of an  $n \times n$ -matrix  $A$ , which at best, are only unique up to scaling. The following theorem generalizes, from  $m = 2$  to  $m \geq 3$ , the familiar linear algebra fact that an  $n \times n$ -matrix  $A$  has precisely  $n$  eigenvalues over the complex numbers, when counted with their algebraic multiplicities.

**Theorem 1.2.** *If a tensor  $A$  has finitely many equivalence classes of eigenpairs over  $\mathbb{C}$  then their number, counted with multiplicity, is equal to*

$$\frac{(m-1)^n - 1}{m-2} = \sum_{i=0}^{n-1} (m-1)^i.$$

*If the entries of  $A$  are generic, then all multiplicities are equal to 1, so there are exactly  $((m-1)^n - 1)/(m-2)$  equivalence classes of eigenpairs.*

For the case when the tensor order  $m$  is even, the above formula was derived in [10, Theorem 3.4] by means of a detailed analysis of the Macaulay matrix for the multivariate resultant. For arbitrary  $m$ , it was stated as a conjecture in line 2 on page 1228 of [10]. We here present a short proof of this conjecture that is based on techniques from toric geometry [3,5]. The adjective “generic” in the statement of Theorem 1.2 is to be understood in the usual sense of algebraic geometry, that is, there exists a polynomial in the entries of  $A$  such that the conclusion holds for all tensors  $A$  at which that polynomial does not vanish. In particular, it will only fail on a set of instances that has measure zero.

This paper is organized as follows. In Section 2, we prove Theorem 1.2. In Section 3, we investigate the characteristic polynomial. Tensors with characteristic polynomial vanishing identically are interpreted as singular tensors. In Section 4, we relate eigenvalues of tensors to dynamics on projective space. Finally, in Section 5, we specialize to the case of symmetric tensors. We show that, in that case, the set of normalized eigenvalues is always finite.

## 2. Intersections in a weighted projective space

We shall formulate the problem of computing the eigenvalues and eigenvectors of the tensor  $A$  as an intersection problem in the  $n$ -dimensional weighted projective space

$$X = \mathbb{P}(1, 1, \dots, 1, m-2).$$

The textbook definition of  $X$  can be found, for example, in [3, Section 2.0] and [5, page 35]. Points in  $X$  are represented by vectors of complex numbers  $(u_1 : \cdots : u_n : \lambda)$ , not all zero, modulo the rescaling  $(tu_1 : \cdots : tu_n : t^{m-2}\lambda)$  for any non-zero complex number  $t$ . The corresponding algebraic representation of our weighted projective space is  $X = \text{Proj}(R)$ , where  $R = \mathbb{C}[x_1, \dots, x_n, \lambda]$  is the polynomial ring with  $x_1, \dots, x_n$  having degree 1 and  $\lambda$  having degree  $m - 2$ . A third description of  $X$  is that it is the projective toric variety corresponding to the  $n$ -dimensional simplex  $\Delta$  with vertices at  $(m - 2)e_i$  for  $1 \leq i \leq n$  and  $e_{n+1}$ , where the  $e_i$  are the basis vectors in  $\mathbb{R}^{n+1}$ . Our proof of Theorem 1.2 uses this description and basic toric intersection theory as in [5, Chapter 5].

**Remark 2.1.** An alternative approach is to perform the substitution  $\lambda = \tilde{\lambda}^{m-2}$  in the equations  $Ax^{m-1} = \lambda x$ . This makes the system of equations homogeneous of degree  $m - 1$ . Bézout's theorem says that there are generically  $(m - 1)^n$  solutions in the projective space  $\mathbb{P}^n$ . If we remove the trivial solution, this leaves  $(m - 1)^n - 1$  solutions. Assuming that none of these have  $\tilde{\lambda} = 0$ , the orbits formed by multiplying  $\tilde{\lambda}$  by  $e^{2\pi i/(m-2)}$  each yield the same value of  $\lambda = \tilde{\lambda}^{m-2}$  and  $x$ . Thus, there are  $((m - 1)^n - 1)/(m - 2)$  classes.

This proves the count in Theorem 1.2 for generic tensors, but to extend this technique to all tensors with finitely many equivalence classes of eigenpairs is delicate. To get the full strength of Theorem 1.2, such a proof would have to argue that when  $\tilde{\lambda} = 0$ , it has multiplicity  $m - 2$  as a solution to  $Ax^{m-1} = \tilde{\lambda}^{m-2}x$ . In effect, toric geometry conveniently does the bookkeeping in the correspondence between solutions to  $Ax^{m-1} = \lambda x$  and solutions to  $Ax^{m-1} = \tilde{\lambda}^{m-2}x$ .

**Proof of Theorem 1.2.** For  $m = 2$ , the summation in the theorem statement simplifies to  $n$ , which is the number of eigenvalues, with algebraic multiplicity, of an ordinary  $n \times n$ -matrix. Hence we shall now assume that  $m \geq 3$ . For a fixed tensor  $A$ , the  $n$  equations determined by  $Ax^{m-1} = \lambda x$  correspond to  $n$  homogeneous polynomials of degree  $m - 1$  with respect to the grading of  $R$ .

