The Entropic Discriminant

Raman Sanyal

Free University Berlin

joint with Bernd Sturmfels (UC Berkeley) and Cynthia Vinzant (U Michigan)

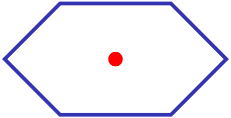
Analytic centers of polytopes

convex polyhedron $P = \{y \in \mathbb{R}^{n-d} : c_i^t y \leq e_i, i = 1, 2, \ldots, n\}$

The Barrier function

$$f(y) \ := \ \sum_{i=1}^{n} \log(e_i - c_i^t y)$$

is strictly concave on the interior of P. If P is a polytope (bounded) then the unique maximum over the interior of P is called the analytic center.



Basis for linear programming $\max\{w^t y : x \in P\}$ with interior point methods.

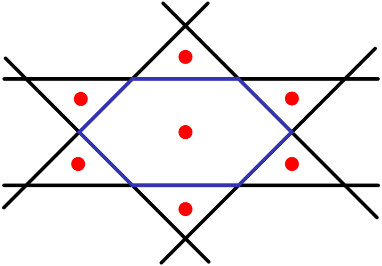
$$f_\lambda(y) = w^t y \ + \lambda \cdot f(y)$$

is strictly concave for $\lambda > 0$ and the maxima $y(\lambda) \to y_{\mathrm{opt}}$ for $\lambda \to 0$.

Analytic centers of hyperplane arrangements

Barrier function extends to the complement the induced hyperplane arr'ment

$$f(y) \;=\; \sum_{i=1}^{n} \log |e_i - c_i^t y|$$



Affine hyperplane arrangement partitions ambient space into regions.

Barrier function has maximum on a region if and only if the region is bounded.

$$\#\text{ analytic centers } \;=\; \#\text{ bounded regions}$$

Analytic centers are solutions to a polynomial system derived from $\{c_i^t y \leq e_i\}$.

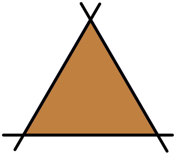
Convienient change of coordinates

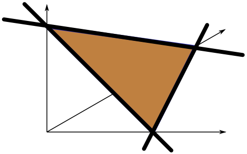
$$P = \{y \in \mathbb{R}^{n-d} : Cy \leq e\} \quad \text{for} \quad C = \begin{pmatrix} — \ c_1^t \ — \\ \vdots \\ — \ c_n^t \ — \end{pmatrix}, e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

Mapping $y \mapsto F(y) = e - Cy$ defines injection $F : \mathbb{R}^{n-d} \to \mathbb{R}^n$

$$\{y \in \mathbb{R}^{n-d} : c_i^t y = e_i\} \ \cong \ \{x \in \mathbb{R}^n : x_i = 0\} \ \cap \ \operatorname{im} F$$

For $A \in \mathbb{R}^{d \times n}$ of full rank such that $AC = 0$ we have $\operatorname{im} F = \{x : Ax = b\}$ with $b = Ae$.





The hyperplane arrangement is isomorphic to the arrangement of coordinate hyperplanes in $\{x : Ax = b\}$. The polytope $P \cong \{x : Ax = b, x \geq 0\}$.

The variety of analytic centers

From here on $A \in \mathbb{R}^{d \times n}$ of full row rank.

For $b \in \mathbb{R}^d$ the coordinate planes induce a hyperplane arrangement in $\{Ax = b\}$.

The Barrier function is the restriction of $x \mapsto f(x) = \log(x_1 x_2 \cdots x_n)$.
The analytic centers are the solutions to

$$\text{maximize} \quad f(x) \quad \text{subject to} \quad Ax = b$$

Karush–Kuhn–Tucker conditions are necessary and sufficient

$$\begin{array}{c} x^* \text{ is an} \\ \text{analytic center} \end{array} \quad \Longleftrightarrow \quad \begin{array}{rcl} Ax^* & = & b \\ \nabla f(x^*) & \in & \operatorname{rowspan}(A) \end{array}$$

