Tensors and Their Eigenvectors

Bernd Sturmfels

My AMS invited address at the SIAM Annual Meeting July 11–15 in Boston discusses the extension of eigenvectors and singular vectors from matrices to higher order tensors.

A real $n \times n$-matrix usually has n independent *eigenvectors* over the complex numbers. When the matrix is symmetric, its eigenvectors have real coordinates and are *orthogonal*. For a rectangular matrix, one considers pairs of *singular vectors*, one on the left and one on the right. The number of these pairs is equal to the smaller of the two matrix dimensions.

Engineers and scientists spice up their linear algebra toolbox with a pinch of algebraic geometry.

Eigenvectors and singular vectors are familiar from linear algebra, where they are taught in concert with *eigenvalues* and *singular values*. Linear algebra is the foundation of applied mathematics and scientific computing. Specifically, the concept of eigenvectors and numerical algorithms for computing them became a key technology during the twentieth century.

However, in our day and age of Big Data, the role of matrices is increasingly often played by *tensors*, that is, multidimensional arrays of numbers. Principal component analysis tells us that eigenvectors of matrices point to directions in which the data is most spread. One hopes to identify similar features in higher-dimensional data. This has encouraged engineers and scientists to spice up their linear algebra toolbox with a pinch of algebraic geometry.

The spectral theory of tensors is the theme of the AMS Invited Address at the SIAM Annual Meeting, held in Boston on July 11–15, 2016. This theory was pioneered around 2005 by Lek-Heng Lim and Liqun Qi. The aim of our introduction is to generalize familiar notions, such as rank, eigenvectors and singular vectors, from matrices to tensors. Specifically, we address the following questions. The answers are provided in Examples 5 and 10 respectively.

Bernd Sturmfels is professor of mathematics, statistics and computer science at the University of California at Berkeley and is past vice president of the American Mathematical Society. His email address is bernd@berkeley.edu.

Question 1. *How many eigenvectors does a* $3 \times 3 \times 3$-*tensor have?*

Question 2. *How many triples of singular vectors does a* $3 \times 3 \times 3$-*tensor have?*

A *tensor* is a d-dimensional array $T = (t_{i_1 i_2 \cdots i_d})$. Tensors of format $n_1 \times n_2 \times \cdots \times n_d$ form a space of dimension $n_1 n_2 \cdots n_d$. For $d = 1, 2$ we get vectors and matrices. A tensor has *rank* 1 if it is the outer product of d vectors, written $T = \mathbf{u} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{w}$ or, in coordinates,

$$t_{i_1 i_2 \cdots i_d} = u_{i_1} v_{i_2} \cdots w_{i_d}.$$

The problem of *tensor decomposition* concerns expressing T as a sum of rank 1 tensors, using as few summands as possible. That minimal number of summands needed is the *rank* of T.

An $n \times n \times \cdots \times n$-tensor $T = (t_{i_1 i_2 \cdots i_d})$ is *symmetric* if it is unchanged under permuting the indices. The space $\mathrm{Sym}_d(\mathbb{R}^n)$ of such symmetric tensors has dimension $\binom{n+d-1}{d}$. It is identified with the space of homogeneous polynomials of degree d in n variables, written as

$$T = \sum_{i_1, \ldots, i_d = 1}^{n} t_{i_1 i_2 \cdots i_d} \cdot x_{i_1} x_{i_2} \cdots x_{i_d}.$$

Example 3. A tensor T of format $3 \times 3 \times 3$ has 27 entries. If T is symmetric, then it has 10 distinct entries, one for each coefficient of the associated cubic polynomial in three variables. This polynomial defines a cubic curve in the projective plane $\mathbb{P}^2$, as indicated in Figure 1.

Symmetric tensor decomposition writes a polynomial as a sum of powers of linear forms:

$$(1) \quad T = \sum_{j=1}^{r} \lambda_j \mathbf{v}_j^{\otimes d} = \sum_{j=1}^{r} \lambda_j (v_{1j} x_1 + v_{2j} x_2 + \cdots + v_{nj} x_n)^d.$$
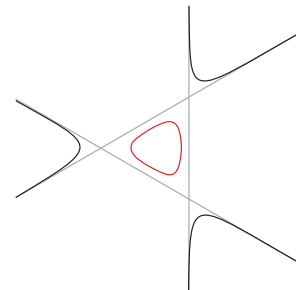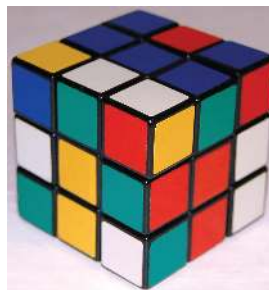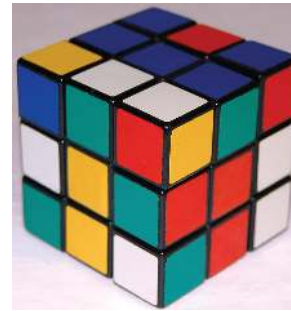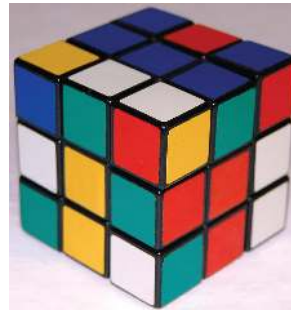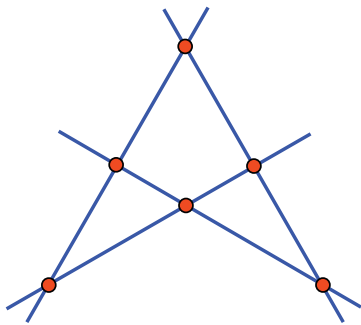
Figure 1. A symmetric $3 \times 3 \times 3$-**tensor represents a cubic curve (drawn by Cynthia Vinzant) in the projective plane.**

