Applications of singularity theory to optimization

Laurențiu Maxim

University of Wisconsin-Madison

Research School “Singularities and Applications”
Lille, 2023

Optimization–Synopsis

Given a *data* point $\underline{u} \in \mathbb{R}^n$ and an *objective function* $f_{\underline{u}} : \mathbb{R}^n \to \mathbb{R}$ depending on $\underline{u}$, a constrained optimization problem has the form

$$\min / \max \; f_{\underline{u}}(\underline{x})$$

subject to polynomial constraints

$$g_1(\underline{x}) = \cdots = g_k(\underline{x}) = 0.$$

Hence, one aims to optimize $f_{\underline{u}}$ over the real algebraic variety

$$X := V(g_1, \ldots, g_k),$$

which oftentimes is a *statistical model*.
To find the optimal solution, consider the *critical points* of $f_{\underline{u}}$ over the smooth locus X_{reg} (or find stratified critical points of $f_{\underline{u}}$ on X).

Optimization: Algebro-geometric approach

In practice, consider $f_{\underline{u}}$ and $g_1, \ldots, g_k$, as *complex* functions, i.e., regard $f_{\underline{u}}$ as a complex function defined on the complex variety (also denoted X) defined by the Zariski closure of X in $\mathbb{C}^n$.
Can (and will!) assume X *irreducible*, and require $f_{\underline{u}}$ to be *holomorphic* and have certain good properties.
For general $\underline{u}$, the number of complex critical points of $f_{\underline{u}}$ on X_{reg} is independent of $\underline{u}$, and is called the algebraic degree of the given optimization problem. (It can be realized as the degree of a map.)
This measures the algebraic complexity of the optimal solution of the optimization problem, and it is a good indicator of the running time needed to solve the problem exactly.

Optimization: examples

The main optimization problems considered in these lectures are:

(a) nearest point problem (NPP) / ED optimization: $X \subset \mathbb{R}^n$ is an algebraic model (i.e., defined by polynomial equations), and

$$f_{\underline{u}}(\underline{x}) = d_{\underline{u}}(\underline{x}) = \sum_{i=1}^{n} (x_i - u_i)^2$$

is the squared Euclidean distance from a general data point $\underline{u} \in \mathbb{R}^n$ to X.

(b) maximum likelihood estimation (MLE) / ML optimization: X is a statistical model (family of probability distributions) and

$$f_{\underline{u}}(\underline{x}) = \ell_{\underline{u}}(\underline{x}) = \prod_{i=1}^{n} p_i(\underline{x})^{u_i}$$

is the likelihood function associated to the data point $\underline{u} = (u_1, \ldots, u_n)$.

Optimization: examples

(c) linear optimization: $X \subset \mathbb{R}^n$ is an algebraic model and

$$\ell_{\underline{u}}(\underline{x}) = \sum_{i=1}^{n} u_i x_i$$

is (the restriction to X_{reg} of) a general linear function.

0. *Preliminaries*

Whitney stratification

Let X be a complex algebraic variety. Then X admits a *Whitney stratification* (i.e., a partition $\mathcal{S}$ into locally closed nonsingular subvarieties (*strata*), along which X is topologically equisingular).

Example

A smooth (irreducible) complex algebraic variety X is Whitney stratified with only one stratum: X.

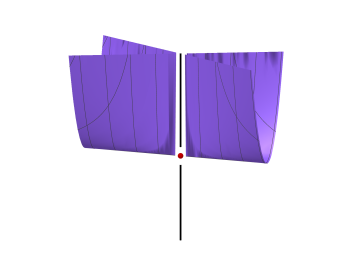
Example

If X is an irreducible complex algebraic variety whose singular locus is a finite set of points $s_1, \dots, s_r$, then a Whitney stratification of X can be given with strata

$$\{X_{\mathrm{reg}}, \{s_1\}, \dots, \{s_r\}\}.$$

Example (Whitney umbrella)

Let $X = \{x^2 = zy^2\} \subset \mathbb{C}^3$. The singular locus of X is the z-axis, but the origin is "more singular" than any other point on the z-axis.



A Whitney stratification of X has strata

$$V_1 = X \setminus \{z - \text{axis}\}, \quad V_2 = \{(0,0,z) \mid z \neq 0\}, \quad V_3 = \{(0,0,0)\}.$$

Constructible functions

Let X be a complex algebraic variety with Whitney stratification $\mathcal{S}$.
A function $\varphi : X \to \mathbb{Z}$ is *$\mathcal{S}$-constructible* if φ is constant along each stratum $S \in \mathcal{S}$.
The Euler characteristic of an $\mathcal{S}$-constructible function φ is:

$$\chi(\varphi) := \sum_{S \in \mathcal{S}} \chi(S) \cdot \varphi(S),$$

with $\varphi(S)$ the (constant) value of φ on the stratum S.
For example,

$$\chi(1_X) = \sum_{S \in \mathcal{S}} \chi(S) = \chi(X).$$

MacPherson's local Euler obstruction

The *local Euler obstruction* $Eu_X : X \to \mathbb{Z}$ is an essential ingredient in MacPherson's definition of Chern classes for singular varieties.

(1) Eu_X is $\mathcal{S}$-constructible for any fixed Whitney stratification $\mathcal{S}$ of X.

(2) If $x \in X$ is a smooth point, then $Eu_X(x) = 1$.

(3) If X is a curve, then $Eu_X(x)$ is the multiplicity of X at x.

(4) If (X, x) is an isolated hypersurface singularity germ, then $Eu_X(x) = \chi(L_{X,x})$, where $L_{X,x}$ is the complex link of x in X.

(5) Eu_X is an analytic invariant.

Example (Nodal curve)

Let $X = \{xy = 0\} \in \mathbb{C}^2$. Then $\mathrm{Sing}(X) = \{(0,0)\}$ with multiplicity 2. A Whitney stratification of X has strata $V_1 = X \setminus \{(0,0)\}$ and $V_2 = \{(0,0)\}$. So $Eu_X|_{V_1} = 1$ and $Eu_X|_{V_2} = 2$.