Since  $R$  is generated in degree  $m - 2$ , the line bundle  $\mathcal{O}_X(m - 2)$  is very ample, and the corresponding lattice polytope  $\Delta$  was described above. The affine hull of  $\Delta$  is the hyperplane  $x_1 + \cdots + x_n + (m - 2)\lambda = m - 2$ . The normalized volume of this simplex equals

$$\text{Vol}(\Delta) = (m - 2)^{n-1}. \quad (2.1)$$

The lattice polytope  $\Delta$  is smooth, except at the vertex  $e_{n+1}$ , where it is simplicial with index  $m - 2$ . Therefore, the projective toric variety  $X$  is simplicial, with precisely one isolated singular point corresponding to the vertex  $e_{n+1}$ . By [5, p. 100], the variety  $X$  has a rational Chow ring  $A^*(X)_{\mathbb{Q}}$ , which we can use to compute intersection numbers of divisors on  $X$ .

Our system of equations  $Ax^{m-1} = \lambda x$  consists of  $n$  polynomials of degree  $m - 1$  in  $R$ . Let  $D$  be the divisor class corresponding to  $\mathcal{O}_X(m - 1)$ , and let  $H$  be the very ample divisor class corresponding to  $\mathcal{O}_X(m - 2)$ . The volume formula (2.1) is equivalent to  $H^n = (m - 2)^{n-1}$  in  $A^*(X)_{\mathbb{Q}}$ , and we compute the self-intersection number of  $D$  as the following rational number:

$$D^n = \left( \frac{m - 1}{m - 2} \cdot H \right)^n = \left( \frac{m - 1}{m - 2} \right)^n \cdot (m - 2)^{n-1} = \frac{(m - 1)^n}{m - 2}.$$

From this count we must remove the trivial solution  $\{x = 0\}$  of  $Ax^{m-1} = \lambda x$ . That solution corresponds to the singular point  $e_{n+1}$  on  $X$ . Since that point has index  $m - 2$ , the trivial solution counts for  $1/(m - 2)$  in the intersection computation, as shown in [5, p. 100]. Therefore the number of non-trivial solutions in  $X$  is equal to

$$D^n - \frac{1}{m - 2} = \frac{(m - 1)^n - 1}{m - 2}, \quad (2.2)$$

Therefore, when the tensor  $A$  admits only finitely many equivalence classes of eigenpairs, then their number, counted with multiplicities, coincides with the positive integer in (2.2).

The set of tensors for which the multiplicities are all 1 is a Zariski open set. Therefore, if we show that it is non-empty, then it is dense in the space of all tensors. In Example 2.2 below we exhibit a tensor  $A$  in which the number of equivalence classes, counted without multiplicities, equals the upper bound (2.2). For that  $A$ , each solution  $(x, \lambda)$  has multiplicity 1, so the same holds for generic  $A$ .  $\square$

**Example 2.2.** Let  $A$  be the diagonal tensor of order  $m$  and size  $n$  defined by setting  $A_{ii\dots i} = 1$  and all other entries zero. An eigenpair  $(\lambda, x)$  is a solution to the equations

$$x_i^{m-1} = \lambda x_i \quad \text{for } 1 \leq i \leq n. \quad (2.3)$$

All non-trivial solutions in  $X = \mathbb{P}(1, \dots, 1, m-2)$  satisfy  $\lambda \neq 0$ . By rescaling, we can assume that  $\lambda = 1$ . Fix the root of unity  $\zeta = e^{2\pi i/(m-2)}$ , and let  $S = \{0, \dots, m-3, *\}$ . For any string  $\sigma$  in  $S^n$  other than the all  $*$ s string, we define  $x_i$  to be  $\zeta^{\sigma_i}$  if  $\sigma_i$  is an integer and to be 0 if  $\sigma_i = *$ . This defines  $(m-1)^n - 1$  eigenpairs. However, some of these are equivalent. Incrementing each integer in our string by 1 modulo  $m-2$  corresponds to multiplying our eigenvector by  $\zeta$ . Thus, we have defined  $((m-1)^n - 1)/(m-2)$  equivalence classes of eigenvalues and eigenvectors. These are all equivalence classes of solutions to (2.3).

More generally, suppose that  $A$  is a diagonal tensor with  $A_{ii\dots i}$  equal to some non-zero complex number  $a_i$ . Then the eigenpairs are similarly given by  $\lambda = 1$  and  $x_i = a_i^{-1/(m-2)} \zeta^{\sigma_i}$  or  $x_i = 0$ , as above, where  $a_i^{-1/(m-2)}$  is a fixed root of  $a_i^{-1}$ . In particular, for generic  $a_i$  all  $((m-1)^n - 1)/(m-2)$  eigenpairs will have distinct normalized eigenvalues.  $\square$

From Theorem 1.2, we get the following result guaranteeing the existence of real eigenpairs.

**Corollary 2.3.** If  $A$  has real entries and either  $m$  or  $n$  is odd, then  $A$  has a real eigenpair.

**Proof.** When either  $m$  or  $n$  is odd, then one can check that the integer  $((m-1)^n - 1)/(m-2)$  in Theorem 1.2 is odd. This implies that  $A$  has a real eigenpair by [4, Corollary 13.2].  $\square$

Corollary 2.3 is sharp, in the sense that there exist real tensors with no real eigenpairs whenever both  $m$  and  $n$  are even. We illustrate this in the following example.

