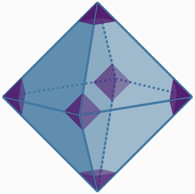
Nearest Points on Toric Varieties

Martin Helmer
University of California at Berkeley
joint work with Bernd Sturmfels



Warm-up example

Suppose we wish to learn four unknown numbers t_1, t_2, t_3, t_4 from noisy pairwise products u_{ij}. Least squares optimization problem:

$$\textbf{Minimize} \quad \begin{array}{l} (t_1 t_2 - u_{12})^2 + (t_1 t_3 - u_{13})^2 + (t_1 t_4 - u_{14})^2 + \\ (t_2 t_3 - u_{23})^2 + (t_2 t_4 - u_{24})^2 + (t_3 t_4 - u_{34})^2 \end{array}$$

Equivalent constrained optimization problem: $(x_{ij} = t_i t_j)$

$$\textbf{Minimize} \sum_{1 \leq i < j \leq 4} (x_{ij} - u_{ij})^2 \quad \textbf{subject to} \quad x_{12}x_{34} = x_{13}x_{24} = x_{14}x_{23}.$$

Warm-up example

Suppose we wish to learn four unknown numbers t_1, t_2, t_3, t_4 from noisy pairwise products u_{ij}. Least squares optimization problem:

$$\textbf{Minimize} \quad \begin{array}{l} (t_1 t_2 - u_{12})^2 + (t_1 t_3 - u_{13})^2 + (t_1 t_4 - u_{14})^2 + \\ (t_2 t_3 - u_{23})^2 + (t_2 t_4 - u_{24})^2 + (t_3 t_4 - u_{34})^2 \end{array}$$

Equivalent constrained optimization problem: $(x_{ij} = t_i t_j)$

$$\textbf{Minimize} \sum_{1 \leq i < j \leq 4} (x_{ij} - u_{ij})^2 \quad \textbf{subject to} \quad x_{12}x_{34} = x_{13}x_{24} = x_{14}x_{23}.$$

The algebraic complexity of this problem is the number of complex solutions to the critical equations. This ED degree equals **28**.

This bounds the number of real critical points, and hence the number of local minima, independently of the data $u_{12}, \ldots, u_{34}$.

The Bézout number of the unconstrained formulation is $3^6 = 729$.

Toric models

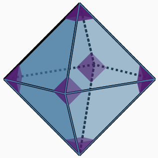
Fix an integer $d \times n$-matrix $A = (a_1, a_2, \ldots, a_n)$ of rank d with $(1, 1, \ldots, 1)$ in its row space. Each column vector a_i represents a monomial $t^{a_i} = t_1^{a_{1i}} t_2^{a_{2i}} \cdots t_d^{a_{di}}$. The *affine toric variety* $\tilde{X}_A$ is the closure of $\{(t^{a_1}, \ldots, t^{a_n}) \in \mathbb{C}^n \,:\, t \in (\mathbb{C}^*)^d\}$.

Write $X_A \subset \mathbb{P}^{n-1}$ for the *projective toric variety* with the same parametrization. Note that $\dim(X_A) = d - 1$ and $\dim(\tilde{X}_A) = d$.

The polytope $P = \mathrm{conv}(A)$ has dimension $d - 1$.

Warm-up example: $d = 4, n = 6$

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



X_A is a toric threefold of degree 4 in $\mathbb{P}^5$.

The (generic) Euclidean distance degree

Fix $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n_{>0}$ and the λ-weighted Euclidean norm

$$||x||_\lambda = \Big(\sum_{i=1}^{n} \lambda_i x_i^2 \Big)^{1/2} \quad \text{for } x \in \mathbb{R}^n.$$

Given a data vector $u \in \mathbb{R}^n$, we seek to find a real point $v \in \tilde{X}_A$ that is closest to u. This is the constrained optimization problem

$$\text{Minimize } ||u - v||_\lambda \text{ subject to } v \in \tilde{X}_A \cap \mathbb{R}^n. \tag{1}$$

This is equivalent to the unconstrained optimization problem

$$\text{Minimize } \sum_{i=1}^{n} \lambda_i (u_i - t^{a_i})^2 \text{ over all } t = (t_1, \ldots, t_d) \in \mathbb{R}^d.$$

The number of complex critical points of (1) is independent of λ and u, provided these are **generic**. This is denoted $\mathrm{EDdegree}(X_A)$ (or $\mathrm{gEDdegree}(X_A)$).

