Boundaries of logarithmic Voronoi cells can be transcendental

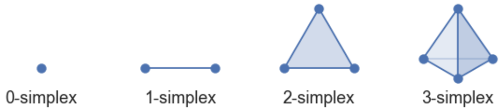
Yulia Alexandr
UCLA $\rightarrow$ University of Melbourne

SIAM Annual Meeting 2026
Cleveland, Ohio

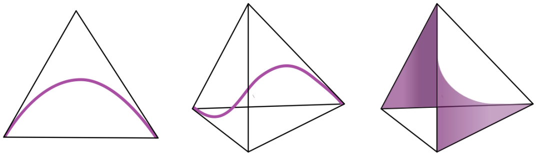
Statistical models

A *probability simplex* is defined as

$$\Delta_{n-1} = \{(p_1, \ldots, p_n) \in \mathbb{R}^n : p_1 + \cdots + p_n = 1, p_i \geq 0 \text{ for } i \in [n]\}.$$



An *algebraic statistical model* is an intersection of some variety $\mathcal{V}$ with the probability simplex.

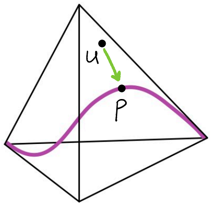


Maximum likelihood estimation

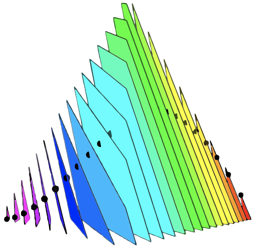
Let $\mathcal{M} = \mathcal{V} \cap \Delta_{n-1}$ be a statistical model.

Given empirical data $u \in \Delta_{n-1}$, the *maximum likelihood estimation (MLE)* problem is to find $p \in \mathcal{M}$ that maximizes:

$$\ell_u(p) = \sum_{i=1}^{n} u_i \log p_i.$$



For a $q \in \mathcal{M}$, the set of all data $u \in \Delta_{n-1}$ that maximize to q is the *logarithmic Voronoi cell* at q.



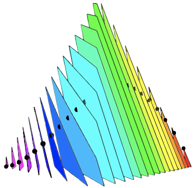
Proposition (A.–Heaton)

Logarithmic Voronoi cells are convex sets.

Logarithmic Voronoi cells

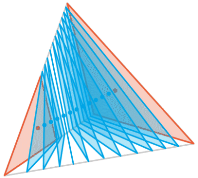
1. When are they simple?

Theorem (A.–Heaton):
If $\mathcal{M}$ is finite, linear, toric, or has ML degree 1, the cell at any $p \in \mathcal{M}$ is a polytope.



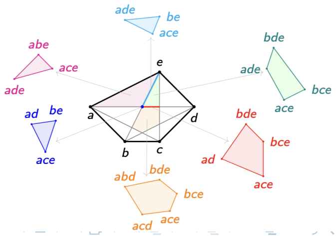
2. What is their combinatorics?

Theorem (A.):
For linear models, cells have the same combinatorial type, described via Gale diagrams.



3. How does it vary?

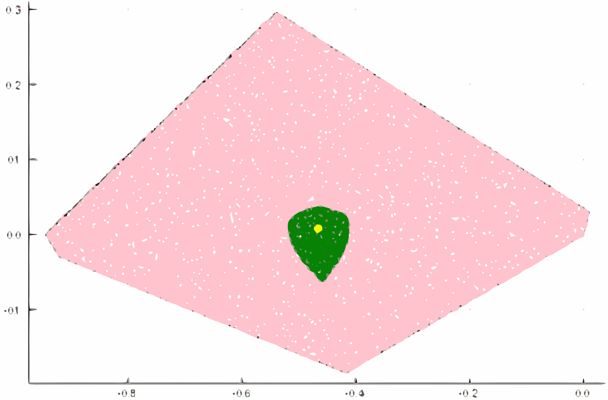
Theorem (A.–Hoșten):
For toric models, cells in the same chamber share the same combinatorial type and vertex supports.



Why do we care?

