Decomposing Tensors into Frames

Elina Robeva
UC Berkeley
joint work with Bernd Sturmfels and Luke Oeding

June 8, 2015

Symmetric Tensors

T is an $\underbrace{n \times \ldots \times n}_{d \text{ times}}$ *symmetric tensor* with elements in a field $\mathbb{K}(= \mathbb{R}, \mathbb{C})$ if

$$T_{i_1 i_2 \ldots i_d} = T_{i_{\sigma_1} i_{\sigma_2} \ldots i_{\sigma_d}}$$

for all permutations σ of $\{1, 2, \ldots, d\}$. Notation: $T \in S^d(\mathbb{K}^n)$.

Example (symmetric matrices ($d = 2$))

$$T = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ T_{12} & T_{22} & \cdots & T_{2n} \\ & & \vdots & \\ T_{1n} & T_{2n} & \cdots & T_{nn} \end{pmatrix}$$

Symmetric Tensor Decomposition

For a tensor $T \in S^d(\mathbb{K}^n)$, a symmetric (or Waring) decomposition has the form

$$T = \sum_{i=1}^{r} \lambda_i v_i^{\otimes d}.$$

The smallest r for which such a decomposition exists is the *Waring rank* of T.

Orthogonal Tensor Decomposition

An *orthogonal decomposition* of a symmetric tensor $T \in S^d(\mathbb{R}^n)$ is a decomposition

$$T = \sum_{i=1}^{r} \lambda_i v_i^{\otimes d}$$

such that the vectors $v_1, \ldots, v_r$ are orthonormal. In particular, $r \leq n$.

Definition

A tensor $T \in S^d(\mathbb{R}^n)$ is *orthogonally decomposable*, for short *odeco*, if it has an orthogonal decomposition.

Examples

1. All symmetric matrices are odeco: by the spectral theorem

$$T = V^T \Lambda V = \begin{bmatrix} | & \cdots & | \\ v_1 & \cdots & v_n \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{bmatrix}$$

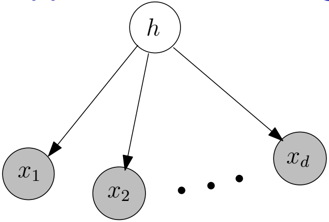
$$= \sum_{i=1}^{n} \lambda_i v_i v_i^T = \sum_{i=1}^{n} \lambda_i v_i^{\otimes 2},$$

   where $v_1, ..., v_n$ is an orthonormal basis of eigenvectors.

2. The Fermat polynomial: If $v_i = e_i$, for $i = 1, ..., n$, then

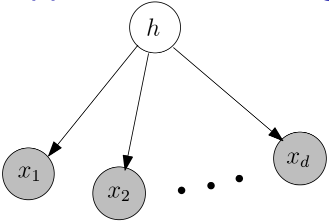
$$T = e_1^{\otimes d} + e_2^{\otimes d} + \cdots + e_n^{\otimes d}.$$

An Application: Exchangeable Single Topic Models



Pick a topic $h \in \{1, 2, ..., k\}$ with distribution $(w_1, ..., w_k) \in \Delta_{k-1}$. Given $h = j$, $x_1, ..., x_d$ are *i.i.d.* random variables taking values in $\{1, 2, ..., n\}$ with distribution $\mu_j = (\mu_{j1}, ..., \mu_{jn}) \in \Delta_{n-1}$.

An Application: Exchangeable Single Topic Models



Pick a topic $h \in \{1, 2, ..., k\}$ with distribution $(w_1, ..., w_k) \in \Delta_{k-1}$. Given $h = j$, $x_1, ..., x_d$ are *i.i.d.* random variables taking values in $\{1, 2, ..., n\}$ with distribution $\mu_j = (\mu_{j1}, ..., \mu_{jn}) \in \Delta_{n-1}$.

