Counting Singular Vectors of a Multidimensional Tensor

SIAM Conference on Applied Algebraic Geometry, Fort Collins, August 1-4, 2013

Colorado State University

MS19, Algebraic Geometry of Tensor Decompositions

Shmuel Friedland, Giorgio Ottaviani[1]

Università di Firenze[1]

Rank one matrices

Recall that a $m \times n$ matrix A, on a field K, has rank one if and only if there are $x \in K^m$, $y \in K^n$ such that $A = x \otimes y$. Let $K = \mathbb{R}$, then we can write $A = \lambda x \otimes y$ where $\lambda \geq 0$, $|x| = |y| = 1$. The variety of rank one matrices is the cone over the Segre variety $\mathbb{P}(K^m) \times \mathbb{P}(K^n)$.

Given a matrix U, we consider the distance function from U to the variety of rank one matrices.

SVD and critical points

Lemma (SVD revisited)

The critical points of the distance function $d_U = d(U, -)$ *from a matrix* U *to the variety of rank one matrices are given by* $\lambda_i x_i \otimes y_i$ *such that* $Ux_i = \lambda_i y_i$, $U^t y_i = \lambda_i x_i$.
Moreover we have the singular value decomposition $U = XDY$ *where* x_i *are the columns of* X, y_j *are the rows of* Y *and* D *has* λ_i *on the diagonal.*

Main remark is that Ux_i and Uy_i coincide with the two contractions of U over $x \otimes y$.

Lemma (SVD revisited)

The critical points of the distance function $d_U = d(U, -)$ *from a matrix* U *to the variety of rank one matrices are given by* $\lambda_i x_i \otimes y_i$ *such that* $Ux_i = \lambda_i y_i$, $U^t y_i = \lambda_i x_i$.
Moreover we have the singular value decomposition $U = XDY$ *where* x_i *are the columns of* X, y_j *are the rows of* Y *and* D *has* λ_i

The distance and the contraction.

We have the formula
$|U - X|^2 = |U|^2 + |X|^2 - 2 < U, X >$,
which implies that, on the sphere $|X| = 1$, the critical points of $d_U = d(U, -)$ are the same than the critical points of the contraction $< U, X >$. Indeed these critical points can be found by asking that the vector $U - X$ is orthogonal to the tangent space at X to the variety corresponding to S.
This language generalizes to tensors.

The singular d-ples

Let $S = S_{m_1-1} \times \ldots \times S_{m_d-1}$ be the product of the unitary euclidean spheres $S_i \subset \mathbb{R}^i$.

Note that S maps in a natural way to the tensor product, and that the cone over S coincides with the cone over the Segre variety. In equivalent way, we may consider the *euclidean distance* d_t of t from the cone over the Segre variety.

Theorem (Lim, Qi)

The critical points of the distance from a tensor t (evaluated on the product of spheres S) correspond to the tensors $(x_1, \ldots, x_d) \in S$ such that

$$t(x_1, \ldots, \hat{x}_i, \ldots, x_d) = \lambda x_i \quad \forall i = 1, \ldots, d \qquad (1)$$

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

The d-ples $(x_1, \ldots, x_d)$ which satisfy (2) are called singular d-ples. In case $d = 2$, they are the usual pairs of singular vectors.

Let $S = S_{m_1-1} \times \ldots \times S_{m_d-1}$ be the product of the unitary

The singular values depend on the scaling in tensor case

In the matrix case, λ's are well defined up to scaling, indeed they are the square roots of eigenvalues of UU^t or U^tU.
In the tensor case, the equation

$$t(x_1, \ldots, \hat{x}_i, \ldots, x_d) = \lambda x_i$$

is no more homogeneous and the λ have no intrinsic (geometric) meaning, in the projective setting we may only divide between the zero or nonzero ones.

Singular vectors may leave the base field.