Example (Whitney umbrella)

Let $X = \{x^2 = zy^2\} \subset \mathbb{C}^3$, with $\mathrm{Sing}(X)$=$\{z\text{-axis}\}$, and Whitney stratification with strata

$$V_1 = X \setminus \{z - \text{axis}\}, \quad V_2 = \{(0,0,z) \mid z \neq 0\}, \quad V_3 = \{(0,0,0)\}.$$

The local Euler obstruction function Eu_X has values 1, 2 and 1 along the strata V_1, V_2 and V_3, respectively (Gonzalez-Sprinberg).

Constructible functions form a group

Let X be a complex algebraic variety with Whitney stratification $\mathcal{S}$.
Denote by $CF(X)$ the free abelian group of constructible functions on X.
A basis of $CF(X)$ is given by $\{1_{\bar{V}} \mid V \in \mathcal{S}\}$.
Another basis consists of $\{Eu_{\bar{V}} \mid V \in \mathcal{S}\}$.
MacPherson Chern class transformation is a homomorphism

$$c_* : CF(X) \to A_*(X),$$

so that if X is smooth $c_*(1_X) = c^*(TX) \cap [X]$.
$c_*(1_X) \in A_*(X)$ is called the Chern class of X.
$c_*(Eu_X) =: c_{Ma}(X) \in A_*(X)$ is called the Chern-Mather class of X.

I. *Applications to NPP / ED optimization*

Nearest Point Problem: given an algebraic model $X \subset \mathbb{R}^n$ and a generic data point $\underline{u} \in \mathbb{R}^n$, find a *nearest point* $u^* \in X$ to $\underline{u}$. Equivalently: minimize over X_{reg} the squared distance function

$$d_{\underline{u}}(\underline{x}) = \sum_{i=1}^{n} (x_i - u_i)^2.$$

Algebraic degree for NPP is called the Euclidean distance (ED) degree of X, denoted $EDdeg(X)$. It was formally introduced by Draisma-Horobeț-Ottaviani-Sturmfels-Thomas (2014), as an algebraic measure of complexity of the nearest point problem.

A general upper bound on the ED degree can be given in terms of the defining polynomials.

Proposition (DHOST)

Let $X \subset \mathbb{C}^n$ be a variety of codim. c, cut out by polynomials $g_1, g_2, \ldots, g_c, \ldots, g_k$ of degrees $d_1 \geq d_2 \geq \cdots \geq d_c \geq \cdots \geq d_k$. Then

$$EDdeg(X) \leq$$

$$\leq d_1 d_2 \cdots d_c \cdot \sum_{i_1+i_2+\cdots+i_c \leq n-c} (d_1-1)^{i_1}(d_2-1)^{i_2}\cdots(d_c-1)^{i_c}$$

Equality holds when X is a general complete intersection of codimension c (hence $c = k$).

Why "degree"?

For an irreducible closed variety $X \subset \mathbb{C}^n$ of codimension c, consider the ED correspondence $\mathcal{E}_X$, i.e., the topological closure in $\mathbb{C}^n \times \mathbb{C}^n$ of

$$\{(\underline{x}, \underline{u}) \in X_{\mathrm{reg}} \times \mathbb{C}^n \mid \underline{x} \text{ is a critical point of } d_{\underline{u}}|_{X_{\mathrm{reg}}}\}$$

The first projection $\pi_1 : \mathcal{E}_X \to X$ is an affine vector bundle of rank c over X_{reg}, whereas for general data points $\underline{u} \in \mathbb{C}^n$ the second projection $\pi_2 : \mathcal{E}_X \to \mathbb{C}^n$ has finite fibers $\pi_2^{-1}(\underline{u})$ of cardinality equal to $EDdeg(X)$.

Example: Low-rank approximation

Fix positive integers $r \leq s \leq t$ and set $n = st$.
Model:

$$X_r := \left\{ A = [a_{ij}] \in \mathbb{R}^{s \times t} \mid \operatorname{rank}(A) \leq r \right\} \subset \mathbb{R}^n.$$

Data point: general $U = [u_{ij}]_{s \times t} \in \mathbb{R}^{s \times t} = \mathbb{R}^n$.
Singular Value Decomposition:

$$U = T_1 \cdot \operatorname{diag}(\sigma_1, \ldots, \sigma_s) \cdot T_2,$$

where $\sigma_1 > \cdots > \sigma_s > 0$ are the *singular values* of a general matrix U, and T_1, T_2 are orthogonal matrices.
Eckart-Young Theorem: the matrix of rank $\leq r$ closest to U is:

$$U^* = T_1 \cdot \operatorname{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) \cdot T_2 \in X_r.$$

$EDdeg(X_r) = \binom{s}{r}$.

Triangulation problem in computer vision

The ED degree was motivated by the triangulation problem in computer vision.

Triangulation problem: Given an object (world point) in 3D space, determine its position from its 2D projections in $n \geq 2$ cameras in general position.

Many practical applications: tourism, robotics, GPS, autonomous driving, cloud modeling, filmmaking, etc.

Building Rome in one day: https://grail.cs.washington.edu/rome/
"[...] we consider the problem of reconstructing entire cities from images harvested from the web. Our aim is to build a [...] system that [...] matches these images to find common points and uses this information to compute the three dimensional structure of the city and the pose of the cameras that captured these images. *All this to be done in a day*."

Triangulation problem and nearest point problem

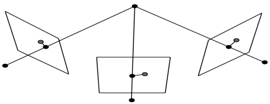
The triangulation problem is *in theory* trivial to solve: if the image points are given with "infinite precision", then two cameras suffice to determine the 3D point (by triangulation).

In practice, "noise" (pixelation, lens distortion, etc.) leads to inaccuracies in the measured image coordinates.

NPP: find the "closest" world point to the image data from multiple camera projections.

Model: The space of all possible n-tuples of such 2D projections is called the affine (real) multiview variety $X_n \subset \mathbb{R}^{2n}$.