Figure 2. The polynomial $T = xyz(x + y + z)$ **represents a symmetric** $3{\times}3{\times}3{\times}3$**-tensor.**

The *gradient of T* defines a map $\nabla T : \mathbb{R}^n \to \mathbb{R}^n$. A vector $\mathbf{v} \in \mathbb{R}^n$ is an *eigenvector* of T if

$$(\nabla T)(\mathbf{v}) = \lambda \cdot \mathbf{v} \qquad \text{for some } \lambda \in \mathbb{R}.$$

Eigenvectors of tensors arise naturally in optimization. Consider the problem of maximizing a polynomial function T over the unit sphere in $\mathbb{R}^n$. If λ denotes a Lagrange multiplier, then one sees that the eigenvectors of T are the critical points of this optimization problem.

Algebraic geometers find it convenient to replace the unit sphere in $\mathbb{R}^n$ by the projective space $\mathbb{P}^{n-1}$. The gradient map is then a rational map from this projective space to itself:

$$\nabla T : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}.$$

The eigenvectors of T are *fixed points* ($\lambda \neq 0$) and *base points* ($\lambda = 0$) of ∇T. Thus the spectral theory of tensors is closely related to the study of dynamical systems on $\mathbb{P}^{n-1}$.

In the matrix case ($d = 2$), the linear map ∇T is the gradient of the quadratic form

$$T = \sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} x_i x_j.$$

By the Spectral Theorem, T has a real decomposition (1) with $d = 2$. Here r is the rank, the λ_j are the eigenvalues of T, and the eigenvectors $\mathbf{v}_j = (v_{1j}, v_{2j}, \ldots, v_{nj})$ are orthonormal. We can compute this by *power iteration*, namely, by applying ∇T until a fixed point is reached.

For $d \geq 3$, one can still use the power iteration to compute eigenvectors of T. However, the eigenvectors are usually not the vectors $\mathbf{v}_i$ in the low-rank decomposition (1). One exception arises when the symmetric tensor is *odeco*, or orthogonally decomposable [4]. This means that T has the form (1), where $r = n$ and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthogonal basis of $\mathbb{R}^n$. These basis vectors are the attractors of the dynamical system ∇T, provided $\lambda_j > 0$.

Theorem 4 ([2])**.** *The number of complex eigenvectors of a general tensor $T \in \mathrm{Sym}_d(\mathbb{R}^n)$ is*

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

Example 5. Let $n = d = 3$. The Fermat cubic $T = x^3 + y^3 + z^3$ is an odeco tensor. Its gradient map squares each coordinate: $\nabla T : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$, $(x : y : z) \mapsto (x^2 : y^2 : z^2)$. This dynamical system has seven fixed points, of which only the first three are attractors:

$$(1 : 0 : 0), (0 : 1 : 0), (0 : 0 : 1),$$
$$(1 : 1 : 0), (1 : 0 : 1), (0 : 1 : 1), (1 : 1 : 1).$$

We conclude that T has seven eigenvectors, and the same holds for $3 \times 3 \times 3$-tensors in general.

It is not known whether all eigenvectors can be real. This holds for $n = 3$ by [1, Theorem 6.1]. Namely, if T is a product of linear forms defining d lines in $\mathbb{P}^2$, then the $\binom{d}{2}$ vertices of the line arrangement are base points of ∇T, and each of the $\binom{d}{2} + 1$ regions contains one fixed point. This accounts for all $1 + (d-1) + (d-1)^2$ eigenvectors, which are therefore real.

Example 6. Let $d = 4$ and fix the product of linear forms $T = xyz(x + y + z)$. Its curve in $\mathbb{P}^2$ is an arrangement of four lines, as shown in Figure 2. This quartic represents a symmetric $3{\times}3{\times}3{\times}3$-tensor. All $13 = 6 + 7$ eigenvectors of this tensor are real. The 6 vertices of the arrangement are the base points of ∇T. Each of the 7 regions contains one fixed point.