$\nabla f(x) = x^{-1} := (x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})$ and the analytic centers are given by

$$\mathcal{C}_b \;=\; \{Ax = b\} \;\cap\; \operatorname{rowspan}(A)^{-1}$$

The variety of analytic centers

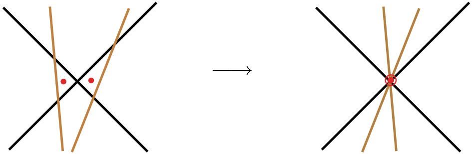
The Zariski closure $\mathcal{L}_A^{-1}$ of $\text{rowspan}(A)^{-1}$ inside $\mathbb{CP}^{n-1}$ was studied by Proudfoot & Speyer'06. The variety $\mathcal{L}_A^{-1}$ is irreducible of degree

$$\deg(\mathcal{L}_A^{-1}) \;=\; \#\text{ bounded regions in } \{Ax = b\}$$

(for $b \in \mathbb{R}^d$ generic).
This is exactly the number of analytic centers and

$$\mathcal{C}_b \;=\; \{Ax = b\} \;\cap\; \mathcal{L}_A^{-1}$$

is the (real) variety of analytic centers. $\{\mathcal{C}_b : b \in \mathbb{C}^d\}$ is a algebraic family.
For special b, $\mathcal{C}_b$ is not reduced. Real geometrically, analytic centers collide



This is a codimension-2 condition.

Question

What is the complex region for which $b \in \mathbb{C}^d$ is $\mathcal{C}_b$ singular?

The entropic discriminant

The entropic discriminant is the locus for which $\mathcal{C}_b$ is singular.

Theorem

Let A be a real $d \times n$-matrix whose columns span $\geq d+1$ distinct lines. The entropic discriminant is a hypersurface given by the vanishing of a homogeneous polynomial $H_A(b)$ of degree

$$\deg H_A \;=\; 2(-1)^d \cdot (d\chi_A(0) \;+\; \chi_A'(0)) \;\leq\; 2(n-d)\binom{n-1}{d-1}.$$

where $\chi_A(t)$ is the characteristic polynomial of A and equality holds for A generic.
Moreover, $H_A(b)$ is non-negative restricted on $\mathbb{R}^d$ and the real locus of $V(H_A) \subset \mathbb{CP}^{d-1}$ is pure of codimension-2 corresponding to colliding analytic centers.

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

Very difficult to compute. Here, the entropic discriminant has degree 8 and 39 terms

$$\chi_A(t) = t^3 - 5t^2 + 8t - 4$$

Polar Cremona transforms

$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{d\times n} \qquad \longrightarrow \qquad f(z) \;=\; \prod_{i=1}^{n} z^t a_i$$

$V_{\mathbb{C}}(f) =$ hyperplane arrangement associated to the columns of A

The polar map associated to f is

$$\nabla_f : \mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{d-1}, \quad z \mapsto \left(\tfrac{\partial f}{\partial z_1}(z) : \tfrac{\partial f}{\partial z_2}(z) : \cdots : \tfrac{\partial f}{\partial z_d}(z)\right) \;=\; A(z^t A)^{-1}$$

Rational map with base locus $\mathrm{Sing}(V(f)) =$ codim-2 subspaces of $V(f)$

Jacobian of ∇_f is the Hessian $\mathrm{Hess}(f) = \det\left(\frac{\partial^2 f}{\partial z_i \partial z_j}\right)$ of f

Proposition

$$V(H_A) \;=\; \textit{closure of } \nabla_f\left(V(\mathrm{Hess}(f)) \setminus V(f)\right).$$

Basic idea: $\{z : \nabla_f(z) = b\} \;=\; \{x : Ax = b, x \in \mathrm{rowspan}(A)^{-1}\}$

For $A = \mathrm{Id}_d$ the identity matrix we have $f(z) \;=\; z_1 z_2 \cdots z_d$ and

$$\mathrm{Hess}(f) \;=\; (-1)^{d-1}(d-2) f^{d-2}$$

and the entropic discriminant is not a hypersurface.
We call a matrix $A \in \mathbb{R}^{d \times n}$ basic if the columns span $\leq d$ lines.

Corollary

If A is not basic, then the entropic discriminant is a hypersurface.