Data point:
$\underline{u} = (u_1, \ldots, u_{2n}) \in \mathbb{R}^{2n}$,
corresponding to n noisy 2D images of a 3D point.



Euclidean distance degree of affine multiview variety

Conjecture (DHOST, 2014)

$$EDdeg(X_n) = \frac{9}{2}n^3 - \frac{21}{2}n^2 + 8n - 4,$$

where X_n is the affine multiview variety of n cameras.

The conjecture was based on the numerical evidence for $n \leq 7$ (Stewénius-Schaffalitzky-Nistér, 2005), and it was proved by M.-Rodriguez-Wang (2020).

The ED degree of the affine multiview variety X_n was studied by Harris-Lowengrub (2017) via characteristic classes. They obtained an upper bound of $EDdeg(X_n)$:

Theorem (Harris-Lowengrub, 2017)

$$EDdeg(X_n) \leq 6n^3 - 15n^2 + 11n - 4.$$

Euclidean distance degree of an affine variety

Theorem (M.-Rodriguez-Wang, 2020)

Let $X \subset \mathbb{C}^n$ be an irreducible closed subvariety. If X is smooth, then for general $u = (u_0, u_1, \dots, u_n) \in \mathbb{C}^{n+1}$,

$$EDdeg(X) = (-1)^{\dim_{\mathbb{C}} X} \chi(X \setminus Q_u),$$

where $Q_u = \{\sum_{i=1}^n (x_i - u_i)^2 = u_0\} \subset \mathbb{C}^n$.
For arbitrary X,

$$EDdeg(X) = (-1)^{\dim_{\mathbb{C}} X} \chi(Eu_X|_{\mathbb{C}^n \setminus Q_u}).$$

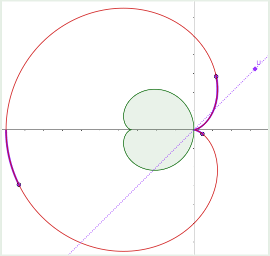
Example ($X = \mathbb{C}$)

$EDdeg(X) = -\chi(X \setminus Q_u) = -(\chi(X) - \chi(X \cap Q_u)) = -(1-2) = 1.$

Example (Cardioid)

Consider the cardioid curve $X \subset \mathbb{C}^2$ defined by

$$(x^2 + y^2 + x)^2 = x^2 + y^2.$$



X has a unique singular point of multiplicity 2 at the origin in $\mathbb{C}^2$, and $X_{\mathrm{reg}} \cong \mathbb{CP}^1 \setminus \{3 \text{ points}\}$. For generic $\underline{u}$, $X \cap Q_{\underline{u}}$ consists of 4 smooth points. Then

$$EDdeg(X) = -\chi(Eu_{X \setminus Q_{\underline{u}}}) = -(2-5) = 3.$$

What's behind the ED degree formula?

The proof of the formula for *EDdeg* amounts to "lineariazing" the problem and using an affine Lefschetz-type theorem.
Consider the closed embedding

$$i : \mathbb{C}^n \hookrightarrow \mathbb{C}^{n+1} , \quad (z_1, \dots, z_n) \mapsto (z_1^2 + \cdots + z_n^2, z_1, \dots, z_n).$$

If $w_0, \dots, w_n$ are the coordinates of $\mathbb{C}^{n+1}$, then the function $\sum_{1 \leq i \leq n}(z_i - u_i)^2 + u_0$ on $\mathbb{C}^n$ is the pullback of the function

$$w_0 + \sum_{1 \leq i \leq n} -2u_i w_i + \sum_{1 \leq i \leq n} u_i^2 + u_0$$

on $\mathbb{C}^{n+1}$. The computation of $EDdeg(X)$ follows by applying the following affine Lefschetz-type Theorem to $i(X) \subset \mathbb{C}^{n+1}$.

Theorem (Seade-Tibar-Verjovsky, 2005)

Let $X \subset \mathbb{C}^N$ be an affine variety. Let $\ell : \mathbb{C}^N \to \mathbb{C}$ be a general linear function, with $H_c = \{\ell = c\} \in \mathbb{C}^N$ for a general $c \in \mathbb{C}$. Then the number of critical points of $\ell|_{X_{\mathrm{reg}}}$ equals $(-1)^{\dim_{\mathbb{C}} X} \chi(Eu_{X \setminus H_c})$.

The DHOST multiview conjecture

This topological formula can be used to confirm the DHOST conjecture:

Theorem (M.-Rodriguez-Wang, 2020)

The ED degree of the affine multiview variety $X_n \subset \mathbb{C}^{2n}$ satisfies:

$$EDdeg(X_n) = -\chi(X_n \setminus Q_u) = \frac{9}{2}n^3 - \frac{21}{2}n^2 + 8n - 4.$$

The computation of $\chi(X_n \setminus Q_u)$ is quite involved and it relies on topological and algebraic techniques from Singularity theory (e.g., Milnor fibration, vanishing cycles, etc).

Projective Euclidean distance degree

Many models in data science, engineering and other applied fields are realized as *affine cones* (defined by homogeneous polynomials), so it is natural to consider such models as *projective varieties*. Examples of such models occur in (structured) low rank matrix approximation, low rank tensor approximation, formation shape control, and all across algebraic statistics.

Example

The variety X_r of $s \times t$ matrices of rank $\leq r$ is an affine cone.

Projective Euclidean distance degree

Definition

If $Y \subset \mathbb{P}^n$ is an irreducible complex projective variety, define the projective Euclidean distance degree of Y by

$$pEDdeg(Y) := EDdeg(C(Y)),$$

where $C(Y)$ is the affine cone of Y in $\mathbb{C}^{n+1}$.

Problem

Compute $pEDdeg(Y)$ in terms of the topology of Y.

The ED degree $pEDdeg(Y)$ of a *smooth* projective variety was related to the Chern classes of the projective variety by Catanese-Trifogli (2000), D-H-O-S-T (2014), Aluffi-Harris (2017).