For special tensors T, two of the eigenvectors in Theorem 4 may coincide. This corresponds to vanishing of the *eigendiscriminant*, which is a big polynomial in the coefficients $t_{i_1 i_2 \cdots i_d}$. In the matrix case ($d = 2$), it is the discriminant of the characteristic polynomial of an $n{\times}n$-matrix. For $3{\times}3{\times}3$-tensors the eigendiscriminant has degree 24. In general we have the following:

Theorem 7 ([1, Corollary 4.2])**.** *The eigendiscriminant is an irreducible homogeneous polynomial of degree $n(n-1)(d-1)^{n-1}$ in the coefficients $t_{i_1 i_2 \cdots i_d}$ of the tensor T.*

Singular value decomposition is a central notion in linear algebra and its applications. Consider a rectangular

Bernd Sturmfels enjoying a spherical blackboard at the Institute Mittag-Leffler in Djursholm, Sweden.

matrix $T = (t_{ij})$ of format $n_1 \times n_2$. The *singular values* σ of T satisfy

$$T\mathbf{u} = \sigma\mathbf{v} \quad \text{and} \quad T^t\mathbf{v} = \sigma\mathbf{u},$$

where $\mathbf{u}$ and $\mathbf{v}$ are the corresponding *singular vectors*. Just as with eigenvectors, we can associate to this a dynamical system. Namely, we interpret the matrix as a bilinear form

$$T = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} t_{ij} x_i y_j.$$

The gradient of T defines a rational self-map of a product of two projective spaces:

$$\nabla T : \quad \begin{matrix} \mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1} & \dashrightarrow & \mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1}, \\ (\mathbf{u}, \mathbf{v}) & \mapsto & (T^t\mathbf{v}, T\mathbf{u}). \end{matrix}$$

The *fixed points* of this map are the pairs of singular vectors of T.

Consider now an arbitrary d-dimensional tensor T in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. It corresponds to a multilinear form. The *singular vector tuples* of T are the fixed points of the gradient map

$$\nabla T : \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1} \dashrightarrow \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1}.$$

Example 8. The trilinear form $T = x_1 y_1 z_1 + x_2 y_2 z_2$ defines a $2\times2\times2$-tensor. Its map ∇T is

$$\begin{matrix} \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 & \to & \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1, \\ \big((x_1{:}x_2),(y_1{:}y_2),(z_1{:}z_2)\big) & \mapsto & \big((y_1 z_1{:}y_2 z_2),(x_1 z_1{:}x_2 z_2),(x_1 y_1{:}x_2 y_2)\big). \end{matrix}$$

This map has no base points, but it has six fixed points, namely $((1{:}0),(1{:}0),(1{:}0))$, $((0{:}1),(0{:}1),(0{:}1))$, $((1{:}1),(1{:}1),(1{:}1))$, $((1{:}1),(1{:}-1),(1{:}-1))$, $((1{:}-1),(1{:}1),(1{:}-1))$, and $((1{:}-1),(1{:}-1),(1{:}1))$. These are the triples of singular vectors of the given $2 \times 2 \times 2$-tensor.

Here is an explicit formula for the expected number of singular vector tuples.

Theorem 9 (Friedland and Ottaviani [3]). *For a general $n_1 \times n_2 \times \cdots \times n_d$-tensor T, the number of singular vector tuples (over $\mathbb{C}$) is the coefficient of $z_1^{n_1-1} \cdots z_d^{n_d-1}$ in the polynomial*

$$\prod_{i=1}^{d} \frac{(\hat{z}_i)^{n_i} - z_i^{n_i}}{\hat{z}_i - z_i} \quad \text{where} \quad \hat{z}_i = z_1 + \cdots + z_{i-1} + z_{i+1} + \cdots + z_d.$$

We conclude our excursion into the spectral theory of tensors by answering Question 2.

Example 10. Let $d = 3$ and $n_1 = n_2 = n_3 = 3$. The generating function in Theorem 9 equals

$$(\hat{z_1}^2 + \hat{z_1} z_1 + z_1^2)(\hat{z_2}^2 + \hat{z_2} z_2 + z_2^2)(\hat{z_3}^2 + \hat{z_3} z_3 + z_3^2)$$
$$= \cdots + \mathbf{37} z_1^2 z_2^2 z_3^2 + \cdots.$$

This means that a general $3\times3\times3$-tensor has exactly 37 triples of singular vectors. Likewise, a general $3\times3\times3\times3$-tensor, as illustrated in Figure 2, has 997 quadruples of singular vectors.

References

[1] H. Abo, A. Seigal, and B. Sturmfels, *Eigenconfigurations of tensors*, arXiv:1505.05729.

[2] D. Cartwright and B. Sturmfels, The number of eigenvalues of a tensor, *Linear Algebra Appl.* **438** (2013), 942–952. MR2996375

[3] S. Friedland and G. Ottaviani, The number of singular vector tuples and uniqueness of best rank-one approximation of tensors, *Found. Comput. Math.* **14** (2014), 1209–1242. MR3273677

[4] E. Robeva, Orthogonal decomposition of symmetric tensors, *SIAM Journal on Matrix Analysis and Applications* **37** (2016), 86–102.

Editor's Note: