Counting critical rank-one approximations to tensors

Jan Draisma (j.w.w. Emil Horobeţ) TU Eindhoven

Tensors of bounded rank

Setting

$V_1, \ldots, V_p$ vector spaces over K, $\dim V_i =: n_i$

$V := V_1 \otimes \cdots \otimes V_p$ has dimension $n_1 \cdots n_p$.

Each $u \in V$ is of the form $\sum_{i=1}^{k} v_{i1} \otimes \cdots \otimes v_{ip}$.

Definition

The minimal k in any such expression is the rank of u.

The border rank is the minimal k such that u lies in the Zariski closure of $\{v \in V \mid \operatorname{rk} v \leq k\}$ (assume K infinite).

Dimension

Tensors of border rank $\leq k$ form a variety of dimension $\leq k \cdot [1 + \sum_{i=1}^{p}(n_i - 1)]$, which is $\ll n_1 \cdots n_p$ for small k.

Low-rank approximation

Setting
Assuming $K = \mathbb{R}$, equip each V_i with a positive definite inner product $(.|.)_i$. Equip V with the inner product determined by $(u_1 \otimes \cdots \otimes u_p | v_1 \otimes \cdots \otimes v_p) := \prod_i (u_i|v_i)_i$.

Inspiring problem
Given $u \in V$ and $k \in \mathbb{N}$, find $x \in V$ of rank $\leq k$ that minimises $d_u(x) := ||u - x||^2$.

Related problem
For (sufficiently general) $u \in V$, count the critical points of d_u on the smooth part of the set $\{v \in V \mid \text{rk } v \leq k\}$.

My talk: $k = 1$.

Rank-one approximation for matrices

Setting
$X := \{v_1 \otimes \cdots \otimes v_p\} \setminus \{0\} \subseteq V$ is the manifold of pure tensors.
Given $u \in V$, we count critical points of d_u on X.

Note
Count doesn't change when acting with $\mathrm{O}(V_1) \times \cdots \times \mathrm{O}(V_p)$.

Case $p = 2, n_1 \leq n_2$ **(Eckart-Young Theorem)**
$\exists g_i \in \mathrm{O}(V_i),\ i = 1, 2$ such that $(g_1 \otimes g_2)u = \sum_{i=1}^{n_1} c_i \cdot e_i \otimes f_i$
where $e_1, \ldots, e_{n_1}$ orthonormal in V_1 and $f_1, \ldots, f_{n_1}$
orthonormal in V_2 and $c_1 \geq \ldots \geq c_{n_1} \geq 0$ (singular values).

The critical points are $(g_1^{-1}, g_2^{-1})c_i \cdot e_i \otimes f_i$ for $i = 1, \ldots, n_1$;
there are n_1 of these.

Main result: ordinary tensors

Fact
For $p \geq 3$ the number of critical points of d_u on X typically jumps as u crosses a hypersurface; we compute an average.

Theorem (D-Horobeţ)
Draw u uniformly from the unit sphere in V centered at 0. Then the expected # of critical points of d_u equals

$$\frac{(2\pi)^{p/2}}{2^{n/2}} \frac{1}{\prod_{i=1}^{p} \Gamma\left(\frac{n_i}{2}\right)} \mathbb{E}(|\det C|)$$

where $n := n_1 + \cdots + n_p$ and where C is a symmetric random $(n-p) \times (n-p)$-matrix with certain structural zeroes.

Main result, continued

Structure of C

$$C = \begin{bmatrix} w_0 I_{n_1-1} & C_{1,2} & \cdots & C_{1,p} \\ C_{1,2}^T & w_0 I_{n_2-1} & \cdots & C_{2,p} \\ \vdots & \vdots & & \vdots \\ C_{1,p}^T & C_{2,p}^T & \cdots & w_0 I_{n_p-1} \end{bmatrix}$$

where C_{ij} is $(n_i - 1) \times (n_j - 1)$, and where w_0 and the entries of all C_{ij} are independent and $\sim \mathcal{N}(0, 1)$.

Sanity check: $p = 2, n_1 = n_2 = 2$

$$C = \begin{bmatrix} w_0 & w_{12} \\ w_{12} & w_0 \end{bmatrix}, \ |\det C| = |w_0^2 - w_{12}^2|, \ \mathbb{E}(|\det C|) = \frac{4}{\pi}$$

$\rightsquigarrow$ get $\frac{(2\pi)^1}{2^2} \cdot 1 \cdot \frac{4}{\pi} = 2$, as given by the Eckart-Young theorem.

More general problem
Given any real algebraic variety X in a Euclidean space V, and given $u \in V$, count the critical points of $d_u(x) := \|u - x\|^2$ on the manifold X_{reg} , i.e., count the x where $u - x \perp T_x X$.