**Example 2.4.** Let  $m$  be even,  $n = 2$ , and  $A$  the  $2 \times \dots \times 2$  tensor which is zero except for the entries  $a_{12\dots 2} = 1$  and  $a_{21\dots 1} = -1$ . The eigenpairs of  $A$  are the solutions to the equations:

$$\begin{aligned} x_2^{m-1} &= \lambda x_1, \\ -x_1^{m-1} &= \lambda x_2. \end{aligned}$$

Eliminating  $\lambda$ , we obtain  $x_1^m + x_2^m = 0$ , which has no non-zero real solutions for even  $m$ .

For  $n$  any even integer, let  $B$  be the tensor whose  $n/2$  diagonal  $2 \times \dots \times 2$  blocks are the tensor  $A$  above, and which is zero elsewhere. A non-trivial eigenpair must be a non-trivial eigenpair for at least one of the blocks, and therefore cannot be real. Thus,  $B$  has no real eigenpairs.  $\square$

### 3. Characteristic polynomial and singular tensors

The characteristic polynomial  $\phi_A(\lambda)$  of a generic tensor  $A$  was defined as follows in [10, 12]. Consider the univariate polynomial in  $\lambda$  that arises by eliminating the unknowns  $x_1, \dots, x_n$  from the system of equations  $Ax^{m-1} = \lambda x$  and  $x \cdot x = 1$ . If  $m$  is even then this polynomial equals  $\phi_A(\lambda)$ . If  $m$  is odd then the symmetry of the eigenvalues under  $\lambda \mapsto -\lambda$  means that this polynomial has the form  $\phi_A(\lambda^2)$ , i.e. the characteristic polynomial evaluated at  $\lambda^2$ . Note that the result of the elimination and thus the characteristic polynomial are only unique up to multiplication by  $-1$ . With these definitions, Theorem 1.2 implies the following:

**Corollary 3.1.** *For a generic tensor  $A$ , the characteristic polynomial  $\phi_A(\lambda)$  is irreducible and has degree  $((m-1)^n - 1)/(m-2)$ . Hence this is the number of normalized eigenvalues.*

For any particular tensor  $A$ , the characteristic polynomial  $\phi_A(\lambda)$  is obtained by specializing the entries in the coefficients of the generic characteristic polynomial. Ni et al. [10] expressed  $\phi_A(\lambda)$  as a Macaulay resultant, which implies a formula as a ratio of determinants. For the present work, we used Gröbner-based software to compute the characteristic polynomials of various tensors. It is tempting to surmise that all zeros of the characteristic polynomial  $\phi_A(\lambda)$  are normalized eigenvalues of the tensor  $A$ . This statement is almost true, but not quite. There is some subtle fine print, to be illustrated by Example 3.2 below.

Qi [12, Question 1] asked whether the set of normalized eigenvalues of a tensor is either finite or all of  $\mathbb{C}$ . We answer this question by showing a tensor where neither of these alternatives holds:

**Example 3.2.** Consider the complex  $2 \times 2 \times 2$  tensor  $A$  whose nonzero entries are

$$a_{111} = a_{221} = 1 \quad \text{and} \quad a_{112} = a_{222} = i = \sqrt{-1}.$$

We claim that any complex number other than 0 is a normalized eigenvalue of  $A$ , but 0 is not a normalized eigenvalue. The equations for an eigenvalue and eigenvector of  $A$  are

$$x_1^2 + ix_1x_2 = \lambda x_1 \quad \text{and} \quad x_1x_2 + ix_2^2 = \lambda x_2.$$

For any  $\lambda \neq 0$  we obtain a matching eigenvector that also satisfies  $x \cdot x = 1$  by taking

$$x = \left( \frac{\lambda^2 + 1}{2\lambda}, \frac{\lambda^2 - 1}{2i\lambda} \right).$$

Hence  $\lambda$  is a normalized eigenvalue. However, if  $\lambda = 0$ , then an eigenvector must satisfy

$$x_1^2 + ix_1x_2 = 0 \quad \text{and} \quad x_1x_2 + ix_2^2 = 0.$$

These imply that  $x \cdot x = x_1^2 + x_2^2$  is zero, so  $\lambda = 0$  cannot be a normalized eigenvalue.  $\square$

However, we have the following weaker statement:

**Proposition 3.3.** *The set of normalized eigenvalues of a tensor is either finite or it consists of all complex numbers in the complement of a finite set.*

**Proof.** The set  $\mathcal{E}(A)$  of normalized eigenvalues  $\lambda$  of the tensor  $A$  is defined by the condition

$$\exists x \in \mathbb{C}^n : Ax^{m-1} = \lambda x \quad \text{and} \quad x \cdot x = 1.$$

Hence  $\mathcal{E}(A)$  is the image of an algebraic variety in  $\mathbb{C}^{n+1}$  under the projection  $(x, \lambda) \mapsto \lambda$ . Chevalley's Theorem states that the image of an algebraic variety under a polynomial map is constructible, that is, defined by a Boolean combination of polynomial equations and inequations [2, Corollary 2 in Section 5.6]. We conclude that the set  $\mathcal{E}(A)$  of normalized eigenvalues is a constructible subset of  $\mathbb{C}$ . This means that  $\mathcal{E}(A)$  is either a finite set or the complement of a finite set.  $\square$

The relationship between the normalized eigenvalues and the characteristic polynomial is summarized in the following proposition.