Then, the joint distribution of $x_1, ..., x_d$ is an $\underbrace{n \times n \times \cdots \times n}_{d \text{ times}}$ symmetric tensor $T \in S^d(\mathbb{R}^n)$ whose entries sum to 1. Moreover,

$$T_{i_1,...,i_d} = \mathbb{P}(x_1 = i_1, ..., x_d = i_d) = \sum_{j=1}^{k} \mathbb{P}(h = j) \prod_{k=1}^{d} \mathbb{P}(x_k = i_k | h = j).$$

$$\text{So,} \quad T = \sum_{j=1}^{k} w_j \mu_j^{\otimes d}.$$

Eigenvectors of Symmetric Tensors

Consider a symmetric tensor $T \in S^d(\mathbb{K}^n)$.

Example ($d = 2$)

T is an $n \times n$ matrix and $w \in \mathbb{K}^n$ is an eigenvector if

$$Tw = \begin{pmatrix} \vdots \\ \sum_{j=1}^{n} T_{i,j} w_j \\ \vdots \end{pmatrix} = \lambda w.$$

Eigenvectors of Symmetric Tensors

Definition

Given a symmetric tensor $T \in S^d(\mathbb{K}^n)$, an *eigenvector* of T with *eigenvalue* λ is a vector $w \in \mathbb{K}^n$ such that

$$Tw^{d-1} := \begin{pmatrix} \vdots \\ \sum_{i_2,\ldots,i_d=1}^{n} T_{i,i_2,\ldots,i_d} w_{i_2} \ldots w_{i_d} \\ \vdots \end{pmatrix} = \lambda w.$$

Two eigenvector-eigenvalue pairs (w, λ) and (w', λ') are equivalent if there exists $t \in \mathbb{K} \setminus \{0\}$ such that $t^{d-2}\lambda = \lambda'$ and $tw = w'$.

Eigenvectors of Symmetric Tensors

Example

Let

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3}.$$

Then, $(x, y, z)^T$ is an eigenvector of T with eigenvalue λ if

$$T \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}^2 = \begin{pmatrix} 3x^2 \\ 3y^2 \\ 3z^2 \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Equivalently, the 2×2 minors of the matrix $\begin{bmatrix} 3x^2 & x \\ 3y^2 & y \\ 3z^2 & z \end{bmatrix}$ vanish. Therefore,

$$x^2y - xy^2 = x^2z - xz^2 = y^2z - yz^2 = 0.$$

The solutions are (up to scaling):

$$\{(1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)\}.$$

Eigenvectors of Odeco Tensors

If $T = \sum_{i=1}^{n} \lambda_i v_i^{\otimes d}$ is an odeco tensor, i.e. $v_1, ..., v_n$ are orthonormal, then $v_1, ..., v_n$ are eigenvectors of T with eigenvalues $\lambda_1, ..., \lambda_n$ respectively.

- Is there an easy way of finding these vectors, i.e. finding the orthogonal decomposition of an odeco tensor?
- Are these all of the eigenvectors of an odeco tensor?

Robust Eigenvectors and the Tensor Power Method

Definition

A unit vector $u \in \mathbb{R}^n$ is a *robust eigenvector* of a tensor $T \in S^d(\mathbb{R}^n)$ if there exists $\epsilon > 0$ such that for all $\theta \in \mathcal{B}_\epsilon(u) = \{u' : ||u - u'|| < \epsilon\}$, repeated iteration of the map

$$\theta \mapsto \frac{T\theta^{d-1}}{||T\theta^{d-1}||}, \tag{1}$$

starting from θ converges to u.

Theorem (Anandkumar et al.)

Let T have an orthogonal decomposition $T = \sum_{i=1}^k \lambda_i v_i^{\otimes d}$ with $\lambda_1, ..., \lambda_k > 0$.
Then, the set of robust eigenvectors of T is equal to $\{v_1, v_2, ..., v_k\}$.

