EIGENVECTORS OF TENSORS

Bernd Sturmfels
University of California at Berkeley

How many eigenvectors does a $3 \times 3 \times 3$-tensor have?
How many singular vector triples does a $3 \times 3 \times 3$-tensor have?
About the cover: Combinatorics, Geometry, and Tensors (see page 613)
Tensors and their rank

A tensor is a $d$-dimensional array of numbers $T = (t_{i_1i_2\ldots i_d})$. For $d = 1$ this is a vector, and for $d = 2$ this is a matrix.

A tensor $T$ of format $n_1 \times n_2 \times \cdots \times n_d$ has $n_1n_2\cdots n_d$ entries.

$T$ has rank 1 if it is the outer product of $d$ vectors $u, v, \ldots, w$:

$$t_{i_1i_2\ldots i_d} = u_{i_1}v_{i_2}\cdots w_{i_d}.$$  

The set of tensors of rank 1 is the Segre variety. We often use $\mathbb{C}$.  

A tensor is a $d$-dimensional array of numbers $T = (t_{i_1i_2\cdots i_d})$. For $d = 1$ this is a vector, and for $d = 2$ this is a matrix.

A tensor $T$ of format $n_1 \times n_2 \times \cdots \times n_d$ has $n_1n_2\cdots n_d$ entries.

$T$ has rank 1 if it is the outer product of $d$ vectors $u, v, \ldots, w$:

$$t_{i_1i_2\cdots i_d} = u_{i_1}v_{i_2}\cdots w_{i_d}.$$ 

The set of tensors of rank 1 is the Segre variety. We often use $\mathbb{C}$.

A tensor has rank $r$ if it is the sum of $r$ tensors of rank 1. (not fewer).

**Tensor decomposition:**

- Express a given tensor as a sum of rank 1 tensors.
- Use as few summands as possible.

Symmetric tensors

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 indices. The vector space of symmetric tensors has dimension \( \binom{n+d-1}{d} \).

\( T \) has rank 1 if it is the \( d \)-fold outer product of a vector \( v \):

\[
t_{i_1 i_2 \cdots i_d} = v_{i_1} v_{i_2} \cdots v_{i_d}.
\]

The set of symmetric tensors of rank 1 is the Veronese variety. A symmetric tensor has rank \( r \) if it is the sum of \( r \) such tensors.
Symmetric tensors

An \( n \times n \times \cdots \times n \)-tensor \( T = (t_{i_1i_2\cdots i_d}) \) is symmetric if it is unchanged under permuting indices. The vector space of symmetric tensors has dimension \( \binom{n+d-1}{d} \).

\( T \) has rank 1 if it is the \( d \)-fold outer product of a vector \( \mathbf{v} \):

\[
t_{i_1i_2\cdots i_d} = v_{i_1}v_{i_2}\cdots v_{i_d}.
\]

The set of symmetric tensors of rank 1 is the Veronese variety.

A symmetric tensor has rank \( r \) if it is the sum of \( r \) such tensors.

**Open Problem** [Comon’s Conjecture] *Is the rank of every symmetric tensor equal to its rank as a general tensor?*

True for \( d = 2 \): every rank 1 decomposition of a symmetric matrix

\[
T = u_1^t\mathbf{v}_1 + u_2^t\mathbf{v}_2 + \cdots + u_r^t\mathbf{v}_r.
\]

transforms into a decomposition into rank 1 symmetric matrices:

\[
T = w_1^t\mathbf{w}_1 + w_2^t\mathbf{w}_2 + \cdots + w_r^t\mathbf{w}_r
\]
Polynomials and their eigenvectors
Symmetric tensors correspond to homogeneous polynomials

\[ T = \sum_{i_1, \ldots, i_d=1}^{n} t_{i_1i_2\ldots i_d} \cdot x_{i_1}x_{i_2} \cdots x_{i_d} \]

The tensor has rank \( r \) if \( T \) is a sum of \( r \) powers of linear forms:

\[ T = \sum_{j=1}^{r} \lambda_j v_j^\otimes d = \sum_{j=1}^{r} \lambda_j (v_{1j}x_1 + v_{2j}x_2 + \cdots + v_{nj}x_n)^d. \]
Polynomials and their eigenvectors

Symmetric tensors correspond to homogeneous polynomials

\[
T = \sum_{i_1, \ldots, i_d = 1}^{n} t_{i_1i_2\ldots i_d} \cdot x_{i_1} x_{i_2} \cdots x_{i_d}
\]

The tensor has rank \( r \) if \( T \) is a sum of \( r \) powers of linear forms:

\[
T = \sum_{j=1}^{r} \lambda_j v_j^\otimes d = \sum_{j=1}^{r} \lambda_j (v_{1j} x_1 + v_{2j} x_2 + \cdots + v_{nj} x_n)^d.
\]

The gradient of \( T \) defines a polynomial map of degree \( d - 1 \):

\[
\nabla T : \mathbb{R}^n \to \mathbb{R}^n.
\]

A vector \( v \in \mathbb{R}^n \) is an eigenvector of the tensor \( T \) if

\[
(\nabla T)(v) = \lambda \cdot v \quad \text{for some } \lambda \in \mathbb{R}.
\]
What is this good for?