**Proposition 3.4.** *For a tensor  $A$ , each of the following conditions implies the next:*

- (1) *The set  $\mathcal{E}(A)$  of all normalized eigenvalues consists of all complex numbers.*
- (2) *The set  $\mathcal{E}(A)$  is infinite.*
- (3) *The characteristic polynomial  $\phi_A(\lambda)$  vanishes identically.*

**Proof.** Clearly, (1) implies (2). By the projection argument in the proof above, the zero set in  $\mathbb{C}$  of the characteristic polynomial  $\phi_A(\lambda)$  contains the set  $\mathcal{E}(A)$ . Hence (2) implies (3).  $\square$

From Example 3.2, we see that (2) does not necessarily imply (1), and in Example 3.5, we shall see that (3) does not imply (2). Qi defines a singular tensor to be one for which the first statement of Proposition 3.4 holds. However, we suggest that the last condition is a better definition: a *singular tensor* is a tensor such that  $\phi_A(\lambda)$  vanishes identically. This more inclusive definition has the advantage that the limit of singular tensors is again singular. In particular, the set of all singular tensors is a closed subvariety in the  $n^m$ -dimensional space of tensors  $\mathbb{C}^{n \times \dots \times n}$ . Its defining polynomial equations are the coefficients of the characteristic polynomial  $\phi_A(\lambda)$  where  $A$  is the tensor whose entries  $a_{i_1 \dots i_n}$  are indeterminates.

**Example 3.5.** Let  $m = 3$  and  $n = 2$ . Here  $A = (a_{ijk})$  is a general tensor of format  $2 \times 2 \times 2$ . The characteristic polynomial  $\phi_A$  is obtained by eliminating  $x_1$  and  $x_2$  from the ideal

$$\langle a_{111}x_1^2 + (a_{112} + a_{121})x_1x_2 + a_{122}x_2^2 - \lambda x_1, a_{211}x_1^2 + (a_{212} + a_{221})x_1x_2 + a_{222}x_2^2 - \lambda x_2, x_1^2 + x_2^2 - 1 \rangle.$$

We find that  $\phi_A$  has degree 3, as predicted by Theorem 1.2. Namely, the elimination yields

$$\phi_A(\lambda^2) = C_2\lambda^6 + C_4\lambda^4 + C_6\lambda^2 + C_8,$$

where  $C_i$  is a certain homogeneous polynomial of degree  $i$  in the eight unknowns  $a_{ijk}$ . The set of singular  $2 \times 2 \times 2$ -tensors is thus the variety in the projective space  $\mathbb{P}^7 = \mathbb{P}(\mathbb{C}^{2 \times 2 \times 2})$  given by

$$\langle C_2, C_4, C_6, C_8 \rangle \subset \mathbb{C}[a_{111}, a_{112}, \dots, a_{222}]. \quad (3.1)$$

This is an irreducible variety of codimension 2 and degree 4, but the ideal (3.1) is not prime. The constant coefficient  $C_8$  is the square of a quartic. That quartic is the *Sylvester resultant*

$$\text{Res}_x(Ax^2) = \det \begin{pmatrix} a_{111} & a_{112} + a_{121} & a_{122} & 0 \\ 0 & a_{111} & a_{112} + a_{121} & a_{122} \\ a_{211} & a_{212} + a_{221} & a_{222} & 0 \\ 0 & a_{211} & a_{212} + a_{221} & a_{222} \end{pmatrix}.$$

The leading coefficient of the characteristic polynomial is the negative of a sum of squares

$$C_2 = -(-a_{111} + a_{122} + a_{212} + a_{221})^2 - (a_{112} + a_{121} + a_{211} - a_{222})^2.$$

This indicates that among singular  $2 \times 2 \times 2$ -tensors those with real entries are scarce. Indeed, the real variety of (3.1) is the union of two linear spaces of codimension 4, with defining ideal

$$\langle a_{122}, a_{211}, a_{112} + a_{121} - a_{222}, a_{212} + a_{221} - a_{111} \rangle \cap \langle a_{111} - a_{122}, a_{211} - a_{222}, a_{112} + a_{121}, a_{212} + a_{221} \rangle.$$

This explains why the singular tensor in Example 3.2 had to have a non-real coordinate.