Decomposing Tensors into Frames

Question: How to enlarge the set of odeco tensors to contain tensors of higher ranks?
Idea:

- Odeco tensors come from an orthonormal basis $V := (\mathbf{v}_1, \cdots, \mathbf{v}_n) \in (\mathbb{R}^n)^n$, i.e.

$$VV^T = I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., n.$$

Decomposing Tensors into Frames

Question: How to enlarge the set of odeco tensors to contain tensors of higher ranks?
Idea:

- Odeco tensors come from an orthonormal basis $V := \big(\mathbf{v}_1, \cdots, \mathbf{v}_n\big) \in (\mathbb{R}^n)^n$, i.e.

$$VV^T = I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., n.$$

- Instead let $V := \big(\mathbf{v}_1, \cdots, \mathbf{v}_r\big) \in (\mathbb{R}^n)^r$ be a *finite unit norm tight frame* (or *funtf*) , i.e.

$$VV^T = \frac{r}{n} I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., r.$$

Decomposing Tensors into Frames

Question: How to enlarge the set of odeco tensors to contain tensors of higher ranks?
Idea:

- Odeco tensors come from an orthonormal basis $V := (\mathbf{v}_1, \cdots, \mathbf{v}_n) \in (\mathbb{R}^n)^n$, i.e.

$$VV^T = I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., n.$$

- Instead let $V := (\mathbf{v}_1, \cdots, \mathbf{v}_r) \in (\mathbb{R}^n)^r$ be a *finite unit norm tight frame* (or *funtf*) , i.e.

$$VV^T = \frac{r}{n} I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., r.$$

A tensor $T \in S^d(\mathbb{R}^n)$ is *frame decomposable* (or *fradeco*) if it can be written as

$$T = \sum_{i=1}^{r} \lambda_i \mathbf{v}_i^{\otimes d},$$

where $(\mathbf{v}_1, ..., \mathbf{v}_r)$ form a finite unit norm tight frame.

Finite Unit Norm Tight Frames

Examples

- The Mercedes Benz Frame $V = \begin{pmatrix} 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$.
- $V = \frac{1}{3\sqrt{3}} \begin{pmatrix} -5 & 1 & 1 & 3 \\ 1 & -5 & 1 & 3 \\ 1 & 1 & -5 & 3 \end{pmatrix}$.
- $V = \begin{pmatrix} 1 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 1 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$.

The *funtf variety*, denoted by $\mathcal{F}_{r,n}$ is the sub variety of $\mathbb{C}^{nr}$ defined by the equations

$$VV^T = \frac{r}{n} I_n \quad \text{and} \quad ||\mathbf{v}_j||^2 = 1, j = 1, ..., r.$$

Finite Unit Norm Tight Frames

Theorem (Cahill, Mixon, Strawn, Dykema)

The funtf variety $\mathcal{F}_{r,n}$ is irreducible when $r \geq n+2 > 4$. We have

$$\dim(\mathcal{F}_{r,n}) \;=\; (n-1)\cdot(r-\frac{n}{2}-1) \qquad \textit{when } \; r > n \geq 2.$$

r	n	$\dim \mathcal{F}_{r,n}$	$\deg \mathcal{F}_{r,n}$	# components & degrees
3	2	1	$8 \cdot 2$	8 components, each degree 2
4	2	2	$12 \cdot 4$	12 components, each degree 4
5	2	3	112	irreducible
6	2	4	240	irreducible
7	2	5	496	irreducible
4	3	3	$16 \cdot 8$	16 components, each degree 8
5	3	5	1024	irreducible
6	3	7	2048	irreducible
7	3	9	4096	irreducible
5	4	6	$32 \cdot 40$	32 components, each degree 40
6	4	9	20800	irreducible
7	4	12	65536	irreducible

Table: Dimension and degree of the funtf variety in some small cases

The Fradeco Variety

Questions:

- Does the tensor power method recover the decomposition of a fradeco tensor?
- Is there another way to recover the decomposition?

The *variety of frame decomposable tensors* or the *fradeco variety*, denoted by $\mathcal{T}_{r,n,d}$ is the Zariski closure in $S^d(\mathbb{C}^n)$ of all tensors that can be written as

$$T = \sum_{j=1}^{r} \lambda_j \mathbf{v}_j^{\otimes d},$$

where $(\mathbf{v}_1, ..., \mathbf{v}_r) \in \mathcal{F}_{r,n}$.