Consider the optimization problem of maximizing a homogeneous polynomial $T$ over the unit sphere in $\mathbb{R}^n$.

Lagrange multipliers lead to the equations

$$(\nabla T)(v) = \lambda \cdot v$$

for some $\lambda \in \mathbb{R}$.

**Fact:** The critical points are the eigenvectors of $T$.

It is convenient to replace $\mathbb{R}^n$ with projective space $\mathbb{P}^{n-1}$.

*In English:* parallel vectors are regarded as the same.
What is this good for?

Consider the optimization problem of maximizing a homogeneous polynomial $T$ over the unit sphere in $\mathbb{R}^n$.

Lagrange multipliers lead to the equations

$$(\nabla T)(v) = \lambda \cdot v$$

for some $\lambda \in \mathbb{R}$.

**Fact:** The critical points are the eigenvectors of $T$.

It is convenient to replace $\mathbb{R}^n$ with projective space $\mathbb{P}^{n-1}$.

*In English: parallel vectors are regarded as the same.*

Eigenvectors of $T$ are fixed points of $\nabla T : \mathbb{P}^{n-1} \rightarrow \mathbb{P}^{n-1}$.

**Fact:** These are nonlinear dynamical systems on $\mathbb{P}^{n-1}$.

[Lim, Ng, Qi: *The spectral theory of tensors and its applications*, 2013]
Linear maps

Real symmetric $n \times n$-matrices $(t_{ij})$ correspond to quadratic forms

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

By the Spectral Theorem, there exists a real decomposition

$$T = \sum_{j=1}^{r} \lambda_j (v_{1j} x_1 + v_{2j} x_2 + \cdots + v_{nj} x_n)^2.$$  

Here $r$ is the rank and the $\lambda_j$ are the eigenvalues of $T$. The eigenvectors $v_j = (v_{1j}, v_{2j}, \ldots, v_{nj})$ are orthonormal.

One can compute this decomposition by the Power Method:

Iterate the linear map $\nabla T : \mathbb{P}^{n-1} \rightarrow \mathbb{P}^{n-1}$

Fixed points of this dynamical system are eigenvectors of $T$. 
Quadratic maps

Symmetric $n \times n \times n$-tensors $(t_{ijk})$ correspond to cubic forms

$$T = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} t_{ijk} x_i x_j x_k$$

We are interested in low rank decompositions

$$T = \sum_{j=1}^{r} \lambda_j (v_{1j} x_1 + v_{2j} x_2 + \cdots + v_{nj} x_n)^3.$$

One idea to find this decomposition is the Tensor Power Method:

Iterate the quadratic map $\nabla T : \mathbb{P}^{n-1} \rightarrow \mathbb{P}^{n-1}$

Fixed points of this dynamical system are eigenvectors of $T$.

**Bad News:** The eigenvectors are usually not the vectors $v_i$ in the low rank decomposition ... unless the tensor is *odeco*.
Odeco tensors

A symmetric tensor $T$ is \textit{odeco} (= orthogonally decomposable) if

$$T = \sum_{j=1}^{n} \lambda_j v_j \otimes d = \sum_{j=1}^{n} \lambda_j (v_{1j}x_1 + \cdots + v_{nj}x_n)^d,$$

where $\{v_1, v_2, \ldots, v_n\}$ is an orthogonal basis of $\mathbb{R}^n$.

The tensor power method works well for odeco tensors:

\textbf{Theorem}

\textit{If $\lambda_j > 0$ then the $v_i$ are precisely the robust eigenvectors of $T$.}

[Anandkumar, Ge, Hsu, Kakade, Telgarsky: \textit{Tensor decompositions for learning latent variable models}, J. Machine Learning Research, 2014]

[Kolda: \textit{Symmetric orthogonal tensor decomposition is trivial}, 2015]

The set of odeco tensors is a very nice \textbf{variety} of dimension $\binom{n+1}{2}$.