Theorem (Aluffi-Harris, 2017)

Let $Y \subset \mathbb{P}^n$ be a smooth complex projective variety. Then

$$pEDdeg(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Y \setminus (Q \cup H))$$

where H is a general hyperplane in $\mathbb{P}^n$, and Q is the isotropic quadric $\{z_0^2 + \cdots + z_n^2 = 0\}$.

This is a generalization of a result from D-H-O-S-T (2014), in which Y was assumed to intersect Q *transversally*. Moreover, Aluffi-Harris conjectured that in the above notations one has:

Theorem (M.-Rodriguez-Wang, 2019)

Let $Y \subset \mathbb{P}^n$ be an irreducible complex projective variety. Then

$$pEDdeg(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Eu_{Y \setminus (Q \cup H)}).$$

Recall: for $Y \subset \mathbb{P}^n$, with $Q = \{z_0^2 + \cdots + z_n^2 = 0\}$ the isotropic quadric and H a general hyperplane,

$$pEDdeg(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Eu_{Y \setminus (Q \cup H)})$$

Example (Nodal curve)

Let $Y = \{x_0^2 x_2 - x_1^2(x_1 + x_2) = 0\} \subset \mathbb{CP}^2$. It has only one singular point $p = [0:0:1]$. Then Eu_Y equals 1 on the smooth locus Y_{reg} of Y, and $Eu_Y(p) = 2$. Y intersects Q transversally at 6 points, and it intersects a generic hyperplane H at 3 points. Moreover, $Y_{\mathrm{reg}} \cong \mathbb{C}^*$. By inclusion-exclusion, get $\chi(Y_{\mathrm{reg}} \setminus (Q \cup H)) = -9$. So $pEDdeg(Y) = (-1) \cdot [(-9) + 2] = 7$.

Recall: for $Y \subset \mathbb{P}^n$, with $Q = \{z_0^2 + \cdots + z_n^2 = 0\}$ the isotropic quadric and H a general hyperplane,

$$pEDdeg(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Eu_{Y \setminus (Q \cup H)})$$

Example (Whitney umbrella)

Consider the (projective) Whitney umbrella $Y = \{x_0^2 x_1 - x_2 x_3^2 = 0\} \subset \mathbb{CP}^3$. Then $\mathrm{Sing}(Y) = \{x_0 = x_3 = 0\}$, and Y has a Whitney stratification with strata: $S_3 := \{[0:1:0:0], [0:0:1:0]\}$, $S_2 = \{x_0 = x_3 = 0\} \setminus S_3$, and $S_1 = Y \setminus \{x_0 = x_3 = 0\}$. One knows that Eu_Y takes the values 1, 2 and 1 along S_1, S_2 and S_3, respectively. Therefore, if $U := \mathbb{P}^3 \setminus (Q \cup H)$, then

$$\chi(Eu_{Y \setminus (Q \cup H)}) = \chi(Y \cap U) + \chi(S_2 \cap U).$$

By inclusion-exclusion, $\chi(Y \cap U) = 13$ and $\chi(S_2 \cap U) = -3$. So $pEDdeg(Y) = 10$.

Recall: for $Y \subset \mathbb{P}^n$, with $Q = \{z_0^2 + \cdots + z_n^2 = 0\}$ the isotropic quadric and H a general hyperplane,

$$pEDdeg(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Eu_{Y \setminus (Q \cup H)})$$

Example (Toric quartic surface)

Let $Y = \{x_0^3 x_1 - x_2 x_3^3 = 0\} \subset \mathbb{CP}^3$. Then Y has a Whitney stratification with three strata: $S_3 := \{[0:1:0:0], [0:0:1:0]\}$, $S_2 := \{x_0 = x_3 = 0\} \setminus S_3$ and $S_1 = Y \setminus \{x_0 = x_3 = 0\}$. Eu_Y takes values 1, 3 and 1 along S_1, S_2 and S_3, respectively. Therefore, if $U := \mathbb{P}^3 \setminus (Q \cup H)$, then

$$\chi(Eu_{Y \setminus (Q \cup H)}) = \chi(Y \cap U) + 2\chi(S_2 \cap U).$$

By inclusion-exclusion, one gets $\chi(Y \cap U) = 16$ and $\chi(S_2 \cap U) = -3$. Hence $pEDdeg(Y) = 10$.

Defect of ED degree

Remark

pEDdeg(Y) *is difficult to compute even if* $Y \subset \mathbb{P}^n$ *is smooth, since* Y *and* Q *may intersect nontransversally in* $\mathbb{P}^n$*.*
Idea: perturb the objective (i.e., squared distance) function to create a transversal intersection.

Definition

The $\underline{\lambda}$-Euclidean distance (ED) degree $EDdeg_{\underline{\lambda}}(X)$ of an *affine* variety $X \subset \mathbb{C}^n$ is the number of critical points of

$$d_{\underline{u}}^{\underline{\lambda}}(\underline{x}) = \sum_{i=1}^{n} \lambda_i (x_i - u_i)^2 \ , \quad \underline{\lambda} = (\lambda_1, \ldots, \lambda_n)$$

on the smooth locus X_{reg} of X (for general $\underline{u} \in \mathbb{C}^n$).
If $Y \subset \mathbb{P}^n$ is an irreducible complex projective variety, define the projective $\underline{\lambda}$-Euclidean distance degree of Y by

$$pEDdeg_{\underline{\lambda}}(Y) := EDdeg_{\underline{\lambda}}(C(Y)),$$

where $C(Y)$ is the affine cone of Y in $\mathbb{C}^{n+1}$.
If $\underline{\lambda}$ is generic, get the generic ED degree.

Theorem (M.-Rodriguez-Wang, 2019)

Let $Y \subset \mathbb{P}^n$ be an irreducible complex projective variety. Then

$$pEDdeg_{\underline{\lambda}}(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Eu_{Y \setminus (Q_{\underline{\lambda}} \cup H)}),$$

where $Q_{\underline{\lambda}} := \{\lambda_0 x_0^2 + \cdots + \lambda_n x_n^2 = 0\}$ and H is a general hyperplane in $\mathbb{P}^n$. In particular, if Y is smooth, then

$$pEDdeg_{\underline{\lambda}}(Y) = (-1)^{\dim_{\mathbb{C}} Y} \chi(Y \setminus (Q_{\underline{\lambda}} \cup H)).$$

Remark

For generic $\underline{\lambda}$, the quadric $Q_{\underline{\lambda}}$ intersects Y transversally in $\mathbb{P}^n$, and the computation of the generic projective ED degree $pEDdeg_{\underline{\lambda}}(Y)$ is more manageable (DHOST, Helmer-Sturmfels, Aluffi-Harris, etc).