Conormal varieties, and some of their friends

Let $X \subset \mathbb{P}^{n-1}$ be a projective variety, it's conormal variety is

$$\mathrm{Con}(X) = \overline{\{(p, L) \mid p \in X_{\mathrm{reg}} \text{ and } L \supseteq T_p X\}} \subset \mathbb{P}^{n-1} \times (\mathbb{P}^{n-1})^{\vee}.$$

Let $c = \mathrm{codim}(X)$, $\mathcal{J}$ be the $(c+1) \times (c+1)$-minors of the matrix $[J(X)\ y]^T$.

The ideal of $\mathrm{Con}(X)$ is $\mathcal{K} = (I_X + \mathcal{J}) : (I_{\mathrm{Sing}(X)})^{\infty}$.

Conormal varieties, and some of their friends

Let $X \subset \mathbb{P}^{n-1}$ be a projective variety, it's conormal variety is

$$\mathrm{Con}(X) = \overline{\{(p, L) \mid p \in X_{\mathrm{reg}} \text{ and } L \supseteq T_pX\}} \subset \mathbb{P}^{n-1} \times (\mathbb{P}^{n-1})^\vee.$$

Let $c = \mathrm{codim}(X)$, $\mathcal{J}$ be the $(c+1) \times (c+1)$-minors of the matrix $[J(X)\ y]^T$.

The ideal of $\mathrm{Con}(X)$ is $\mathcal{K} = (I_X + \mathcal{J}) : (I_{\mathrm{Sing}(X)})^\infty$.

It's class in the Chow ring $A^*(\mathbb{P}^{n-1} \times (\mathbb{P}^{n-1})^\vee) \cong \mathbb{Z}[H, h]/(H^n, h^n)$ is

$$[\mathrm{Con}(X)] = \delta_0 H^{n-1} h + \cdots + \delta_{n-2} H h^{n-1}.$$

The integers $\delta_0 = \delta_0(X), \ldots, \delta_{n-2} = \delta_{n-2}(X)$ are the polar degrees of X.

In the language of books such as:

E. Miller, and B. Sturmfels: *Combinatorial commutative algebra*

the polar degrees are the multidegree of the bigraded ideal $\mathcal{K}$.

Our results

Theorem

The (generic) Euclidean distance degree of the toric variety X_A equals

$$\mathrm{EDdegree}(X_A) \quad = \quad \sum_{j=0}^{d-1} (-1)^{d-j-1} \cdot (2^{j+1} - 1) \cdot V_j,$$

where V_j is the sum of the Chern-Mather volumes of j-faces of P.

If X_A is smooth then this is the normalized lattice volume. Our theorem is based on (Aluffi, 2016) and (Matsui-Takeuchi, 2011). The latter gives a formula for the dimension and degree of the *A-discriminant* $X_A^\vee$.

Theorem

The polar degrees of the projective toric variety X_A are

$$\delta_i(X_A) \ = \ \sum_{j=i+1}^{d} (-1)^{d-j} \binom{j}{i+1} V_{j-1} \qquad \text{for } i = 0, 1, \ldots, d-1.$$

We have $\mathrm{codim}(X_A^\vee) = \min\{c : \delta_{c-1} > 0\}$ *and* $\deg(X_A^\vee) = \delta_{c-1}$.

Software

A Macaulay2 package is available at

`https://math.berkeley.edu/~mhelmer/Software/toricED`

Its input is the integer matrix A. Its output consists of

- the Chern-Mather volumes $V = (V_0, V_1, \ldots, V_{d-1})$,
- the polar degrees $\delta = (\delta_0, \delta_1, \ldots, \delta_{d-1})$,
- $\mathrm{EDdegree}(X_A) = \sum \delta_i$.

Example

Input: $A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$

Output: $V = (12, 12, 8, 4)$, $\delta = (4, 12, 8, 4)$, $\mathrm{EDdegree} = 28$.

Calculus

A Calculus student might want to minimize functions such as $L =$

$(t-11)^2+(st-1)^2+(s^2t-3)^2+(s^3t-1)^2+(s^4t-3)^2+(s^5t-1)^2+(s^6t-11)^2.$

The equations $\frac{1}{t}\frac{\partial L}{\partial s} = \frac{\partial L}{\partial t} = 0$ have **16** complex solutions.

Eight critical points are real.

Four of these are local minima, by the Second Derivative Test.