Example

When $r = n$, $\mathcal{T}_{n,n,d}$ is the odeco variety.

- Can we give equations that define the variety $\mathcal{T}_{r,n,d}$.

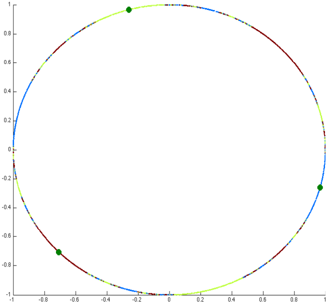
The tensor power method

Conjecture

Let $r = n+1 < d$ and $T = \sum_{j=1}^{n+1} \lambda_j \mathbf{v}_j^{\otimes d} \in \mathcal{T}_{n+1,n,d}$ with $\lambda_1, \ldots, \lambda_{n+1} > 0$. Then, $\mathbf{v}_1, \ldots, v_{n+1}$ are the robust eigenvectors of T, so they are found by the tensor power method.

Example (The Mercedes Benz Frame)

Let $T = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{\otimes 5} + \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}^{\otimes 5} + \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}^{\otimes 5}$. The dynamics of the power method looks like this



The tensor power method

Example

Let $n = 2, r = 4, d = 5$ and consider the tensor

$$T = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\otimes 4} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{\otimes 4} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{\otimes 4} + \begin{pmatrix} 1 \\ -1 \end{pmatrix}^{\otimes 4} \in \mathcal{T}_{4,2,5},$$

where $\alpha > 6$. The vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector, but none of the other eigenvectors are real. Therefore, the frame decomposition of T cannot be recovered from its eigenvectors.

Symmetric $2 \times 2 \times \cdots \times 2$ fradeco tensors

Given a symmetric tensor $T \in S^d(\mathbb{C}^2)$, denote

$$t_i = T_{i_1,\ldots,i_d}$$

whenever i of $i_1, ..., i_d$ are equal to 1 and the rest are equal to 0.

Theorem

Fix $r \in \{3, \ldots, 9\}$ and $d \geq 2r - 2$. There exists a matrix $\mathcal{M}_r$ with

(a) *... $r-1$ rows and $d-r+1$ columns, whose entries are linear forms in the coordinates $t_0, t_1, \ldots, t_d$ on $\mathbb{P}^d$.*

(b) *The columns of $\mathcal{M}_r$ involve $r+1$ of the unknowns t_i, and they are identical up to index shifts.*

(c) *The maximal minors of $\mathcal{M}_r$ form a Gröbner basis for the homogeneous prime ideal of $\mathcal{T}_{r,2,d}$.*

Lemma

The fradeco variety $\mathcal{T}_{r,2,d}$ has dimension $2r-3$ in $\mathbb{P}^d$.

Matrices for Binary Fradeco Tensors

$$\mathcal{M}_3 = \begin{pmatrix} t_0 - 3t_2 & t_1 - 3t_3 & t_2 - 3t_4 & t_3 - 3t_5 & \cdots & t_{d-3} - 3t_{d-1} \\ 3t_1 - t_3 & 3t_2 - t_4 & 3t_3 - t_5 & 3t_4 - t_6 & \cdots & 3t_{d-2} - t_d \end{pmatrix}$$

$$\mathcal{M}_4 = \begin{pmatrix} t_0 + t_4 & t_1 + t_5 & t_2 + t_6 & t_3 + t_7 & \cdots & t_{d-4} + t_d \\ t_1 - t_3 & t_2 - t_4 & t_3 - t_5 & t_4 - t_6 & \cdots & t_{d-3} + t_{d-1} \\ t_2 & t_3 & t_4 & t_5 & \cdots & t_{d-2} \end{pmatrix}$$