Definition (Defect of ED degree)

If $Y \subset \mathbb{P}^n$ is an irreducible projective variety and $\underline{\lambda}$ is generic, the *defect of Euclidean distance degree* of Y is defined as:

$$\mathit{EDdefect}(Y) := \mathit{pEDdeg}_{\underline{\lambda}}(Y) - \mathit{pEDdeg}(Y).$$

Problem

Compute EDdefect(Y) in terms of the topology of $\mathrm{Sing}(Y \cap Q)$, the locus where Y intersects Q nontransversally.

Example

The projective Whitney umbrella $Y = \{x_0^2 x_1 - x_2 x_3^2 = 0\} \subset \mathbb{CP}^3$ is transversal to the isotropic quadric Q, so its projective Euclidean distance degree coincides with the generic Euclidean distance degree. In particular, $EDdefect(Y) = 0$.

Example

Recall that the projective ED degree of the quartic surface $Y = \{x_0^3 x_1 - x_2 x_3^3 = 0\} \subset \mathbb{CP}^3$ equals 10. Moreover, the generic ED degree is equal to 14 (Aluffi-Harris, Helmer-Sturmfels). Therefore, $EDdefect(Y) = 4$.

Theorem (M.-Rodriguez-Wang, 2019)

Let $Y \subset \mathbb{P}^n$ be a smooth irreducible variety, with $Y \not\subseteq Q$, and let $Z = \mathrm{Sing}(Y \cap Q)$. Let $\mathcal{V}$ be the collection of strata of a Whitney stratification of $Y \cap Q$ which are contained in Z, and choose $\underline{\lambda}$ generic. Then:

$$EDdefect(Y) = \sum_{V \in \mathcal{V}} \alpha_V \cdot pEDdeg_{\underline{\lambda}}(\bar{V}),$$

where, for any stratum $V \in \mathcal{V}$,

$$\alpha_V = (-1)^{\mathrm{codim}_{Y\cap Q} V} \cdot \left(\mu_V - \sum_{\{S \mid V < S\}} \chi_c(L_{V,S}) \cdot \mu_S \right),$$

with $\mu_V = \chi(\widetilde{H}^*(F_V; \mathbb{Q}))$ the Euler characteristic of the reduced cohomology of the Milnor fiber F_V of the hypersurface $Y \cap Q \subset Y$ at some point in V, and $L_{V,S}$ the complex link of a pair of distinct strata (V, S) with $V \subset \bar{S}$.

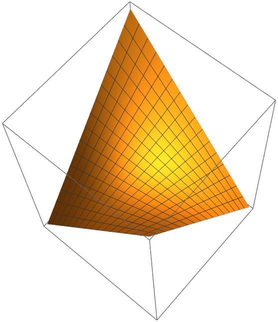
The proof relies on the theory of vanishing cycles, adapted to a pencil of quadrics on Y.

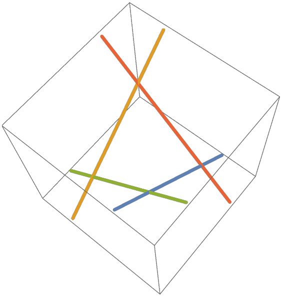
Corollary (Aluffi-Harris)

If $Z = \mathrm{Sing}(Y \cap Q)$ has only isolated singularities, then

$$EDdefect(Y) = \sum_{x \in Z} \mu_x,$$

where μ_x is the Milnor number of the IHS $(Y \cap Q, x)$ in Y.





Example (2×2 matrices of rank 1)

- $Y = \{x_0x_3 - x_1x_2 = 0\} \subset \mathbb{P}^3$, $Q = \{\sum_{i=0}^{3} x_i^2 = 0\}$.
- $Y \cap Q$ consists of 4 lines with 4 isolated double point singularities (each having $\mu = 1$).
- $EDdefect(Y) = 4$.
- (Here, $pEDdeg(Y) = 2$ and $pEDdeg_{\underline{\lambda}}(Y) = 6$ for generic $\underline{\lambda}$.)

Corollary

Assume that $Z = \mathrm{Sing}(Y \cap Q)$ is connected and $Y \cap Q$ is equisingular along Z. Then:

$$\mathit{EDdefect}(Y) = \mu \cdot \mathit{pEDdeg}_{\underline{\lambda}}(Z),$$

where μ is the Milnor number of the isolated transversal singularity at some point $x \in Z$ (i.e., the Milnor number of the isolated hypersurface singularity in a normal slice to Z at x).

Remark

- *Computing the ED degree defect of $Y \subset \mathbb{P}^n$ yields a formula for pEDdeg(Y) only in terms of generic ED degrees (which are easier to compute).*
- *Applications in structured low-rank approximation (duality conjecture of Ottaviani-Spaenlehauser-Sturmfels).*
- *Similar methods apply to other objective functions, leading to the computation of other important invariants in algebraic statistics (e.g., maximum likelihood (ML) degree).*

Remark

NPP for non-generic data point $\underline{u}$ can be studied via a generic linear morsification $d_{\underline{u}}(\underline{x}) - t\ell(\underline{x}) \simeq d_{\underline{u}+t\underline{\epsilon}}(\underline{x})$, with $\underline{\epsilon} \in \mathbb{C}^n$ generic. One then studies "limits" of Morse points of $d_{\underline{u}}(\underline{x}) - t\ell(\underline{x})$ as $t \to 0$ (M.-Rodriguez-Wang, M.-Tibăr).

