Algebraic Degree of Optimization Over a Variety

Olga Kuznetsova, Aalto University

Joint work with Kaie Kubjas (Aalto University) and Luca Sodomaco (Aalto University)

Abstract

We study an optimization problem with the feasible set being a real algebraic variety X and whose parametric objective function f_u is gradient-solvable with respect to the parametric data u. This class of problems includes Euclidean distance optimization as well as maximum likelihood optimization. The algebraic degree of optimization is the number of complex restricted critical points for a general data point u and reflects the complexity of a given problem. We use rational parametrization to study the algebraic degree and give some formulas using the polar classes.

Parametric optimization over a variety

$\min_x f_u(x)$ s.t. x is a point on an irreducible variety X

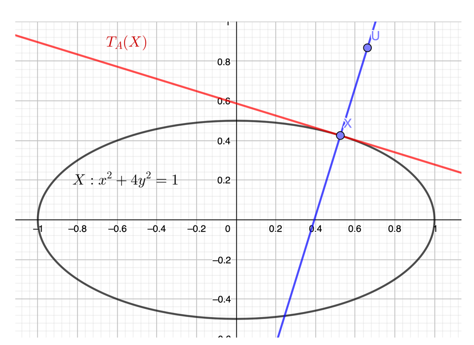

Figure 1: Euclidean distance minimization over ellipse X

Well-known examples

- (Squared) Euclidean distance minimization: $f_u = \sum_{i=1}^n (u_i - x_i)^2$
- Discrete maximum likelihood estimation $f_u = u/x$

Restricted critical points

- The argmin are in the set of restricted critical points, i.e., $x \in X$ s.t. the tangent space $T_x(X)$ is perpendicular to ∇f_u

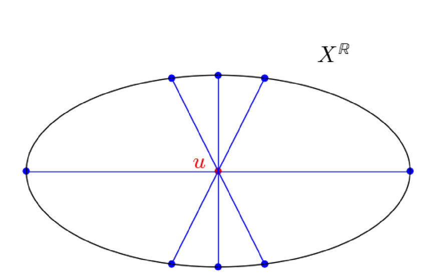

Figure 2: The 8 real critical points of $f_u = (u_1 - x_1)^4 + (u_2 - x_2)^4$ on the ellipse X.

Algebraic degree of gradient-solvable optimization

- Goal: study the algebraic complexity of optimization problems where f_u is gradient-solvable.
- Gradient-solvable: for all i, $\partial f_u/\partial x_i \in \mathbb{C}(u_i, x)$ and it is possible to write an explicit formula for u_i in radicals.
- Algebraic degree: number of complex restricted critical points for a general data point u.

Application: p-norm minimization

$$\min_x \|u - x\|_p = \left(\sum_{i=1}^n |u_i - x_i|^p\right)^{\frac{1}{p}}.$$

s.t. x is a point on an real irreducible variety $X^{\mathbb{R}}$

- If p is even, this is equivalent to minimizing $f_u = \sum_{i=1}^n (u_i - x_i)^p$
- If p is odd, then $|u_i - x_i|^p \neq (u_i - x_i)^p$
- **Idea for odd p:** find the global minimum among the solutions to the optimization problems with objective functions of the form $f_u = \pm(u_1 - x_1)^p \pm \cdots \pm (u_n - x_n)^p$
- Up to the simultaneous change of signs, there are 2^{n-1} such optimization problems

Example 1: $p = 3$ over ellipse $x_1^2 + 4x_2^2 - 1$

Need to solve the two optimization problems

$$\min_{x \in X^{\mathbb{R}}} d_u^{3,+} := [(u_1 - x_1)^3 + (u_2 - x_2)^3] \quad \text{and} \quad \min_{x \in X^{\mathbb{R}}} d_u^{3,-} := [(u_1 - x_1)^3 - (u_2 - x_2)^3]\,. \tag{1}$$

Let $u = (-6/10, 6/10)$. Each optimization problems gives two real critical points: $A_1 = (-0.228, -0.487)$, $A_2 = (0.998, 0.032)$, $A_3 = (-0.508, 0.431)$ and $A_4 = (0.997, -0.040)$ shown in Figure 4.