$$\mathcal{M}_5 = \begin{pmatrix} t_0 + 5t_2 & t_1 + 5t_3 & t_2 + 5t_4 & t_3 + 5t_5 & \cdots & t_{d-5} + 5t_{d-3} \\ t_1 - 3t_3 & t_2 - 3t_4 & t_3 - 3t_5 & t_4 - 3t_6 & \cdots & t_{d-4} - 3t_{d-2} \\ 3t_2 - t_4 & 3t_3 - t_5 & 3t_4 - t_6 & 3t_5 - t_7 & \cdots & 3t_{d-3} - t_{d-1} \\ 5t_3 + t_5 & 5t_4 + t_6 & 5t_5 + t_7 & 5t_6 + t_8 & \cdots & 5t_{d-2} + t_d \end{pmatrix}$$

$$\mathcal{M}_6 = \begin{pmatrix} t_0 + 3t_2 & t_1 + 3t_3 & t_2 + 3t_4 & t_3 + 3t_5 & \cdots & t_{d-6} + 3t_{d-4} \\ t_1 + t_5 & t_2 + t_6 & t_3 + t_7 & t_4 + t_8 & \cdots & t_{d-5} + t_{d-1} \\ t_2 - t_4 & t_3 - t_5 & t_4 - t_6 & t_5 - t_7 & \cdots & t_{d-4} - t_{d-2} \\ t_3 & t_4 & t_5 & t_6 & \cdots & t_{d-3} \\ 3t_4 + t_6 & 3t_5 + t_7 & 3t_6 + t_8 & 3t_7 + t_9 & \cdots & 3t_{d-2} + t_d \end{pmatrix}$$

$$\cdots \quad \cdots \quad \cdots$$

We conjecture that this works for all r.

Higher Dimensions

Theorem

The following table describes the fradeco varieties $\mathcal{T}_{r,n,d}$ in the cases when $n \geq 3$ and $1 \leq \dim(\mathcal{T}_{r,n,d}) \cdot \operatorname{codim}(\mathcal{T}_{r,n,d}) \leq 100$.

variety	dim	codim	degree	known equations
$\mathcal{T}_{4,3,3}$	6	3	17	3 cubics, 6 quartics
$\mathcal{T}_{4,3,4}$	6	8	74	6 quadrics, 37 cubics
$\mathcal{T}_{4,3,5}$	6	14	191	27 quadrics, 104 cubics
$\mathcal{T}_{5,3,4}$	9	5	210	1 cubic, 6 quartics
$\mathcal{T}_{5,3,5}$	9	11	1479	20 cubics, 213 quartics
$\mathcal{T}_{6,3,4}$	12	2	99	none in degree ≤ 5
$\mathcal{T}_{6,3,5}$	12	8	4269	one quartic
$\mathcal{T}_{7,3,5}$	15	5	≥ 38541	none in degree ≤ 4
$\mathcal{T}_{8,3,5}$	18	2	690	none in degree ≤ 5
$\mathcal{T}_{10,3,6}$	24	3	≥ 16252	none in degree ≤ 7
$\mathcal{T}_{5,4,3}$	10	9	830	none in degree ≤ 4
$\mathcal{T}_{6,4,3}$	14	5	1860	none in degree ≤ 3
$\mathcal{T}_{7,4,3}$	18	1	194	one in degree 194

Proposition

For all $r > n, d \geq 3$, the dimension of $\mathcal{T}_{r,n,d}$ is bounded above by

$$\min\left\{(r-n)(n-1) + \binom{n-1}{2} + r - 1\,,\ \binom{n+d-1}{d} - 1\right\}.$$

Thank you!

References

- A. Anandkumar, R. Ge, D. Hsu, S. Kakade, and M. Telegarsky: *Tensor Decompositions for Learning Latent Variable Models*
- A. Anandkumar, D. Hsu, and S. Kakade: *A Method of Moments for Mixture Models and Hidden Markov Models*
- J. Brachat, P. Common, B. Mourrain, and E. Tsigaridas: *Symmetric Tensor Decomposition*
- P. Common, G. Golub, L. Lim, and B. Mourrain: *Symmetric Tensors and Symmetric Tensor Rank*
- D. Cartwright and B. Sturmfels: *The Number of Eigenvalues of a Tensor*
- D. Eisenbud and B. Sturmfels: *Binomial Ideals*
- L. Oeding and G. Ottaviani: *Eigenvectors of Tensors and Algorithms for Waring Decomposition*