II. *Applications to MLE / ML optimization*

MLE / ML optimization

Example (Biased coin)

Let θ be the probability of observing tail (T) on a biased coin.
Experiment: flip a biased coin twice and record the outcomes.
$p_i(\theta) :=$ probability of observing i heads (H), $i = 0, 1, 2$.
$p_0(\theta) = \theta^2$, $p_1(\theta) = 2\theta(1-\theta)$, $p_2(\theta) = (1-\theta)^2$.
Repeat the experiment a number of times.
$u_i :=$ the number of times i heads were observed, $i = 0, 1, 2$.
MLE problem: estimate θ by maximizing the likelihood function
$\ell_{\underline{u}} = p_0(\theta)^{u_0} p_1(\theta)^{u_1} p_2(\theta)^{u_2}$.
Solve $d \log \ell_{\underline{u}} = 0$ for θ, unique solution $\hat{\theta} = \frac{2u_0+u_1}{2u_0+2u_1+2u_2}$.
The distribution p lives in a statistical model $X = V(g)$ defined by
$g(p_0, p_1, p_2) = 4p_0p_2 - p_1^2$. This is the *Hardy-Weinberg curve*
(important in population genetics).

In general, suppose $X \subset \Delta_n$ is a family of probability distributions, where Δ_n is the n-dimensional *probability simplex*:

$$\Delta_n = \{\underline{p} = (p_0, \dots, p_n) \in \mathbb{R}^{n+1} \mid p_i > 0, \sum_i p_i = 1\}.$$

Given N i.i.d. samples, summarize the outcome in the data vector $\underline{u} = (u_0, \dots, u_n)$, with $N = \sum_i u_i$ and $u_i :=$ the number of times state i was observed. Let p_i be the probability of observing state i.
MLE / ML optimization: maximize the likelihood function

$$\ell_{\underline{u}}(\underline{p}) := \prod_{i=0}^{n} p_i^{u_i},$$

subject to $p \in X$ (a parametrization of X may not be available).
Algebraic degree of ML optimization, i.e., the number of critical points of $\ell_{\underline{u}}$ on X_{reg} ($\subset (\mathbb{C}^*)^{n+1}$), is the ML degree $MLdeg(X)$.
One allows $\underline{u} \in \mathbb{C}^{n+1}$, so $MLdeg(X)$ counts the degeneration points of $d \log \ell_{\underline{u}}$ on X_{reg}.
(Introduced by Catanese-Hosten-Khetan-Sturmfels in 2006.)

Likelihood geometry in $\mathbb{P}^n_{\mathbb{C}}$

Let $p_0, \ldots, p_n$ be coordinates (representing probabilities).
Observed data vector: $\underline{u} = (u_0, \ldots, u_n)$, $u_i :=$ number of samples in state i.
Likelihood function on $\mathbb{P}^n_{\mathbb{C}}$:

$$\ell_{\underline{u}}(\underline{p}) = \frac{p_0^{u_0} p_1^{u_1} \cdots p_n^{u_n}}{(p_0 + \cdots + p_n)^{u_0 + \cdots + u_n}}.$$

$\ell_{\underline{u}}$ is a rational function on $\mathbb{P}^n_{\mathbb{C}}$, regular on $\mathbb{P}^n_{\mathbb{C}} \setminus \mathcal{H}$, where

$$\mathcal{H} := \{p_0 \cdots p_n (p_0 + \cdots + p_n) = 0\}.$$

Consider the restriction of $\ell_{\underline{u}}$ to a closed irreducible subvariety $X \subset \mathbb{P}^n_{\mathbb{C}}$ (e.g., defined over $\mathbb{R}$), so that $X^\circ := X \setminus \mathcal{H} \neq \emptyset$.
When X is a statistical model, the ML problem is to maximize $\ell_{\underline{u}}$ over $X \cap \Delta_n$.
Note: X° is a very affine variety (a closed subvariety of $(\mathbb{C}^*)^{n+1}$).
$MLdeg(X) := \#$ of critical points of $\ell_{\underline{u}}$ on $X_{\mathrm{reg}} \setminus \mathcal{H} = X^\circ_{\mathrm{reg}}$.

Geometric and topological interpretation of MLdeg

As before, let $X \subset \mathbb{P}^n$ with $X^\circ := X \setminus \mathcal{H} \neq \emptyset$, where $\mathcal{H} := \{p_0 \cdots p_n(p_0 + \cdots + p_n) = 0\}$.

Theorem (Huh, 2013)

If X° is smooth, $d = \dim_{\mathbb{C}} X$, then $MLdeg(X) = (-1)^d \cdot \chi(X^\circ)$.

More generally,

Theorem (Rodriguez-Wang, 2017)

If $d = \dim_{\mathbb{C}} X$, then $MLdeg(X) = (-1)^d \cdot \chi(Eu_{X^\circ})$.

ML bidegrees

Likelihood correspondence $\mathcal{L}_X$ is the closure in $\mathbb{P}^n \times \mathbb{P}^{n+1}$ of the set

$$\{(\underline{p}, \underline{u}) \in X^\circ_{\mathrm{reg}} \times \mathbb{C}^{n+1} \mid \underline{p} \text{ is a critical point of } \ell_{\underline{u}}|_{X^\circ_{\mathrm{reg}}}\}$$

Definition (ML bidegrees)

The *i*-th ML bidegree b_i of X, $i = 0, \ldots, d = \dim_{\mathbb{C}} X$, is given by:

$$[\mathcal{L}_X] = \sum_{i=0}^{d} b_i [\mathbb{P}^i \times \mathbb{P}^{n+1-i}] \in A_*(\mathbb{P}^n \times \mathbb{P}^{n+1}).$$

♣ b_i is the degree of the preimage of $\mathcal{L}_X \to \mathbb{P}^{n+1}$ over a general codim i linear subspace.

♣ $b_0 = MLdeg(X)$.

♣ $b_d = \deg(X)$.