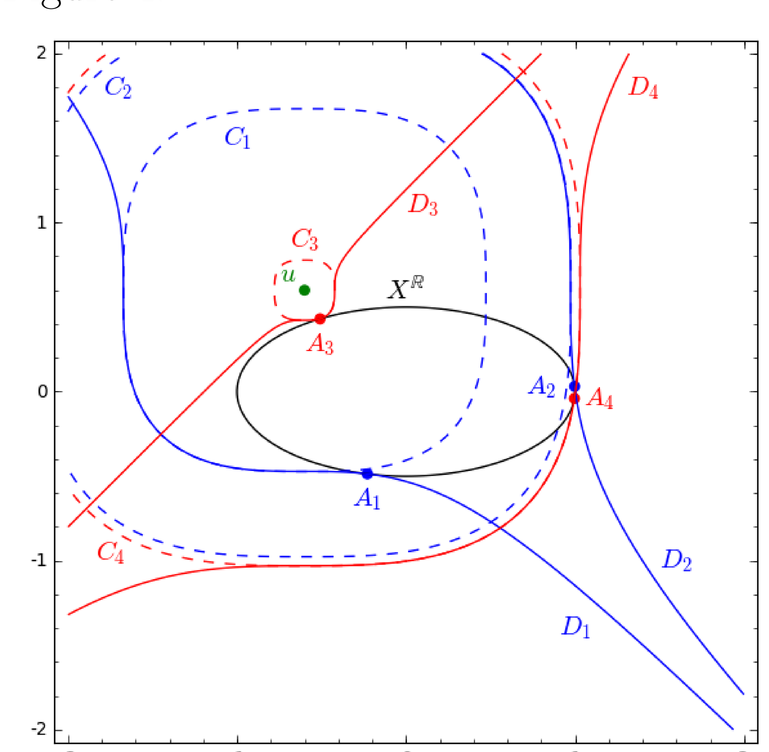

Figure 3: 3-norm minimization over ellipse with gradient-solvable functions

The dashed lines describe the boundaries of the 3-norm circles $C_i := \{x\colon \|x - u\|_3 \leq \|A_i - u\|_3\}$ for all $i \in \{1, \ldots, 4\}$. The relaxed optimization problems in (1) determine four algebraic curves

$$D_1 := \mathbb{V}(d_u^{3,+}(x) - d_u^{3,+}(A_1)),\ D_2 := \mathbb{V}(d_u^{3,+}(x) - d_u^{3,+}(A_2)),$$
$$D_3 := \mathbb{V}(d_u^{3,-}(x) - d_u^{3,-}(A_3)),\ D_4 := \mathbb{V}(d_u^{3,-}(x) - d_u^{3,-}(A_4))\,.$$

The four curves D_i are tangent to $X^{\mathbb{R}}$ at the corresponding critical points A_i and only partially overlap the corresponding 3-norm balls B_i. The global minimum of the 3-norm is attained at A_3 and is equal to $\|u - A_3\|_3 \approx 0.178$.

Main tool: optimization correspondence

- Let $I(X) = \langle g_1, \ldots, g_s\rangle \subset \mathbb{C}[x]$ and f_u gradient solvable.
- **Optimization correspondence** of X associated with f_u is the variety
$$\mathcal{F}_X := \overline{\{(x, u) \in (X \times \mathbb{C}^n) \setminus H \mid x \in X_{\text{sm}},\ x \text{ critical point of } f_u\}}\,,$$
where H is defined by $(\partial f_u/\partial x_1)_D \cdots (\partial f_u/\partial x_n)_D = 0$ and the subscript D refers to denominators.
- **Critical ideal** in $\mathbb{C}[x]$
$$\big(I(X) + \langle (c+1) \text{ minors of } J(f_u, g)\rangle\big) : \left(I(X_{\text{sing}}) \cdot \left\langle \prod_{i=1}^n \left(\frac{\partial f_u}{\partial x_i}\right)_D \right\rangle\right)$$
where
$$J(f_u, g) = \begin{pmatrix} \nabla f_u \\ \text{Jac}(g) \end{pmatrix} \cdot \text{diag}\left(\left(\frac{\partial f_u}{\partial x_1}\right)_D, \ldots, \left(\frac{\partial f_u}{\partial x_n}\right)_D\right).$$
- **Algebraic degree of X w.r.t. f_u:** cardinality of the fiber $\pi_2^{-1}(u)$ over general data point u, where $\pi_2 : \mathcal{F}_X \to \mathbb{C}^n$

Example 2: pth power of p-norm

- Let $f_u = \sum_{i=1}^n (u_i - x_i)^p$
- Critical ideal
$$\left(I(X) + \left\langle (c+1) \text{ minors of } \begin{pmatrix} \text{Jac}(g) \\ (u - x)^{p-1} \end{pmatrix} \right\rangle\right) : I(X_{\text{sing}})^{\infty}.$$
- Formula for the principal branch of u_i
$$u_i = x_i + \Big(\sum_{j=1}^c \alpha_j \text{Jac}(g)_{ji}\Big)^{1/(p-1)},$$
where α_j is a parameter and $c := \text{codim}\, X$.