We now look more closely at the real singular tensors defined by the second ideal in this intersection. These are tensors  $A$  for which  $Ax^2 = (a_{111}(x^2 + y^2), a_{211}(x^2 + y^2))$ . It is easy to see that, so long as  $a_{111}$  and  $a_{211}$  are not both zero, the only normalized eigenvector is

$$\left( \frac{a_{111}}{\sqrt{a_{111}^2 + a_{211}^2}}, \frac{a_{211}}{\sqrt{a_{111}^2 + a_{211}^2}} \right), \text{ which has eigenvalue } \lambda = \sqrt{a_{111}^2 + a_{211}^2}.$$

In particular, the number of eigenvalues of such a tensor must be finite. This example shows that (3) does not imply (2) in Proposition 3.4.  $\square$

The reader will not have failed to notice that the notion of “singular” used here (and in [12]) is more restrictive than the one familiar from the classical case  $m = 2$ . Indeed, a matrix is singular if it has  $\lambda = 0$  as an eigenvalue, or, equivalently, if the constant term of the characteristic polynomial vanishes. That constant term is a power of the resultant  $\text{Res}_x(Ax^{m-1})$ , and its vanishing means that the homogeneous equations  $Ax^{m-1} = 0$  have a non-trivial solution  $x \in \mathbb{P}^{n-1}$ . This holds when the tensor  $A$  is singular but not conversely.

Yet another possible notion of singularity for a tensor  $A$  arises from its *hyperdeterminant*  $\text{HDet}(A)$ , as defined in [6]. For example, the hyperdeterminant of a  $2 \times 2 \times 2$ -tensor equals

$$\begin{aligned} \text{HDet}(A) = & a_{122}^2 a_{211}^2 + a_{121}^2 a_{212}^2 + a_{112}^2 a_{221}^2 + a_{111}^2 a_{222}^2 \\ & - 2a_{121} a_{122} a_{211} a_{212} - 2a_{112} a_{122} a_{211} a_{221} - 2a_{112} a_{121} a_{212} a_{221} - 2a_{111} a_{122} a_{211} a_{222} \\ & - 2a_{111} a_{121} a_{212} a_{222} - 2a_{111} a_{112} a_{221} a_{222} + 4a_{111} a_{122} a_{212} a_{221} + 4a_{112} a_{121} a_{211} a_{222}. \end{aligned}$$

The hyperdeterminant vanishes if the hypersurface defined by the multilinear form associated with  $A$  has a singular point in  $(\mathbb{P}^{n-1})^m$ . This property is unrelated to the coefficients of the characteristic polynomial  $\phi_A(\lambda)$ . In particular,  $\text{HDet}(A)$  and  $\text{Res}_x(Ax^{m-1})$  have no common factors.

For an example take the  $2 \times 2 \times 2$ -tensor  $A$  all of whose entries are  $a_{ijk} = 1$  is not singular but  $\text{HDet}(A) = 0$ . On the other hand, the following tensor  $A$  is singular but has  $\text{HDet}(A) = -1$ :

$$\begin{aligned} a_{111} &= -1, & a_{112} &= 0, & a_{121} &= 0, & a_{122} &= -1, \\ a_{211} &= 1, & a_{212} &= -1, & a_{221} &= 0, & a_{222} &= 1 - i. \end{aligned}$$

This highlights the distinction between our setting here and that in [9, Proposition 2].

#### 4. Dynamics on projective space

The purpose of this short section is to point out a connection to dynamical systems. Dynamics on projective space is a well-established field of mathematics [1,8]. We believe that the interpretation of eigenpairs of tensors in terms of fixed points on  $\mathbb{P}^{n-1}$  could be of interest to applied mathematicians as a tool for modeling and numerical computations.

We consider the map  $\psi_A$  defined by the formula  $\psi_A(x) = Ax^{m-1}$ . This is a rational map from complex projective space  $\mathbb{P}^{n-1}$  to itself. The fixed points of the map  $\psi_A: \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}$  are exactly the eigenvectors of the tensor  $A$  with non-zero eigenvalue, and the base locus of  $\psi_A$  is the set of eigenvectors with eigenvalue zero. In particular, the map  $\psi_A$  is defined everywhere if and only if 0 is not an eigenvalue of  $A$ .

Note that every such rational map arises from some tensor  $A$ , but the tensor is not unique. Indeed,  $A$  has  $n^m$  entries while the map is determined by  $n$  polynomials, which have only  $n \binom{n+m-2}{m-1}$  distinct coefficients. For instance, in Example 3.5, with  $m = 3$ ,  $n = 2$ , the eight entries of the tensor translate into six distinct coefficients of the two binary quadrics that specify the self-map of the projective line  $\psi_A: \mathbb{P}^1 \dashrightarrow \mathbb{P}^1$ .

It is instructive to revisit the classical case  $m = 2$ , where  $\psi_A: \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}$  is a linear map. The condition that every eigenvalue of the matrix  $A$  is zero is equivalent to saying that  $A$  is *nilpotent*, that is, some matrix power of  $A$  is zero. Geometrically, this means that some iterate of the rational map  $\psi_A$  is defined nowhere in projective space  $\mathbb{P}^{n-1}$ . We use the same definition for tensors:  $A$  is *nilpotent* if some iterate of  $\psi_A$  is nowhere defined.