[Robeva: \textit{Orthogonal decomposition of symmetric tensors}, 2015]
**Associativity**

**Fact:** Every $n \times n \times n$-tensor $T$ defines an algebra structure on $\mathbb{R}^n$.

**Example:** Fix $\mathbb{R}^2$ with basis $\{a, b\}$. A $2 \times 2 \times 2$-tensor $T = (t_{ijk})$ defines

\[
\begin{align*}
a \star a &= t_{000} a + t_{001} b \\
b \star a &= t_{100} a + t_{101} b \\
a \star b &= t_{010} a + t_{011} b \\
b \star b &= t_{110} a + t_{111} b
\end{align*}
\]

This algebra is generally not associative:

\[
\begin{align*}
b \star (a \star a) &= (t_{000} t_{100} + t_{001} t_{110}) a + (t_{000} t_{101} + t_{001} t_{111}) b \\
(b \star a) \star a &= (t_{000} t_{100} + t_{101} t_{100}) a + (t_{100} t_{001} + t_{101} t_{101}) b
\end{align*}
\]

Suppose that $T$ is a symmetric tensor, corresponding to a binary cubic

\[
t_{000} x^3 + (t_{001} + t_{010} + t_{100}) x^2 y + (t_{011} + t_{101} + t_{110}) x y^2 + t_{111} y^3
\]

\[
= t_{000} x^3 + 3 t_{001} x^2 y + 3 t_{011} x y^2 + t_{111} y^3
\]
**Associativity**

**Fact:** Every $n \times n \times n$-tensor $T$ defines an algebra structure on $\mathbb{R}^n$.

**Example:** Fix $\mathbb{R}^2$ with basis $\{a, b\}$. A $2 \times 2 \times 2$-tensor $T = (t_{ijk})$ defines

$$a \star a = t_{000}a + t_{001}b \quad a \star b = t_{010}a + t_{011}b$$
$$b \star a = t_{100}a + t_{101}b \quad b \star b = t_{110}a + t_{111}b$$

This algebra is generally not associative:

$$b \star (a \star a) = (t_{000}t_{100} + t_{001}t_{110})a + (t_{000}t_{101} + t_{001}t_{111})b$$
$$(b \star a) \star a = (t_{000}t_{100} + t_{101}t_{100})a + (t_{100}t_{001} + t_{101}^2) b$$

Suppose that $T$ is a symmetric tensor, corresponding to a binary cubic

$$t_{000}x^3 + (t_{001} + t_{010} + t_{100})x^2y + (t_{011} + t_{101} + t_{110})xy^2 + t_{111}y^3$$

$$= t_{000}x^3 + 3t_{001}x^2y + 3t_{011}xy^2 + t_{111}y^3$$

$$b \star (a \star a) = (b \star a) \star a \text{ iff } t_{000}t_{011} + t_{001}t_{111} = t_{001}^2 + t_{011}^2 \text{ iff } T \text{ odeco}$$

**Theorem (Boralevi-Draisma-Horobeț-Robeva 2015)**

The odeco equations say that $T$ defines an **associative** algebra.
Our question

How many eigenvectors does a symmetric $3 \times 3 \times 3$-tensor have?

[Diagram of a Rubik's cube]

How many critical points does a cubic have on the unit 2-sphere?

Answer: Seven.
Our question

How many eigenvectors does a symmetric $3 \times 3 \times 3$-tensor have?

How many critical points does a cubic have on the unit 2-sphere?

Fermat: Odeco tensor: $T = x^3 + y^3 + z^3$

$\nabla T : \mathbb{P}^2 \rightarrow \mathbb{P}^2, \ (x : y : z) \mapsto (x^2 : y^2 : z^2)$

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

Answer: Seven.
Let's count

**Theorem**

*Consider a general symmetric tensor $T$ of format $n \times n \times \cdots \times n$. The number of complex eigenvectors in $\mathbb{P}^{n-1}$ equals*

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

[Cartwright, St: The number of eigenvalues of a tensor, 2013]  
[Fornaess, Sibony: Complex dynamics in higher dimensions, 1994]

**Q**: How many *eigenvectors* does a $3 \times 3 \times 3 \times 3$-tensor have?  
**A**: Plug $n = 3$ and $d = 4$ into the formula. The answer is $13$. 

Discriminant

The *eigendiscriminant* is the irreducible polynomial in the entries $t_{i_1i_2...i_d}$ which vanishes when two eigenvectors come together.

**Theorem**

*The degree of eigendiscriminant is* $n(n-1)(d-1)^{n-1}$.  

[Abo, Seigal, St: Eigenconfigurations of tensors, 2015]

**Example 1** ($d=2$) The discriminant of the characteristic polynomial of an $n \times n$-matrix is an equation of degree $n(n-1)$.

**Example 2** ($n=3$, $d=4$) The eigendiscriminant for $3 \times 3 \times 3 \times 3$ tensors is an equation of degree 54.