Definition (D-H-Ottaviani-Sturmfels-Thomas)
Complexify $(.|.)$ to a symmetric bilinear form on $V_{\mathbb{C}}$. Then for general $u \in V_{\mathbb{C}}$ the number of smooth points $x \in X_{\mathbb{C}}$ with $u - x \perp T_x X_{\mathbb{C}}$ is a constant called the ED degree of X (or $X_{\mathbb{C}}$).

The average ED degree of X w.r.t. a probability measure on V is the expected number of critical points of d_u for random $u \in V$. This is a real, rather than complex count.

ED degree for rank-one tensors

Setting
Complexify (.|.) from before to a symmetric bilinear form on $V_{\mathbb{C}} = (V_1 \otimes \cdots \otimes V_p)_{\mathbb{C}}$; $X_{\mathbb{C}}$ consists of complex pure tensors.

Theorem (Friedland-Ottaviani)
$\text{EDdegree}(X) =$ coefficient of $s_1^{n_1-1} \cdots s_p^{n_p-1}$ in $\prod_{i=1}^{p} \frac{\hat{s}_i^{n_i} - s_i^{n_i}}{\hat{s}_i - s_i}$,
where $\hat{s}_i = s_1 + \cdots + s_{i-1} + s_{i+1} + \cdots + s_p$.

Sanity check: $p = 2, n_1 \leq n_2$
$\rightsquigarrow$ We get the coefficient of $s_1^{n_1-1} s_2^{n_2-1}$ in
$\frac{s_2^{n_1} - s_1^{n_1}}{s_2 - s_1} \cdot \frac{s_1^{n_2} - s_2^{n_2}}{s_1 - s_2} = (s_2^{n_1-1} + \cdots + s_1^{n_1-1}) \cdot (s_1^{n_2-1} + \cdots + s_2^{n_2-1})$
which equals n_1 as given by the Eckart-Young theorem.

(Average) ED degrees for rank-one tensors

Tensor format	average ED degree (/$\mathbb{R}$)	ED degree (/$\mathbb{C}$)
$n_1 \times n_2$	$\min(n_1, n_2)$	$\min(n_1, n_2)$
$2^3 = 2 \times 2 \times 2$	4.287	6
2^4	11.06	24
2^5	31.56	120
2^6	98.82	720
2^7	333.9	5040
2^8	$1.206 \cdot 10^3$	40320
$2 \times 2 \times 3$	5.604	8
$2 \times 2 \times 4$	5.556	8
$2 \times 2 \times 5$	5.536	8
$2 \times 3 \times 3$	8.817	15
$2 \times 3 \times 4$	10.39	18
$2 \times 3 \times 5$	10.28	18
$3 \times 3 \times 3$	16.03	37
$3 \times 3 \times 4$	21.28	55

General setting, $\mathbb{R}$ or $\mathbb{C}$ (D-H-O-S-T)
$\mathcal{E}_X := \{(x, u) \mid x \in X_{\mathrm{reg}} \text{ critical for } d_u\} \subseteq X \times V$ is called the ED correspondence of X.

Observation
$\pi_1 : \mathcal{E}_X \to X$ is an affine vector bundle $/X_{\mathrm{reg}}$ of rank codim$X$.
$\pi_2 : \mathcal{E}_X \to V$ has fibres whose cardinalities we try to count.

Cones
Assume X is closed under scalar multiplication, so $\mathbb{P}X \subseteq \mathbb{P}V$ is a projective variety. Let $\mathbb{P}\mathcal{E}_X$ be the image of $\mathcal{E}_X$ in $\mathbb{P}X \times V$. Its fibre over $[x] \in \mathbb{P}X$ with $(x|x) \neq 0$ equals $\langle x \rangle + (T_x X)^\perp$.

$\rightsquigarrow$ the projective ED correspondence $\mathbb{P}\mathcal{E}_X$ is a vector bundle over $\mathbb{P}X_{\mathrm{reg}}$ of rank codim$X + 1$ away from $Q := \{v \mid (v|v) = 0\}$.

Towards vector bundle methods

$u \in V$ gives a section of the quotient bundle $(\mathbb{P}X \times V)/\mathbb{P}\mathcal{E}_X$, and the ED degree counts the zeroes of this section (but there may be problems at Q and outside X_{reg}).

Zeroes of a general sections of a vector bundle $\mathcal{E} \to \mathbb{P}X$ of rank $\dim \mathbb{P}X$ are counted by the degree of the top Chern class of $\mathcal{E}$. This class lives in the Chow ring of $\mathbb{P}X$.

So control over the Chow ring of $\mathbb{P}X$ and the behaviour at singular points and at Q can yield the ED degree.

For $X = \{\text{pure tensors}\}$, $\mathbb{P}X = \mathrm{Seg}(\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_p-1})$. The Chow ring is $\mathbb{Z}[s_1, \ldots, s_p]/(s_1^{n_1}, \ldots, s_p^{n_p})$.