Calculus

A Calculus student might want to minimize functions such as $L =$

$(t-11)^2+(st-1)^2+(s^2t-3)^2+(s^3t-1)^2+(s^4t-3)^2+(s^5t-1)^2+(s^6t-11)^2$.

The equations $\frac{1}{t}\frac{\partial L}{\partial s} = \frac{\partial L}{\partial t} = 0$ have **16** complex solutions.

Eight critical points are real.

Four of these are local minima, by the Second Derivative Test.

This is the ED problem for the *rational normal curve*:

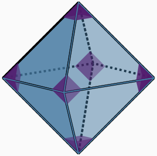
$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Our software reports $V = (2,6)$, $\delta = (10,6)$, $\mathrm{EDdegree}(X_A) = \mathbf{16}$.

The Calculus problem above finds the point on the surface $\tilde{X}_A \cap \mathbb{R}^7$ that is located closest to the data vector $u = (11,1,3,1,3,1,11)$.

Combinatorics

In a Combinatorics class one might study the lattice of faces of a polytope $P = \mathrm{conv}(A)$.



Definition

The *Euler obstruction* of a face β of P is an integer $\mathrm{Eu}(\beta)$ that measures how singular X_A is along β. It is defined recursively:

1. $\mathrm{Eu}(P) = 1$,
2. $\mathrm{Eu}(\beta) = \sum_{\substack{\alpha \text{ s.t. } \beta \text{ is a} \\ \text{proper face of } \alpha}} (-1)^{\dim(\alpha)-\dim(\beta)-1} \cdot \mu(\alpha/\beta) \cdot \mathrm{Eu}(\alpha).$

Here $\mu(\alpha/\beta)$ is the normalized relative subdiagram volume.

Characteristic classes

The *Chern-Mather volume* of a face β is the product of the normalized volume $\mathrm{Vol}(\beta)$ and the Euler obstruction $\mathrm{Eu}(\beta)$.

Our `Macaulay2` code sums these up over all j-faces of P:

$$V_j \quad = \sum_{\substack{\beta \text{ face of } P \\ \dim(\beta)=j}} \mathrm{Vol}(\beta)\mathrm{Eu}(\beta).$$

General context: MacPherson (1974) introduced *local Euler obstructions* in singularity theory. Ernström (1997) related this to polar degrees and dual varieties. The *Chern-Mather class* generalizes the total Chern class (of the tangent bundle) to singular varieties. Piene (1988) wrote the Chern-Mather class of a projective variety as an alternating sum of polar degrees. Aluffi (2016) points out the connection to ED degrees.

We recast all of this in terms of polyhedral combinatorics.

Your Algebraic Geometry teacher will prepare you to read Books like

JP. Brasselet, J. Seade, T. Suwa: *Vector Fields on Singular Varieties*,

W. Fulton: *Intersection Theory*.

Toric hypersurfaces

Let $X_A \subset \mathbb{P}^{n-1}$ be defined by one binomial equation

$$x_1^{c_1} \cdots x_r^{c_r} \;=\; x_{r+1}^{c_{r+1}} \cdots x_n^{c_n}.$$

Here $c_1, \ldots, c_n \in \mathbb{Z}_{>0}$, $\deg(X_A) = c_1 + \cdots + c_r = c_{r+1} + \cdots + c_n$.

Theorem

The i^{th} polar degree of the toric hypersurface X_A equals

$$\delta_i \;=\; \binom{n-1}{i+1} \deg(X_A) \;-\; \sum_{\tau \,:\, |\tau| = n-i-1} \min\Big(\sum_{j \in \tau \cap \{1,\ldots,r\}} c_j \,,\; \sum_{j \in \tau \cap \{r+1,\ldots,n\}} c_j \Big).$$

Example

Given a list $(u_1, u_2, u_3, u_4, u_5, u_6)$ of six real measurements, we seek to find the best approximation by the model

$$x_1^{22} x_2^{23} x_3^{64} \;=\; x_4^{26} x_5^{14} x_6^{69}.$$

$\mathrm{EDdegree}(X_A) = 1348$. *Numerical Algebraic Geometry* enables us to compute **all** critical points, and hence **all** local minima.

Discriminant

We examine the genericity condition on the weight vector $\lambda = (\lambda_1, \ldots, \lambda_n)$ that specifies the norm $||x||_\lambda = (\sum_{i=1}^n \lambda_i x_i^2)^{1/2}$. We can define the ED degree of the toric variety X_A for any positive λ. However, it may be smaller than the **generic** one:

$$\mathrm{EDdegree}_\lambda(X_A) \ \leq \ \mathbf{g}\mathrm{EDdegree}(X_A). \qquad (2)$$

The principal A-determinant variety $V(E_A) = \bigcup_{\substack{\alpha \text{ a face} \\ \text{of } \mathrm{conv}(A)}} X^\vee_{A \cap \alpha}$.