**Proposition 4.1.** *If the tensor  $A$  is nilpotent then 0 is the only eigenvalue of  $A$ . The converse is not true: there exist tensors with only eigenvalue 0 that are not nilpotent.*

**Proof.** Suppose  $\lambda \neq 0$  is an eigenvalue and  $x \in \mathbb{C}^n \setminus \{0\}$  a corresponding eigenvector. Then  $x$  represents a point in  $\mathbb{P}^{n-1}$  that is fixed by  $\psi_A$ . Hence it is fixed by every iterate  $\psi_A^{(r)}$  of  $\psi_A$ . In particular,  $\psi_A^{(r)}$  is defined at (an open neighborhood of)  $x \in \mathbb{P}^{n-1}$ , and  $A$  is not nilpotent.



Let  $A$  be the  $2 \times 2 \times 2$ -tensor with  $a_{111} = a_{211} = a_{212} = 1$  and the other five entries zero. The eigenpairs of  $A$  are the solutions to  $x_1^2 = \lambda x_1$  and  $x_1^2 + x_1 x_2 = \lambda x_2$ . Up to equivalence, the only eigenpair is  $x = (0, 1)$  and  $\lambda = 0$ . However, the self-map  $\psi_A$  on  $\mathbb{P}^1$  is dominant. To see this, note that  $\psi_A$  acts by translation on the affine line  $\mathbb{A}^1 = \{x_1 \neq 0\}$  because  $(x_1^2 : x_1^2 + x_1 x_2) = (x_1 : x_1 + x_2)$ . All iterates of  $\psi_A$  are defined on  $\mathbb{A}^1$ , i.e. there are no base points with  $x_1 \neq 0$ , and hence  $A$  is not nilpotent.  $\square$

The example in the previous proof works because the two binary quadrics in  $Ax^{m-1}$  have  $x_1$  as a common factor. Indeed, whenever  $n = 2$ , an eigenvector  $x$  has eigenvalue zero if and only if  $x$  is a solution to a common linear factor of the two binary forms of  $Ax^{m-1}$ .

However, for  $n \geq 3$ , this is no longer true. Work in dynamics by Ivashkovic [8, Theorem 1] implies that, for  $n = 3$ , one can construct tensors  $A$  such that zero is the only eigenvalue, the polynomials in  $Ax^{m-1}$  have no common factors, but  $A$  is not nilpotent.

**Example 4.2.** This example is taken from [8, Example 4.1]. Let  $m = n = 3$  and take  $A$  to be any tensor whose corresponding map is the Cremona transformation

$$\psi_A: \mathbb{P}^2 \dashrightarrow \mathbb{P}^2 : (x_1, x_2, x_3) \mapsto (x_1 x_2 : x_1 x_3 : 2x_2 x_3).$$

This map has no fixed points, but it is not nilpotent. The base locus of  $\psi_A$  consists of the three points  $(1, 0, 0)$ ,  $(0, 1, 0)$ , and  $(0, 0, 1)$ , and, up to scaling, these are the only eigenvectors of  $A$ , all with eigenvalue 0.  $\square$

## 5. Symmetric tensors

Of particular interest in numerical multilinear algebra is the situation when the tensor  $A$  is symmetric and has real entries. Here  $A$  being *symmetric* means that the entries  $a_{i_1 i_2 \dots i_m}$  are invariant under permuting the  $m$  indices  $i_1, i_2, \dots, i_m$ . Each symmetric  $n \times \dots \times n$  tensor  $A$  of order  $m$  corresponds to a unique homogeneous polynomial  $f(x)$  of degree  $m$  in  $n$  unknowns. The symmetric case is of interest because of the following result due to Qi [11, Theorem 5(a)]:

**Theorem 5.1.** *A real polynomial  $f(x)$  of even degree  $m$  is positive semidefinite if and only if every real eigenpair of the corresponding symmetric tensor  $A$  has non-negative eigenvalue.*

Theorem 5.1 will be illustrated in Example 5.9 below. In the notation of [11], the homogeneous polynomial  $f(x)$  can be identified with the symmetric tensor  $A$  by the equation

$$Ax^m = mf(x), \quad (5.1)$$

where  $Ax^m$  is defined to be

$$\sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_m=1}^n a_{i_1 \dots i_m} x_{i_1} x_{i_2} \dots x_{i_m} = x \cdot Ax^{m-1}.$$

By differentiating both sides of (5.1), it follows that  $Ax^{m-1} = \nabla f(x)$ , which says that the coordinates of the gradient of  $f(x)$  are precisely the entries of  $Ax^{m-1}$ . The gradient  $\nabla f(x)$  vanishes at a point  $x$  in  $\mathbb{P}^{n-1}$  if and only if  $x$  is a singular point of the hypersurface in  $\mathbb{P}^{n-1}$  defined by the polynomial  $f(x)$ . This implies:

**Proposition 5.2.** *The singular points of the projective hypersurface  $\{x \in \mathbb{P}^{n-1} : f(x) = 0\}$  are precisely the eigenvectors of the corresponding symmetric tensor  $A$  which have eigenvalue 0.*

The other eigenvectors of  $A$  can also be characterized in terms of the polynomial  $f(x)$ .