Proof sketch of the Friedland-Ottaviani formula

EDdegree(X) = coefficient of $s_1^{n_1-1} \cdots s_p^{n_p-1}$ in $\prod_{i=1}^{p} \frac{\hat{s}_i^{n_i} - s_i^{n_i}}{\hat{s}_i - s_i}$,
where $\hat{s}_i = s_1 + \cdots + s_{i-1} + s_{i+1} + \cdots + s_p$.

Tangent space

$x = v_1 \otimes \cdots \otimes v_p \rightsquigarrow T_x X = \sum_{j=1}^{p} v_1 \otimes \cdots \otimes V_j \otimes \cdots \otimes v_p$.

Proof sketch

$([x], u) \in \mathbb{P}\mathcal{E}_X \Leftrightarrow \exists c \forall_i : u - cx \perp v_1 \otimes \cdots \otimes V_i \otimes \cdots \otimes v_p$
$\Leftrightarrow \exists c \forall_i : u - cx \perp x$ and $u \perp v_1 \otimes \cdots \otimes v_i^{\perp} \otimes \cdots \otimes v_p$
$\Leftrightarrow u \perp v_1 \otimes \cdots \otimes v_i^{\perp} \otimes \cdots \otimes v_p$ (if $(x|x) \neq 0$)
Define $\mathcal{E}_i$: bundle on $\mathbb{P}X$ with fibre $(v_1 \otimes \cdots \otimes v_i^{\perp} \otimes \cdots \otimes v_p)^*$
The tensor u gives a section of the bundle $\bigoplus_i \mathcal{E}_i$, whose top Chern class has degree = coefficient. $\square$

Average ED degrees by double counting

Recall (now over $\mathbb{R}$)

$\pi_1 : \mathcal{E}_X \to X$ is an affine vector bundle $/X_{\text{reg}}$ of rank $\operatorname{codim} X$.
$\pi_2 : \mathcal{E}_X \to V$ has fibres whose cardinalities we try to count.

Assume that the probability distribution on V is given by a density f, and that we have a generically one-to-one map φ from $\mathbb{R}^{\dim V}$ to $\mathcal{E}_X$. Then the average ED degree of X equals

$$\int_V \#|\pi_2^{-1}(u)| \cdot f(u)\mathrm{d}u = \int_{\mathbb{R}^{\dim V}} |J(\pi_2 \circ f)(t)| \cdot f(\pi_2(\varphi(t)))\mathrm{d}t.$$

For rational varieties X we can take φ to be birational. If X is a cone, we can also work with $\mathbb{P}\mathcal{E}_X$.

Proof sketch of the D-Horobeţ formula

Fix a norm-1 vector $e_i \in V_i$ for $i = 1, \ldots, p$.

Set $W := (\bigoplus_{i=1}^{p} e_1 \otimes \cdots \otimes e_i^{\perp} \otimes \cdots \otimes e_p)^{\perp} \subseteq V$, the fibre over $[e_1 \otimes \cdots \otimes e_p] \in \mathbb{P}X$ in $\mathbb{P}\mathcal{E}_X$.

Define the birational map $\psi : e_1^{\perp} \times \cdots \times e_p^{\perp} \to \mathbb{P}X$ by $(v_1, \ldots, v_p) \mapsto [(e_1 + v_1) \otimes \cdots \otimes (e_p + v_p)]$. By symmetry, the π_2-fibre over $\psi(v_1, \ldots, v_p)$ equals gW, where $g \in \prod_i O(V_i)$ is any element such that g_i maps $[e_i]$ into $[e_i + v_i]$.

Choose $g_i := \left(I - e_i e_i^T - \frac{v_i}{\|v_i\|} \frac{v_i^T}{\|v_i\|}\right) + \left(\frac{e_i + v_i}{\sqrt{1+\|v_i\|^2}} e_i^T + \frac{v_i - \|v_i\|^2 e_i}{\|v_i\| \sqrt{1+\|v_i\|^2}} \frac{v_i^T}{\|v_i\|}\right)$
and apply double counting to
$\psi : \prod_i (e_i)^{\perp} \times W \to \mathbb{P}\mathcal{E}_X, (v_1, \ldots, v_p) \mapsto (\varphi(v_1, \ldots, v_p), gw)$. $\square$

Further result: symmetric tensors

Setting

Equip $V := \mathrm{Sym}^p\mathbb{R}^n$ with the positive inner product where $(v_1^p|v_2^p) = (v_1|v_2)^p$ (the Bombieri inner product). Now we approximate $u \in V$ by an element of $X := \{\pm v^p \mid v \in \mathbb{R}^n \setminus \{0\}\}$.