Proposition

Let $\lambda \in \mathbb{R}^n_{>0}$ be such that $\lambda \notin V(E_A)$. Then equality holds in (2).

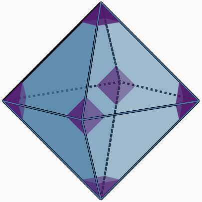
Discriminant

Remark

If all proper faces α of A are affinely independent and if all the $X^\vee_{A\cap\alpha}$ are hypersurfaces then the polynomials defining $X^\vee_A$ and $V(E_A)$ vanish at the same locus in $\mathbb{R}^n_{>0}$.

Warm-up example:

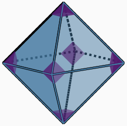
$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$



$\lambda = (1, 1, \ldots, 1)$ is not in $X^\vee_A = \left\{ \det \begin{pmatrix} 0 & \lambda_{12} & \lambda_{13} & \lambda_{14} \\ \lambda_{12} & 0 & \lambda_{23} & \lambda_{24} \\ \lambda_{13} & \lambda_{23} & 0 & \lambda_{34} \\ \lambda_{14} & \lambda_{24} & \lambda_{34} & 0 \end{pmatrix} = 0 \right\}$

The hypersimplex

Let A be the matrix whose columns are the vectors in $\{0,1\}^d$ with precisely k entries equal to 1. The polytope $P = \mathrm{conv}(A)$ is the *hypersimplex* $\Delta_{d,k}$. The toric variety X_A represents generic torus orbits on the Grassmannian of k-dim'l linear subspaces in $\mathbb{C}^d$.



Proposition

The Chern-Mather volumes for the hypersimplex $\Delta_{d,k}$ are

$$V_0 \;=\; \binom{d}{k} \cdot \min(k, d-k)$$

$$V_\ell \;=\; \sum_{i=1}^{\min(k,\ell)} \binom{d}{\ell+1}\binom{d-\ell-1}{k-i} \cdot A(\ell, i-1) \qquad \text{for } \ell = 1, \ldots, d-1.$$

For $\ell = d-1$ this formula gives the Eulerian number

$$V_{d-1} \;=\; A(d-1, k-1) \;=\; \mathrm{Vol}(\Delta_{d,k}).$$

The hypersimplex

d	k	Chern-Mather volumes	Polar degrees	EDdegree
4	2	$(12, 12, 8, 4)$	$(4, 12, 8, 4)$	28
5	2	$(20, 30, 30, 25, 11)$	$(5, 20, 40, 30, 11)$	106
6	2	$(30, 60, 80, 90, 72, 26)$	$(6, 30, 80, 120, 84, 26)$	346
6	3	$(60, 90, 120, 150, 132, 66)$	$(\mathbf{96}, 300, 480, 480, 264, 66)$	1686
7	2	$(42, 105, 175, 245, 273, 189, 57)$	$(7, 42, 140, 280, 336, 210, 57)$	1072
7	3	$(105, 210, 350, 560, 714, \ldots, 302)$	$(315{,}1302{,}2940{,}\ldots, 302)$	13441
8	2	$(56, 168, 336, 560, 784, \ldots, 120)$	$(8, 56, 224, 560, 896, \ldots 120)$	3256
8	3	$(168, 420, 840, 1610, \ldots, 1191)$	$(848, 4256, 12096, \ldots, 1191)$	86647
8	4	$(280, 560, 1120, 2240, \ldots, 2416)$	$(3816, 16016, 38976, \ldots 2416)$	236104

Table 1: The ED degree for the toric variety of the hypersimplex $\Delta_{d,k}$

If we wish to learn 7 unknown numbers $t_1, t_2, t_3, t_4, t_5, t_6, t_7$ from noisy triple products u_{ijk} by solving the critical equations we would obtain 13441 complex solutions.

Conclusion

The algebraic degree of an optimization problem governs the intrinsic complexity of representing the coordinates of the optimal solution, and of identifying all local solutions.

A priori knowledge of that degree is very useful for methods from computational algebraic geometry, either symbolic or numeric.

A great way to study characteristic classes of singular varieties.

In this project we resolved the ED problem for all toric varieties.