Idea: $V(\mathrm{Hess}(f)) \setminus V(f) \neq \emptyset$ and $\nabla_f : \mathbb{P}^{d-1} \setminus V(f) \to \mathbb{P}^{d-1}$ is finite-to-one.

Corollary

$H_A(b)$ is a homogeneous polynomial, non-negative on $\mathbb{R}^d$.

Question

Is $H_A(b)$ a sum of squares

$$H_A(b) \;=\; h_1(b)^2 \;+\; h_2(b)^2 \;+\; \cdots \;+\; h_s(b)^2$$

for some $h_1, h_2, \ldots, h_s \in \mathbb{R}[b_1, \ldots, b_d]$?

The codimension-1 case

Theorem

If A is real non-basic matrix with $d-1$ rows and d columns, then $H_A(b)$ is an explicit sum of squares.

Sufficient to consider $A = \begin{pmatrix} 1 & & 1 \\ & \ddots & \vdots \\ & 1 & 1 \end{pmatrix}$. The reciprocal plane $\mathcal{L}_A^{-1}$ is the hypersurface defined by the symmetric determinantal form

$$\det \begin{pmatrix} x_1 + x_d & x_d & & x_d \\ x_d & x_2 + x_d & & x_d \\ \vdots & & \ddots & \vdots \\ x_d & x_d & \cdots & x_{d-1} + x_d \end{pmatrix}$$

$\mathcal{C}_b = \mathcal{L}_A^{-1} \cap \{Ax = b\}$ are the eigenvalues of a symmetric matrix $M(b)$.

$\mathcal{C}_b$ is singular iff $M(b)$ has a double eigenvalue.

The discriminant of $\det(M(b) - t \cdot \mathrm{Id})$ is a sum of squares (Borchart 1846, Newell'72).

This is arrangements, analytic centers, and collisions on the line! Boring?

Discriminants of derivatives

Let $p(t) \;=\; t^n + c_1 t^{n-1} + \cdots + c_{n-1} t + c_n = \prod_{i=1}^n (t - r_i)$ be a univariate polynomial.

The discriminant of $p(t)$ is the 'unique' polynomial in the coefficients such that

$$\Delta(c_1, \ldots, c_n) \;=\; \prod_{i<j}^{n} (r_i - r_j)^2$$

Conjecture (Sottile, Mukhin'10)

The discriminant of the derivative $p'(t)$ is a sum of squares in the $r_i - r_j$.

True!

Real case: The roots of $p(t)$ give an affine hyperplane arrangement in $\mathbb{R}^1$.

Deformation of n linear hyperplanes in $\mathbb{R}$ with $b_i = r_i - r_1$ for $i = 2, \ldots, n$.

The roots of $p'(t)$ are the analytic centers of the arrangement.

The discriminant of $p'(t)$ is the sum-of-squares $H_A(b)$.

Double eigenvalues

Question

Given a net $A(x,y) = A_0 + xA_1 + yA_2$ of real symmetric $n \times n$-matrices.
For how many real (x,y) does $A(x,y)$ have a double eigenvalue?

Theorem (Lax'98)

At least one, if $n = 4k + 2$.

Double eigenvalue $\Leftrightarrow$ Discriminant of $p_{(x,y)}(t) = \det(A(x,y) - t \cdot \mathrm{Id})$ vanishes.

Theorem (Sturmfels'02)

For a general net $A(x,y)$ the number of critical double eigenvalues is $\binom{n+1}{3}$
($=$ the degree of the singular set of the discriminant).

Conjecture (Sturmfels'02)

There is a net $A(x,y) \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ with precisely $\binom{n+1}{3}$ real double eigenvalues.

Theorem

There is a net $A(x,y) \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ with precisely $\binom{n+1}{3}$ real double eigenvalues.

Sketch of Proof.

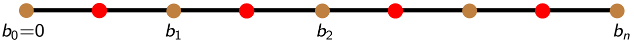
In codim 1, $H_A(b)$ is the discriminant of a real symmetric $n \times n$-matrix map

$$M(b) \;=\; b_1 M_1 + b_2 M_2 + \cdots + b_n M_n$$

for symmetric matrices $M_1, M_2, \ldots, M_n$.