Theorem (D-Horobeţ)

Draw u from the uniform distribution on the unit sphere in V centered at 0. Then $\mathbb{E}(\#\text{critical points of } d_u \text{ on } X)$ equals

$$\frac{1}{2^{(n^2+3n-2)/4}\prod_{i=1}^n \Gamma(i/2)} \int_{\lambda_2\le\ldots\le\lambda_n} \int_{-\infty}^{\infty} \left(\prod_{i=2}^n |\sqrt{p}w_0 - \sqrt{p-1}\lambda_i|\right)\cdot \left(\prod_{i<j}(\lambda_j-\lambda_i)\right) e^{-w_0^2/2-\sum_{i=2}^n \lambda_i^2/4} \mathrm{d}w_0 \mathrm{d}\lambda_2 \cdots \mathrm{d}\lambda_n$$

ED degrees for symmetric rank-1 tensors

ED degrees (right) and average ED degrees (left)

$p\backslash n$	1	2	3	4	1	2	3	4
1	1	1	1	1	1	1	1	1
2	1	2	3	4	1	2	3	4
3	1	$\sqrt{7}$	$1+4\cdot\frac{2}{7}\cdot\sqrt{7\cdot 2}$	$9.3951\ldots$	1	3	7	15
4	1	$\sqrt{10}$	$1+4\cdot\frac{3}{10}\cdot\sqrt{10\cdot 3}$	$16.254\ldots$	1	4	13	40
5	1	$\sqrt{13}$	$1+4\cdot\frac{4}{13}\cdot\sqrt{13\cdot 4}$	$24.300\ldots$	1	5	21	85
6	1	$\sqrt{16}$	$1+4\cdot\frac{5}{16}\cdot\sqrt{16\cdot 5}$	$33.374\ldots$	1	6	31	156
7	1	$\sqrt{19}$	$1+4\cdot\frac{6}{19}\cdot\sqrt{19\cdot 6}$	$43.370\ldots$	1	7	43	259
8	1	$\sqrt{22}$	$1+4\cdot\frac{7}{22}\cdot\sqrt{22\cdot 7}$	$54.21\ldots$	1	8	57	400

Theorem (Cartwright-Sturmfels)
The ED degree of $X \subseteq S^p V$ is $1+(p-1)+\cdots+(p-1)^{n-1}$.

Further remarks and problems

- Relation to singular vector tuples and eigenvectors. **(Lim)**
- Other inner products on V?
- Closed form expression for the average ED degree for sym case?
- What is the expected number of local minima of d_u?
- Describe hypersurface where $|\pi_2^{-1}(u)|$ jumps (the ED discriminant).
- Find all typical values of $|\pi_2^{-1}(u)|$ (over $\mathbb{R}$).
- Give a geometric proof for "stabilisation" when $n_p - 1 \geq \sum_{i=1}^{p-1}(n_i - 1)$.
- Can knowledge of (average) ED degrees be used in algorithms?
- How about rank two?

Further remarks and problems

- Relation to singular vector tuples and eigenvectors. **(Lim)**
- Other inner products on V?
- Closed form expression for the average ED degree for sym case?
- What is the expected number of local minima of d_u?
- Describe hypersurface where $|\pi_2^{-1}(u)|$ jumps (the ED discriminant).
- Find all typical values of $|\pi_2^{-1}(u)|$ (over $\mathbb{R}$).
- Give a geometric proof for "stabilisation" when $n_p - 1 \geq \sum_{i=1}^{p-1}(n_i - 1)$.
- Can knowledge of (average) ED degrees be used in algorithms?
- How about rank two?

References

- Jan Draisma and Emil Horobeţ: *The average number of critical rank-one approximations to a tensor*, `arxiv:1408.3507`
- JD, EH, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas: *The Euclidean distance degree of an algebraic variety*, Found. Comput. Math., to appear, `arxiv:1309.0049`

Further remarks and problems

- Relation to singular vector tuples and eigenvectors. **(Lim)**
- Other inner products on V?
- Closed form expression for the average ED degree for sym case?
- What is the expected number of local minima of d_u?
- Describe hypersurface where $|\pi_2^{-1}(u)|$ jumps (the ED discriminant).
- Find all typical values of $|\pi_2^{-1}(u)|$ (over $\mathbb{R}$).
- Give a geometric proof for "stabilisation" when $n_p - 1 \geq \sum_{i=1}^{p-1}(n_i - 1)$.
- Can knowledge of (average) ED degrees be used in algorithms?
- How about rank two?

References

- Jan Draisma and Emil Horobeţ: *The average number of critical rank-one approximations to a tensor*, `arxiv:1408.3507`
- JD, EH, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas: *The Euclidean distance degree of an algebraic variety*, Found. Comput. Math., to appear, `arxiv:1309.0049`

Thank you!