If t is a real tensor, we may get complex singular d-ples for $d \geq 3$. Since the general d-ple does not contain isotropic vectors, still we can count properly the singular $d$ples by using algebraic geometry techniques over $\mathbb{C}$.
We get an upper bound for the real solutions, which is attained by some real tensor.

The problem of counting singular d-ples

How many are the singular d-ples of a general tensor?

In the format $(2, 2, 2)$ they are 6, in the format $(3, 3, 3)$ they are 37. Note they are more than the dimension of the factors, and even more than the dimension of the ambient space.

The number of singular vector d-ples

Theorem (Friedland-O)

The number of singular d-ples of a general tensor $t \in \mathbb{P}(\mathbb{R}^{m_1}) \times \ldots \times \mathbb{P}(\mathbb{R}^{m_d})$ over $\mathbb{C}$ of format $(m_1, \ldots, m_d)$ is equal to the coefficient of $\prod_{i=1}^{d} t_i^{m_i-1}$ in the polynomial

$$\prod_{i=1}^{d} \frac{\hat{t}_i^{\,m_i} - t_i^{m_i}}{\hat{t}_i - t_i}$$

where $\hat{t}_i = \sum_{j \neq i} t_j$

Amazingly, for $d = 2$ this formula gives the expected value $\min(m_1, m_2)$.

Interpretation with vector bundles

For the proof, we express the d-ples of singular vectors as zero loci of sections of a suitable vector bundle on the Segre variety. Precisely, let $X = \mathbb{P}(\mathbb{C}^{m_1}) \times \ldots \times \mathbb{P}(\mathbb{C}^{m_d})$ and let $\pi_i \colon X \to \mathbb{P}(\mathbb{C}^{m_i})$ be the projection on the i-th factor. Let $\mathcal{O}(\underbrace{1, \ldots, 1}_{d})$ be the very ample line bundle which gives the Segre embedding.

Then the bundle is $\oplus_{i=1}^{d} (\pi_i^* Q)(1, 1, \ldots, 1, \underset{\substack{\uparrow \\ i}}{0}, 1, \ldots, 1)$

We may conclude with a Chern class computation.

In the format $(\underbrace{2, \ldots, 2}_{d})$ the number of singular d-ples is $d!$.

List for tensors of order 3

List of the number of singular triples in the format (d_1, d_2, d_3)

d_1, d_2, d_3	$c(d_1, d_2, d_3)$	
$2, 2, 2$	6	
$2, 2, n$	8	$n \geq 3$
$2, 3, 3$	15	
$2, 3, n$	18	$n \geq 4$
$3, 3, 3$	37	
$3, 3, 4$	55	
$3, 3, n$	61	$n \geq 5$
$3, 4, 4$	104	
$3, 4, 5$	138	
$3, 4, n$	148	$n \geq 6$

The stabilization property

The output stabilizes for (a, b, c) with $c \geq a + b - 1$.
For a tensor of size $2 \times 2 \times n$ there are 6 singular vector triples for $n = 2$ and 8 singular vector triples for $n > 2$.
The format $(a, b, a + b - 1)$ is the *boundary format*, well known in hyperdeterminant theory [Gelfand-Kapranov-Zelevinsky]. It generalizes the square case.

The boundary format

For any d , the format $k_1 \times k_2 \times \ldots \times k_d$ with $k_1 = \max_j k_j$ is called *boundary format* if

$$k_1 - 1 = \sum_{i=2}^{d} (k_i - 1)$$

This is the format where it is possible to define the diagonal, as in the square case.

The basic example is given by the multiplication tensor

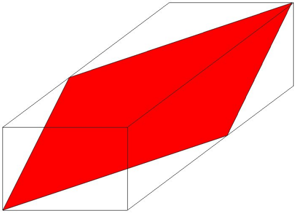
$$S^{k_2-1}\mathbb{C}^2 \otimes \ldots \otimes S^{k_d-1}\mathbb{C}^2 \to S^{\sum_{i=2}^{d}(k_i-1)}\mathbb{C}^2$$

which belongs to

$$\otimes_{i=1}^{d} \left(S^{k_i-1}\mathbb{C}^2 \right)$$

The diagonal is well defined in the boundary format case.