**Proposition 5.3.** Fix a non-zero  $\lambda$  and suppose  $m \geq 3$ . Then  $x \in \mathbb{C}^n$  is a normalized eigenvector with eigenvalue  $\lambda$  if and only if  $x$  is a singular point of the affine hypersurface defined by the polynomial

$$f(x) - \frac{\lambda}{2} x \cdot x - \left( \frac{1}{m} - \frac{1}{2} \right) \lambda. \quad (5.2)$$

**Proof.** The gradient of (5.2) is  $\nabla f - \lambda x = Ax^{m-1} - \lambda x$ , so every singular point  $x$  is an eigenvector with eigenvalue  $\lambda$ . Furthermore, if we substitute  $f(x) = \frac{1}{m} x \cdot \nabla f = \frac{\lambda}{m} x \cdot x$  into (5.2), then we obtain  $x \cdot x = 1$ . This argument is reversible: if  $x$  is a normalized eigenvector of  $A$  then  $x \cdot x = 1$  and  $\nabla f(x) = \lambda x$ , and this implies that (5.2) and its derivatives vanish.  $\square$

**Corollary 5.4.** The characteristic polynomial  $\phi_A(\lambda)$  is a factor of the discriminant of (5.2).

Here we mean the classical multivariate discriminant [6] of an inhomogeneous polynomial of degree  $m$  in  $n$  variables  $x$  evaluated at (5.2), where  $\lambda$  is regarded as a parameter. Besides the characteristic polynomial  $\phi_A(\lambda)$ , this discriminant may contain other irreducible factors.

**Example 5.5** (Discriminantal representation of the characteristic polynomial of a symmetric tensor). If  $n = 2$  and  $m = 3$  then the discriminant at the bivariate cubic in (5.2) equals  $\lambda^4 \cdot \phi_A(\lambda)$ . If  $n = 2$  and  $m = 4$  then we evaluate the discriminant of the ternary quartic using Sylvester's formula [6, Section 3.4.D]. The output has the discriminant of binary quartic as an extraneous factor:

$$\text{Discriminant of (5.2)} = (\phi_A(\lambda))^2 \cdot \lambda^9 \cdot \text{Discriminant of } f(x)$$

It would be interesting to determine the analogous factorization for arbitrary  $m$  and  $n$ .  $\square$

The subject of this paper is the number of normalized eigenvalues of a tensor. In Section 2 we gave an upper bound for that number under the hypothesis that the number is finite. Remarkably, this hypothesis is not needed if we restrict our attention to symmetric tensors.

**Theorem 5.6.** Every symmetric tensor  $A$  has at most  $((m-1)^n - 1)/(m-2)$  distinct normalized eigenvalues. This bound is attained for generic symmetric tensors  $A$ .

**Proof.** It suffices to show that the number of normalized eigenvalues of every symmetric tensor  $A$  is finite. Recall from the proof of Theorem 1.2 that the set of eigenpairs is the intersection of  $n$  linearly equivalent divisors on a weighted projective space. Since these divisors are ample, each connected component of the set of eigenpairs contributes at least one to the intersection number. Therefore, the number of connected components of eigenpairs can be no more than  $((m-1)^n - 1)/(m-2)$ . We conclude that the number of normalized eigenvalues of  $A$ , if finite, must be bounded above by that quantity as well. Finally, Example 2.2 shows that the bound is tight.

We now prove that the number of normalized eigenvalues of a symmetric tensor  $A$  is finite. Let  $S$  be the affine hypersurface in  $\mathbb{C}^n$  defined by the equation  $x_1^2 + \cdots + x_n^2 = 1$ . We claim that a point  $x \in S$  is an eigenvector of  $A$  if and only if  $x$  is a critical point of  $f$  restricted to  $S$ , in which case, the corresponding eigenvalue  $\lambda$  equals  $\frac{1}{m} f(x)$ . By definition, a point  $x \in S$  is a critical point of  $f|_S$  if and only if the gradient  $\nabla(f|_S)$  is zero at  $x$ . The latter condition is equivalent to the gradient  $\nabla f$  being a multiple of  $\nabla(x_1^2 + \cdots + x_n^2 - 1) = 2x$ . This is exactly the definition of an eigenvector. Finally, if  $x \in S$  is a critical point of  $f|_S$ , then  $mf(x) = x \cdot \nabla f(x) = \lambda x \cdot x = \lambda$ , and hence  $\lambda = \frac{1}{m} f(x)$ .

Finally, to prove Theorem 5.6, we note that, by generic smoothness [7, Corollary III.10.7], a polynomial function on a smooth variety has only finitely many critical values. Equivalently, Sard's theorem in differential geometry says that the set of critical values of a differentiable function has measure zero, so, by Proposition 3.3, that set must be finite.  $\square$

We note two subtleties about Theorem 5.6. First, it does not imply that the characteristic polynomial of every symmetric tensor is non-trivial. Second, the result is intrinsically tied to the normalization  $x \cdot x = 1$ . We begin with an example of the first assertion.