Theorem (M.-Rodriguez-Wang-Wu, 2022)

Let $X \subset \mathbb{P}^n$ be a d-dimensional closed irreducible subvariety with $X^\circ = X \setminus \mathcal{H} \neq \emptyset$. Then the total Chern-Mather class of X° is:

$$c_{Ma}(X^\circ) = \sum_{i=0}^{d} (-1)^{d-i} b_i [\mathbb{P}^i] \in A_*(\mathbb{P}^n).$$

Here, $c_{Ma}(X^\circ) := c_*(Eu_{X^\circ})$, with $c_* : CF(\mathbb{P}^n) \to A_*(\mathbb{P}^n)$ the MacPherson-Chern class transformation.

At degree zero,

Corollary (Rodriguez-Wang, 2017)

$$MLdeg(X) = (-1)^d \chi(Eu_{X^\circ}).$$

Remark

If X° is smooth, $Eu_{X^\circ} = 1_{X^\circ}$, so we recover Huh's result.

Sectional ML degrees & Involution Conjecture

Definition (Sectional ML degrees)

Let $X \subset \mathbb{P}^n$ be a d-dimensional closed irreducible subvariety with $X^\circ = X \setminus \mathcal{H} \neq \emptyset$. The i-th sectional ML degree of X is:

$$s_i := MLdeg(X \cap L_{n-i}),$$

where L_{n-i} is a general linear subspace of $\mathbb{P}^n$ of codimension i.

- ♣ $s_d = \deg(X)$.
- ♣ $s_0 = MLdeg(X)$.

Huh-Sturmfels Involution Conjecture

Theorem (M.-Rodriguez-Wang-Wu, 2022)

Let $X \subset \mathbb{P}^n$ be a d-dimensional closed irreducible subvariety with $X^\circ = X \setminus \mathcal{H} \neq \emptyset$, and set

$$B_X(\mathrm{p}, \mathrm{u}) = (b_0 \cdot \mathrm{p}^d + b_1 \cdot \mathrm{p}^{d-1}\mathrm{u} + \cdots + b_d \cdot \mathrm{u}^d) \cdot \mathrm{p}^{n-d}.$$

$$S_X(\mathrm{p}, \mathrm{u}) = (s_0 \cdot \mathrm{p}^d + s_1 \cdot \mathrm{p}^{d-1}\mathrm{u} + \cdots + s_d \cdot \mathrm{u}^d) \cdot \mathrm{p}^{n-d}.$$

Then

$$B_X(\mathrm{p}, \mathrm{u}) = \frac{\mathrm{u} \cdot S_X(\mathrm{p}, \mathrm{u} - \mathrm{p}) - \mathrm{p} \cdot S_X(\mathrm{p}, 0)}{\mathrm{u} - \mathrm{p}},$$

$$S_X(\mathrm{p}, \mathrm{u}) = \frac{\mathrm{u} \cdot B_X(\mathrm{p}, \mathrm{u} + \mathrm{p}) + \mathrm{p} \cdot B_X(\mathrm{p}, 0)}{\mathrm{u} + \mathrm{p}}.$$

Conjectured by Huh-Sturmfels (2014). Proved by Huh when X° is smooth and Schön. Our proof follows from our geometric interpretation of $c_{Ma}(X^\circ)$ together with Aluffi's involution formula.

For a constructible function φ on $\mathbb{P}^n$, let $\varphi_j := \varphi|_{L_{n-j}}$ be the restriction of φ to a codimension j generic linear subspace.
E.g., if $\varphi = Eu_Z$ for a locally closed subvariety Z of $\mathbb{P}^n$, then $\varphi_j = Eu_{Z \cap L_{n-j}}$.
The Euler polynomial of φ is defined as:

$$\chi_\varphi(t) := \sum_{j \geq 0} \chi(\varphi_j) \cdot (-t)^j.$$

For $c_* : CF(\mathbb{P}^n) \to A_*(\mathbb{P}^n)$ the Chern class transformation, let

$$c_*(\varphi) = \sum_{j \geq 0} c_j [\mathbb{P}^j] \in A_*(\mathbb{P}^n),$$

and define the Chern polynomial of φ as

$$c_\varphi(t) := \sum_{j \geq 0} c_j t^j.$$

Deligne's prediction: The polynomials $\chi_\varphi(t)$ and $c_\varphi(t)$ carry precisely the same information!

Theorem (Ohmoto, 2003; Aluffi, 2013)

The involution on polynomials (of the same degree)

$$p(t) \longmapsto \mathcal{I}(p) := \frac{t \cdot p(-t-1) + p(0)}{t+1},$$

interchanges $c_\varphi(t)$ and $\chi_\varphi(t)$, i.e., $c_\varphi = \mathcal{I}(\chi_\varphi)$ and $\chi_\varphi = \mathcal{I}(c_\varphi)$.

Proof of the Involution Conjecture

Let $\varphi := Eu_{X^\circ}$, regarded as a constructible function on $\mathbb{P}^n$.

Step 1: Our geometric interpretation of $c_{Ma}(X^\circ) := c_*(Eu_{X^\circ})$, yields (with $d = \dim_{\mathbb{C}} X$):

$$B_X(\mathrm{p}, \mathrm{u}) = (-1)^d c_\varphi \left(-\frac{\mathrm{u}}{\mathrm{p}}\right) \mathrm{p}^n,$$

$$S_X(\mathrm{p}, \mathrm{u}) = (-1)^d \chi_\varphi \left(\frac{\mathrm{u}}{\mathrm{p}}\right) \mathrm{p}^n,$$

Step 2: Apply Aluffi's involution formula.