**Note:** The eigendiscriminant divides tensor space into regions where the number of **real** solutions is constant. Average number?
[Breiding: The expected number of eigenvalues of a real Gaussian tensor, 2016] gives an exact formula in terms of hypergeometric integrals.
Line Arrangements

Open Problem: Can all eigenvectors be real?

Yes, if $n = 3$: All $1 + (d-1) + (d-1)^2$ fixed points can be real.
Open Problem: Can all eigenvectors be real?

Yes, if $n = 3$: All $1 + (d-1) + (d-1)^2$ fixed points can be real.

Proof: Let $T$ be a product of $d$ linear forms.

The $\binom{d}{2}$ vertices of the line arrangement are the base points. The analytic centers of the $\binom{d}{2} + 1$ regions are the fixed points.
Singular vectors

Given a rectangular matrix $T$, one seeks to solve the equations

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

The scalar $\sigma$ is a singular value and $(u, v)$ is a singular vector pair.
Singular vectors

Given a rectangular matrix $T$, one seeks to solve the equations

$$Tu = \sigma v \quad \text{and} \quad T^tv = \sigma u.$$ 

The scalar $\sigma$ is a singular value and $(u, v)$ is a singular vector pair.

**Gradient Dynamics:** Matrices correspond to bilinear forms

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

This defines a rational map

$$\left(\nabla_x T, \nabla_y T\right) : \mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1} \rightarrow \mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1}$$

$$(u, v) \mapsto (T^tv, Tu)$$

The fixed points of this map are the singular vector pairs of $T$. 
Multilinear forms

Tensors $T$ in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ correspond to multilinear forms. The singular vector tuples of $T$ are fixed points of the gradient map $\nabla T : \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1} \rightarrow \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1}$. 

Example: $d = 3$, $n_1 = n_2 = n_3 = 3$:

$$\left( \hat{z}_1 z_2 + \hat{z}_1 z_3 + z_2^3 \right) \left( \hat{z}_2 z_3 + \hat{z}_2 z_4 + z_3^2 \right) \left( \hat{z}_3 z_4 + \hat{z}_3 z_5 + z_2 z_3 \right) = \cdots + 37 z_2 z_3 z_4.$$
Multilinear forms

Tensors $T$ in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ correspond to multilinear forms. The singular vector tuples of $T$ are fixed points of the gradient map

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

Theorem

For a general $n_1 \times n_2 \times \cdots \times n_d$-tensor $T$, the number of singular vector tuples is the coefficient of $z_1^{n_1-1} \cdots z_d^{n_d-1}$ in the polynomial

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

[Friedland, Ottaviani: The number of singular vector tuples...., 2014]
**Multilinear forms**

Tensors $T$ in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ correspond to multilinear forms. The singular vector tuples of $T$ are fixed points of the gradient map

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

**Theorem**

*For a general $n_1 \times n_2 \times \cdots \times n_d$-tensor $T$, the number of singular vector tuples is the coefficient of $z_1^{n_1-1} \cdots z_d^{n_d-1}$ in the polynomial*

$$\prod_{i=1}^{d} \frac{(\hat{z}_i)^{n_i} - z_i^{n_i}}{\hat{z}_i - z_i}$$

where $\hat{z}_i = z_1 + \cdots + z_{i-1} + z_{i+1} + \cdots + z_d$.

[Friedland, Ottaviani: The number of singular vector tuples..., 2014]

**Example:** $d = 3$, $n_1 = n_2 = n_3 = 3$:

$$(\hat{z}_1^2 + \hat{z}_1 z_1 + z_1^2)(\hat{z}_2^2 + \hat{z}_2 z_2 + z_2^2)(\hat{z}_3^2 + \hat{z}_3 z_3 + z_3^2) = \cdots + 37 z_1^2 z_2^2 z_3^2 + \cdots$$
Odeco Tensors

A general tensor of format $3 \times 3 \times 2 \times 2$ has 98 singular vector tuples. What happens for orthogonally decomposable tensors

$$T = x_0y_0z_0w_0 + x_1y_1z_1w_1$$

[Robeva, Seigal: Singular vectors of odeco tensors, 2016]

The gradient map $\nabla T : \mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1 \rightarrow \mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ has only 18 fixed points. In addition, there is a surface of base points:
Conclusion

Eigenvectors of square matrices are central to linear algebra.

Eigenvectors of tensors are a natural generalization. Pioneered in numerical multilinear algebra, these now have many applications.

[Lek-Heng Lim: Singular values and eigenvalues of tensors...., 2005]
[Liqun Qi: Eigenvalues of a real supersymmetric tensor, 2005]

Fact: This lecture serves as an invitation to applied algebraic geometry.

The word variety is not scary.

The terms Segre variety and Veronese variety refer to tensors of rank 1. Given some data, getting close to these is highly desirable.