**Example 5.7.** Let  $A$  be the symmetric  $2 \times 2 \times 2$  tensor with

$$a_{111} = -2i, \quad a_{112} = a_{121} = a_{211} = 1, \quad a_{122} = a_{212} = a_{221} = 0, \quad a_{222} = 1.$$

Then, up to equivalence, the only eigenvectors are  $(0, 1)$  with eigenvalue 1 and  $(1, i)$  with eigenvalue 0. Note that the second cannot be rescaled to be a normalized eigenvector, so the only normalized eigenvalue is 1. Moreover, the characteristic polynomial of  $A$  is identically zero. The reason is that, for a small perturbation of  $A$ , the perturbation of the eigenvector  $(1, i)$  can take on any given normalized eigenvalue.  $\square$

No analogue of Theorem 5.6 is possible with the alternative normalization of requiring  $x \cdot \bar{x} = 1$ . In this case, each equivalence class yields infinitely many eigenvalues, which nonetheless have the same magnitude. Therefore, for generic tensors, the set of magnitudes of the eigenvalues with  $x \cdot \bar{x} = 1$  will be finite and have cardinality  $((m-1)^n - 1)/(m-2)$  by Theorem 5.6. However, the following example shows that this is not true for all tensors.

**Example 5.8.** Let  $A$  be the symmetric  $3 \times 3 \times 3$  tensor whose non-zero entries are

$$a_{111} = 2 \quad \text{and} \quad a_{122} = a_{212} = a_{221} = a_{133} = a_{313} = a_{331} = 1.$$

The eigenpairs of  $A$  are the solutions to the equations

$$2x_1^2 + x_2^2 + x_3^2 = \lambda x_1,$$

$$2x_1x_2 = \lambda x_2,$$

$$2x_1x_3 = \lambda x_3.$$

For any  $\alpha \in \mathbb{C}$ , the vector  $x = (1, i\alpha, \alpha)$  is an eigenvector with eigenvalue  $\lambda = 2$ . Rescaling,  $x/\sqrt{x \cdot \bar{x}}$  is an eigenvector with unit length and eigenvalue

$$\frac{2}{\sqrt{1+2|\alpha|}}.$$

The magnitude of this eigenvalue can be any real number in the interval  $(0, 2]$ . Note that the family of eigenvectors above all have  $x \cdot x = 1$ , so  $\lambda = 2$  is the only normalized eigenvalue.  $\square$

As promised, we conclude by showing how the eigenvalues of symmetric tensors can be used to decide if  $f$  is *positive semidefinite*, i.e., whether  $f(x) \geq 0$  for all  $x \in \mathbb{R}^n$ .

**Example 5.9.** The *Motzkin polynomial*  $f(x, y, z) = z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2$  is a well-known example of a positive semidefinite polynomial which cannot be written as a sum of squares. Let  $A$  be the corresponding  $3 \times 3 \times 3 \times 3 \times 3$ -tensor. This tensor has 25 eigenvalues, counting multiplicities, six less than our upper bound of 31. Disregarding multiplicities, there are only four distinct eigenvalues. All four are real and they are equal to: 0 (with multiplicity 14),  $3/32$  (with multiplicity 8),  $3/2$  (with multiplicity 2), and 6 (with multiplicity 1). By [11, Theorem 5(a)], this confirms the fact that the Motzkin polynomial  $f$  is positive semidefinite.  $\square$

We applied this criterion for testing positivity to a range of homogeneous polynomials in three variables of modest degree, and we found it to be competitive with other methods.

## Acknowledgments

We thank Tamara Kolda for inspiring this project, with a question she asked us at the *Berkeley Optimization Day* on March 6, 2010. Both authors were supported in part by the National Science Foundation (DMS-0456960 and DMS-0757207).

## References

- [1] E. Bedford, M. Jonsson, Dynamics of regular polynomial endomorphisms of  $\mathbb{C}C^k$ , *Amer. J. Math.* 122 (2000) 153–212.
- [2] D.A. Cox, J.B. Little, D. O’Shea, *Ideals, Varieties, and Algorithms*, third ed., Springer, New York, NY, 2007.
- [3] D.A. Cox, J.B. Little, H. Schenck, *Toric varieties*, graduate studies in mathematics, Am. Math. Soc. 124 (2011).
- [4] W. Fulton, *Intersection Theory*, second ed., Springer, New York, NY, 1998.
- [5] W. Fulton, Introduction to toric varieties, in: William H. Roever (Ed.), in: *Lectures in Geometry*, Annals of Mathematics Studies, vol. 131, Princeton University Press, Princeton, NJ, 1993.
- [6] I. Gel’fand, M. Kapranov, A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.
- [7] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics, vol. 22, Springer, New York, 1977.
- [8] S. Ivashkovich: On fixed points of rational self-maps of the complex projective plane, *Indiana Univ. Math. J.*, in press. Available from: arXiv:0911.5084.
- [9] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach, in: *Proc. IEEE Int. Workshop on Comput. Advances in Multi-Sensor Adaptive Processing (CAMSAP 2005)*, pp. 129–132.
- [10] G. Ni, L. Qi, F. Wang, Y. Wang, The degree of the E-characteristic polynomial of an even order tensor, *J. Math. Anal. Appl.* 329 (2007) 1218–1229.
- [11] L. Qi, Eigenvalues of a real supersymmetric tensor, *J. Symbolic Comput.* 40 (2005) 1302–1324.
- [12] L. Qi, Eigenvalues and invariants of tensors, *J. Math. Anal. Appl.* 325 (2007) 1363–1377.