III. *Applications to linear optimization*

Linear optimization

Definition (Linear optimization (LO) degree)

The linear optimization (LO) degree $LOdeg(X)$ of an affine variety $X \subset \mathbb{C}^n$ is the number of critical points of a general linear function ℓ restricted to the smooth locus X_{reg} of X. It equals the cardinality of the general fiber of the projection of $T^*_X\mathbb{C}^n$ to the second factor $\mathbb{C}^n$, where $T^*_X\mathbb{C}^n$ is the conormal space of X, i.e., the closure in $T^*\mathbb{C}^n \cong \mathbb{C}^n \times \mathbb{C}^n$ of

$$T^*_{X_{reg}}\mathbb{C}^n := \{(\underline{x}, \underline{u}) \in T^*\mathbb{C}^n \mid \underline{x} \in X_{reg},\ \underline{u} \in T^*_{\underline{x}}\mathbb{C}^n, \underline{u}|_{T_{\underline{x}}X_{\mathrm{reg}}} = 0\}.$$

LO bidegrees

Definition (LO bidegrees)

The LO bidegrees of an irreducible affine variety $X \subset \mathbb{C}^n$, denoted by $b_i(X)$, are the bidegrees of $T^*_X\mathbb{C}^n$, i.e.,

$$[\overline{T^*_X\mathbb{C}^n}] = b_0[\mathbb{P}^0 \times \mathbb{P}^n] + b_1[\mathbb{P}^1 \times \mathbb{P}^{n-1}] + \cdots + b_d[\mathbb{P}^d \times \mathbb{P}^{n-d}] \in A_*(\mathbb{P}^n \times \mathbb{P}^n)$$

where $d = \dim X$, and $\overline{T^*_X\mathbb{C}^n}$ is the closure of $T^*_X\mathbb{C}^n \subset \mathbb{C}^n \times \mathbb{C}^n$ in $\mathbb{P}^n \times \mathbb{P}^n$.

♣ $b_0(X) = LOdeg(X)$.

LO bidegrees determine the Chern-Mather class

Fixing the standard compactification $\mathbb{C}^n \subset \mathbb{P}^n$, regard Eu_X for $X \subset \mathbb{C}^n$ as a constructible function on $\mathbb{P}^n$, with value 0 on $\mathbb{P}^n \setminus X$. The value $c_*(Eu_X)$ of the Chern-MacPherson transformation $c_* : CF(\mathbb{P}^n) \to A_*(\mathbb{P}^n)$ is the Chern-Mather class of X:

$$c_{Ma}(X) := c_*(Eu_X) = a_0[\mathbb{P}^0] + a_1[\mathbb{P}^1] + \cdots + a_d[\mathbb{P}^d] \in A_*(\mathbb{P}^n),$$

where $d = \dim(X)$.

Theorem (M.-Rodriguez-Wang-Wu, 2023)

$$\sum_{0 \leq i \leq d} b_i t^{n-i} = \sum_{0 \leq i \leq d} a_i (-1)^{d-i} t^{n-i} (1+t)^i.$$

Identifying the top degree coefficients of t yields:

Corollary (Seade-Tibăr-Verjovsky)

If $X \subset \mathbb{C}^n$ is a d-dimensional irreducible affine variety and $H \subset \mathbb{C}^n$ is a general affine hyperplane, one has $b_0 = (-1)^d \cdot \chi(Eu_X|_{\mathbb{C}^n \setminus H})$.

Sectional LO degrees

Definition (Sectional linear optimization (LO) degrees)

Let $X \subset \mathbb{C}^n$ be a d-dimensional irreducible affine variety. For any $0 \leq i \leq d$, define the *i-th sectional LO degree* $s_i(X)$ of X by

$$s_i(X) := LOdeg(X \cap H_1 \cap \cdots \cap H_i),$$

where $H_1, \ldots, H_i$ are generic affine hyperplanes.

- $s_0(X) = LOdeg(X)$.
- $s_d(X) = \deg(X)$.

Topological interpretation of LO bidegrees

Theorem (M.-Rodriguez-Wang-Wu, 2023)

Let $X \subset \mathbb{C}^n$ be any irreducible affine variety, and let b_i and s_i be its LO bidegrees and LO sectional degrees, respectively. Then $s_i = b_i$ for all i.

Corollary

$$b_i(X) = (-1)^{d-i}\chi(Eu_{X\cap H_1\cap\cdots\cap H_i}|_{\mathbb{C}^n\setminus H_{i+1}}).$$

Remark

Further relations to the polar degrees of $\bar{X} \subset \mathbb{P}^n$.

Some references

F. Catanese, **S. Hoșten**, **A. Khetan**, **B. Sturmfels**, *The maximum likelihood degree*, Amer. J. Math. 128 (2006), no. 3, 671–697.

J. Draisma, **E. Horobeț**,**G. Ottaviani**, **B. Sturmfels**, **R. Thomas**, *The Euclidean distance degree of an algebraic variety*, Found. Comput. Math. 16 (2016), no. 1, 99–149.

J. Huh, *The maximum likelihood degree of a very affine variety*, Compos. Math. 149 (2013), no. 8, 1245–1266.

J. Huh, **B. Sturmfels**, *Likelihood geometry*, in Combinatorial algebraic geometry, 63–117, Lecture Notes in Math., 2108, Fond. CIME/CIME Found. Subser., Springer, Cham, 2014.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, *Euclidean distance degree of the multiview variety*, SIAM J. Appl. Algebra Geometry 4 (2020), no. 1, 28–48.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, *Euclidean distance degree of projective varieties*, Int. Math. Res. Not. IMRN 2021, no. 20, 15788–15802.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, *Defect of Euclidean distance degree*, Adv. in Appl. Math. 121 (2020), 102101, 22 pp.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, *A Morse theoretic approach to non-isolated singularities and applications to optimization*, J. Pure Appl. Algebra 226 (2022), no. 3, Paper No. 106865, 23 pp.

L. Maxim, **M. Tibăr**, *Euclidean distance degree and limit points in a Morsification*, arXiv:2210.06022, Adv. Appl. Math. (to appear).

L. Maxim, **J. I. Rodriguez**, **B. Wang**, **L. Wu**, *Logarithmic cotangent bundles, Chern-Mather classes, and the Huh-Sturmfels Involution conjecture*, Comm. Pure Appl. Math. (to appear); arXiv:2202.00554.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, **L. Wu**, *Linear optimization on varieties and Chern-Mather classes*, arXiv:2208.09073.

L. Maxim, **J. I. Rodriguez**, **B. Wang**, *Applications of singularity theory in applied algebraic geometry and algebraic statistics*, arXiv:2305.19842.

That's All Folks !!!