Real locus $V(H_A) \cap \mathbb{R}^n$ is pure of codim 2 with $\binom{n+1}{3}$ linear components

$$\sqrt[\mathbb{R}]{\langle H_A(b)\rangle} \;=\; \bigcap_{1\le i<j\le n} \langle b_i, b_j\rangle \quad\cap\quad \bigcap_{1\le i<k<j\le n} \langle b_i - b_k, b_j - b_k\rangle$$



For generic vectors $a_0, a_1, a_2 \in \mathbb{R}^n$, the net

$$M(a_0 + x a_1 + y a_2) \;=\; A_0 + x A_1 + y A_2$$

has exactly $\binom{n+1}{3}$ real double eigenvalues.

Characteristic polynomials

$$\text{Think of } A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \subset \mathbb{R}^d \text{ as spanning collection of vectors}$$

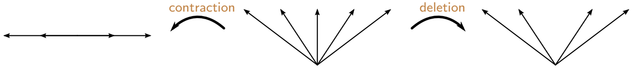
For a $J \subseteq [n]$, denote by A_J the corresponding subcollection. The characteristic polynomial of A is defined by

$$\chi_A(t) \;=\; \sum_{J \subseteq [n]} (-1)^{|J|} t^{d - \mathrm{rk}\,(A_J)}$$

depends only on the independence structure – deletion-contraction invariant.

Deleting the i-th column $A_{\setminus i} := A_{[n]\setminus i}$

Contracting the i-th column $A_{/i} \in \mathbb{R}^{(d-1)\times(n-1)}$ is obtained by projecting to $a_i^{\perp}$.



For $a_i \neq 0$ deletion-contraction gives

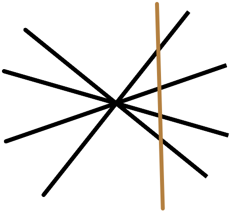
$$\chi_A(t) \;=\; \chi_{A_{\setminus i}}(t) \;-\; \chi_{A_{/i}}(t)$$

Geometric interpretations

Consider the linear hyperplane arr'ment $\mathcal{A}$ given by $H_i = \{a_i^t x = 0\} \subset \mathbb{R}^d$
Let G be a general affine hyperplane.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 2 \end{pmatrix}$$

$$\chi_A(t) = t^2 - 4t + 3$$



Theorem (Zaslavsky'70s)

$(-1)^d \chi_A(0)$ *is the number bounded components induced by* $\mathcal{A}$ *in* G.

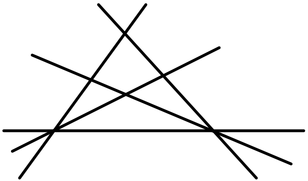
Deletion $A_{\setminus i}$ yields deletion $\mathcal{A} \setminus H_i$.
Contraction $A_{/i}$ yields restriction $\mathcal{A}|_{H_i}$ of $\mathcal{A} \setminus \{H_i\}$ to $H_i \cong \mathbb{R}^{d-1}$.

$$\#\left\{\begin{matrix}\text{bounded components} \\ \text{of } \mathcal{A}\end{matrix}\right\} = \#\left\{\begin{matrix}\text{bounded components} \\ \text{of } \mathcal{A} \setminus H_i\end{matrix}\right\} + \#\left\{\begin{matrix}\text{bounded components} \\ \text{of } \mathcal{A}|_{H_i}\end{matrix}\right\}$$

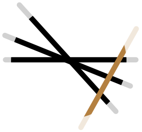
$$(-1)^d \chi_A(0) = (-1)^d \chi_{A_{\setminus i}}(0) + (-1)^{d-1} \chi_{A_{/i}}(0)$$

For $A = \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array}$ we get $\chi_A(t) = t^3 - 5t^2 + 8t - 4.$

The intersection of the corresponding arrangement with a general hyperplane G is this.



Analogous interpretations for all coefficients by summing 'local' contributions of closed sets A_J.