In the boundary format it is well defined a unique "diagonal" given by the elements $a_{i_1 \ldots i_d}$ which satisfy $i_1 = \sum_{j=2}^{d} i_j$



(indices start from zero).

The symmetric case

Theorem (Cartwright-Sturmfels)

In the symmetric case, a tensor in $S^d(\mathbb{C}^m)$ has

$$\frac{(d-1)^m - 1}{d-2}$$

singular vectors (which can be called eigenvectors).

For $d = m = 3$ the number of eigenvectors is 7. In general we compute [Oeding-O]

$$c_{m-1}(T\mathbb{P}^{m-1}(d-2)) = \frac{(d-1)^m - 1}{d-2}$$

The first proof of the formula in the symmetric case has been given by [Cartwright-Sturmfels] through the computation of a toric volume. It counts the number of eigenvectors of a symmetric tensor.

We have the same geometric interpretation with the Veronese

Advantages of the geometric interpretation with bundles

In the symmetric format $3 \times 3 \times 3$, the possible cases for the zero loci of a section are

- 7 points (with multiplicity)
- 1 line and 3 points (with multiplicity)
- 1 conic

These correspond to the possible configurations of the eigenvectors of the tensor.

The Segre-Veronese case

We consider the Segre-Veronese variety which is the embedding of the Segre variety $X = \mathbb{P}(\mathbb{R}^{m_1}) \times \ldots \times \mathbb{P}(\mathbb{R}^{m_p})$ by the line bundle $\mathcal{O}(\omega_1, \ldots, \omega_p)$.

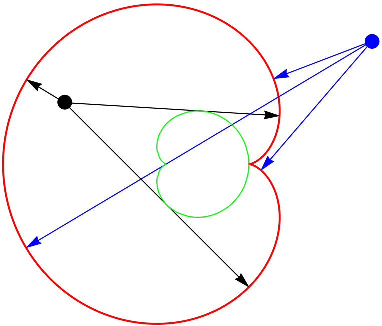
Theorem (Friedland-O)

Let $X \subset \mathbb{P}(S^{\omega_1}\mathbb{C}^{m_1} \otimes \cdots \otimes S^{\omega_p}\mathbb{C}^{m_p})$ be the Segre-Veronese variety of partially symmetric tensors of rank one. The number of singular partially symmetric d-ples is the coefficient of the monomial $z_1^{m_1-1} \cdots z_p^{m_p-1}$ in the polynomial

$$\prod_{i=1}^{p} \frac{(\widehat{z}_i)^{m_i} - z_i^{m_i}}{\widehat{z}_i - z_i} \qquad \text{where} \quad \widehat{z}_i = \left(\sum_{j=1}^{p} \omega_j z_j\right) - z_i .$$

Euclidean Distance Degree

The construction of critical points of the distance from a point u, can be generalized to any affine (real) algebraic variety.
I report some work in progress with J. Draisma, E. Horobet, B. Sturmfels and R. Thomas.
We call Euclidean Distance Degree (shortly *ED degree*) the number of critical points of $d_u = d(u, -)\colon X \to \mathbb{R}$. As before, the number of critical points does not depend on u, provided u is generic.



Duality property

We complexify, considering again the euclidean distance, and we consider the projective closure.
There is a formula for ED degree in terms of Chern classes, provided X is transversal to the quadric $\sum x_i^2 = 0$ of isotropic vectors.
One important result is

Theorem (Draisma-Horobet-O-Sturmfels-Thomas)

Let X^* be the dual variety of (the projective) X. Then

$$\textit{ED degree } X = \textit{ ED degree } X^*.$$

As a corollary, the list of the numbers of singular vectors gives the number of critical points of the distance function from the dual of the Segre variety (which is the hyperdeterminant, when it is a hypersurface).

Thanks !!