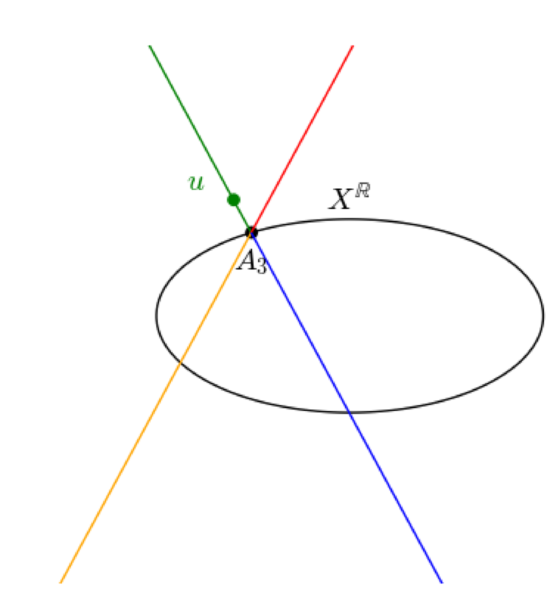

Figure 4: Fiber of point A_3 of the map $\pi_1 : \mathcal{F}_X \to X$ in Example 1

Note: the fiber of $\pi_1 : \mathcal{F}_X \to X$ is linear even though ∇f_u is non-linear in this example!

Intuition: from gradient to fibers

- Need to confirm that the notion of algebraic degree is well-defined
- To do that, need to know the dimension of $\mathcal{F}_X$
- Gradient-solvability of f_u allows to parametrize $\mathcal{F}_X$ in $\mathbb{C}(u, x)$, and hence, compute the dimension
- If $f_u = \sum_{i=1}^n (u_i - x_i)^p$, this allows to move give a 'linear' definition of critical points

s-conormal variety

- When $f_u = \sum_{i=1}^n (u_i - x_i)^2$, one has the classical conormal variety
$$\mathcal{N}_X := \overline{\{(x, y) \in \mathbb{C}^n \times \mathbb{C}^n \mid x \in X_{\text{sm}} \text{ and } y \perp T_x X\}}\,.$$
and the dual variety equals the projection on the second factor.
- If $p > 2$, define the s-conormal variety as
$$\mathcal{N}_X^{(s)} := \overline{\{(x, y) \in \mathbb{C}^n \times \mathbb{C}^n \mid x \in X_{\text{sm}} \text{ and } y^s \perp T_x X\}}$$
and, correspondingly, the s-dual variety.
- Inspired by the connection between the fibers of π_1 and ∇f_u
- **Open question:** which properties of the classical conormal variety extend to the s-conormal variety?

Formulas for p-degree

- p-degree: algebraic degree when $f_u = \sum_{i=1}^n (u_i - x_i)^p$
- Under technical conditions, the algebraic degree equals the weighted sum of polar classes of a projective variety
$$\deg_p(X) = \sum_{j=0}^{n-2} (p-1)^{n-1-j} \delta_{n-2-j}(X)\,,$$
where $\delta_i(X)$ are the coefficients of the class in the cohomology of X.

References

- Kubjas, Kaie, Olga Kuznetsova, and Luca Sodomaco. 'Algebraic degree of optimization over a variety with an application to p-norm distance degree.' arXiv preprint arXiv:2105.07785 (2021).
- Draisma, Jan, et al. "The Euclidean distance degree of an algebraic variety." Foundations of computational mathematics 16.1 (2016): 99-149.
- Sendra, J. Rafael, David Sevilla, and Carlos Villarino. 'Algebraic and algorithmic aspects of radical parametrizations.' Computer Aided Geometric Design 55 (2017): 1-14.

Contact Information

- Web: okuznetsova.com
- olga.kuznetsova@aalto.fi