$J \subseteq [n]$ is closed if $\operatorname{rk} A_J < \operatorname{rk} A_K$ for all $K \supset J$.
The linear coefficient equals

$$\sum \{\chi_{A_J}(0) : J \text{ closed}, \operatorname{rk}(A_J) = d - 1\} = \chi'_A(0)$$

Geometry of $\mathcal{L}_A^{-1}$ = *closure of* rowspan$(A)^{-1} \subseteq \mathbb{P}^{n-1}$

In the torus $\mathbb{T}^{n-1} = (\mathbb{C}^*)^n$ this is a complete intersection given by $Bx^{-1} = 0$ where $\ker(B) = \text{rowspan}(A)$.

The boundary components are $\mathcal{L}_{A_J}^{-1}$ for $J \subseteq [n]$ is closed.

$$\mathcal{L}_A^{-1} \;=\; \bigcup_{J \text{ closed}} \text{rowspan}(A_J)^{-1} \;\subseteq\; \mathbb{P}^{n-1}$$

Singular set $\text{Sing}(\mathcal{L}_A^{-1})$ is the union of strata $\mathcal{L}_{A_J}^{-1}$ for which $A_{/J}$ is non-basic.

Degree $\deg(\mathcal{L}_A^{-1}) = (-1)^d \chi_A(0)$. In fact, $\chi_A(t)$ is the multidegree for

$$\textit{closure of } \left\{(u, u^{-1}) : u \in \text{rowspan}(A)\right\} \;\subset\; \mathbb{P}^{n-1} \times \mathbb{P}^{n-1}$$

The map $A : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{d-1}$ is basepoint free on $\mathcal{L}_A^{-1}$. The ramification locus is the (closure of) the set of smooth points $p \in \mathcal{L}_A^{-1}$ at which the tangent space $T_p\mathcal{L}_A^{-1}$ and $\{Ax = Ap\}$ are not transverse.

The entropic discriminant $V(H_A)$ is the image of the ramification locus of A on $\mathcal{L}_A^{-1}$. The degree of H_A equals the degree of the ramification cycle $\mathcal{R}_A$ of $\mathcal{L}_A^{-1}$.

Degree of ramification cycle

The ramification cycle is hard to get but over the torus the ramification cycle is defined by

$$g_A(x) \;=\; \det \begin{pmatrix} B \cdot \operatorname{diag}(x)^{-2} \\ A \end{pmatrix} \;=\; \det(A \cdot \operatorname{diag}(x)^2 \cdot A^T)$$

The tangent space at $p \in \operatorname{rowspan}(A)^{-1}$ is exactly $\ker(B \cdot \operatorname{diag}(x)^{-2})$.

Let $\widehat{\mathcal{R}}_A$ be the subscheme in $\mathcal{L}_A^{-1}$ cut out by g_A.

$$\widehat{\mathcal{R}}_A \;=\; \mathcal{R}_A \;\cup\; 2 \cdot \bigcup \left\{ \mathcal{L}_{A_J}^{-1} : J \text{ closed}, \operatorname{rk}(A_J) = d-1 \right\}$$

By Bézout's theorem $\deg(\widehat{\mathcal{R}}_A) = \deg(g_A) \cdot \deg(\mathcal{L}_A^{-1}) = (-1)^d 2d\chi_A(0)$ and thus

$$\deg(H_A) \;=\; \deg(\mathcal{R}_A) \;= (-1)^d 2\left(d\chi_A(0) + \chi_A'(0)\right).$$

Open Questions

What is the Newton polytopes of $H_A(b)$?
In codim-1 this is permutohedron $\Pi(2, 4, \ldots, 2d)$.

Is $H_A(b)$ always a sum-of-squares?
True in codim-1 and $d = 2$ but open even for $(d, n) = (3, 5)$.

Generators for the ramification scheme?

Conjecture

For a general $d \times n$-matrix A with $n \geq d + 2$, the ramification scheme in $\mathcal{L}_A^{-1}$ is defined by the polynomials(!) $\frac{g_A(x)}{x_i x_j}$ for $i, j \in [n]$.

What is a discrete-geometric interpretation for $2(-1)^d (d\chi_A(0) + \chi_A'(0))$?

The entropic discriminant

Raman Sanyal, Bernd Sturmfels, Cynthia Vinzant

arXiv:1108.2925