- Data privacy, especially for statistical disclosure limitation
  - Logarithmic Voronoi cells at singular and boundary points of the model

- Estimate sensitivity to data perturbation
  - Boundaries of logarithmic Voronoi cells



Example 1: two lines

Two lines

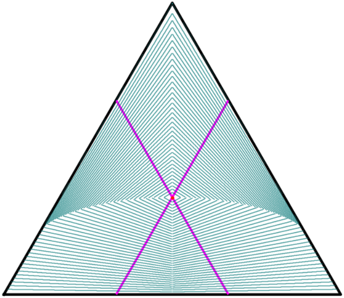
We take a model with two linear pieces:

$$\mathcal{M} = L_1 \cup L_2 \subset \Delta_2^\circ,$$

where

$$L_1 = \left\{ \left(\frac{1}{3} + t, \frac{1}{3}, \frac{1}{3} - t \right) : -\frac{1}{3} < t < \frac{1}{3} \right\},$$

$$L_2 = \left\{ \left(\frac{1}{3}, \frac{1}{3} + s, \frac{1}{3} - s \right) : -\frac{1}{3} < s < \frac{1}{3} \right\}.$$



Choose the smooth rational point

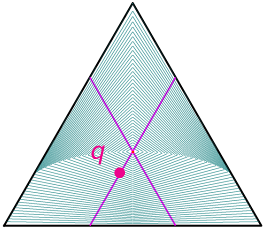
$$q = \left(\frac{4}{9}, \frac{1}{3}, \frac{2}{9} \right) \in L_1.$$

Claim: the boundary of $\mathrm{logVor}_{\mathcal{M}}(q)$ contains a transcendental point.

Fixing the MLE on the first line

Write a data point in ratio coordinates:

$$u = \frac{(\alpha, \beta, 1)}{\alpha + \beta + 1} \quad \text{where} \quad \alpha = \frac{u_1}{u_3}.$$



Along L_1, the likelihood only depends on α. The unique maximizer on L_1 is:

$$\widehat{p}_{L_1}(u) = \left(\frac{2\alpha}{3(\alpha+1)}, \frac{1}{3}, \frac{2}{3(\alpha+1)} \right).$$

We see that $\widehat{p}_{L_1}(u) = q \iff \alpha = 2.$

The data points whose MLE on L_1 is q form a one-parameter family:

$$u(\beta) = \frac{(2, \beta, 1)}{3 + \beta}, \qquad \beta > 0.$$

We move along this family and ask when the second line catches up.

The tie point is transcendental

Along the family $u(\beta) = \frac{(2,\beta,1)}{3+\beta}$, the boundary occurs when q ties with the MLE on L_2:

$$K(\beta) = K(2)$$

where

$$K(r) = r \log\left(\frac{2r}{r+1}\right) + \log\left(\frac{2}{r+1}\right).$$

Since $K'(r) < 0$ on $(0,1)$ and

$$\log 2 = K(0) > K(2) > K(1) = 0,$$

there is a unique solution $\beta^* \in (0,1)$.

Exponentiating the tie equation gives:

$$\left(\frac{2\beta^*}{\beta^*+1}\right)^{\beta^*} = \frac{16(\beta^*+1)}{27}.$$

Gelfond–Schneider: If $a \neq 0, 1$ is algebraic and b is algebraic irrational, then a^b is transcendental.

- 3-adic valuations show $\beta^* \notin \mathbb{Q}$.
- If β^* were algebraic irrational, G–S forces the LHS to be transcendental, but the RHS is algebraic. Contradiction!

Therefore, $u^* = \dfrac{(2,\beta^*,1)}{3+\beta^*} \in \partial \log\mathrm{Vor}_{\mathcal{M}}(q)$ is transcendental.

Example 2: a cusp

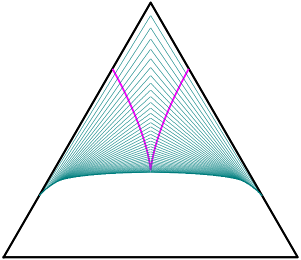
A cuspidal curve

Now consider a one-dimensional model with a singular point:

$$\mathcal{M} = \left\{ (x, y, z) \in \Delta_2 : \left(z - \frac{1}{3} \right)^3 = (x - y)^2 \right\}$$

We study the boundary of the (2-dim) logarithmic Voronoi cell at the singularity:

$$p_0 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right).$$



Claim:

the boundary of $\text{logVor}_{\mathcal{M}}(p_0)$ contains a non-algebraic real-analytic branch.

Tying with a smooth point

Parametrize the cusp by

$$p(t) = \left(\frac{1}{3} - \frac{1}{2}t^2 + \frac{1}{2}t^3, \frac{1}{3} - \frac{1}{2}t^2 - \frac{1}{2}t^3, \frac{1}{3} + t^2\right).$$

Then $p(0) = p_0$, while $p(t)$ is smooth for $t \neq 0$.

Pull the likelihood back to this parametrization:

$$L_u(t) = \ell_u(p(t)).$$

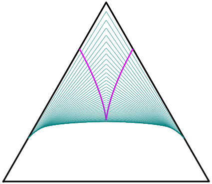
A smooth point $p(t)$ ties with p_0 and is critical precisely when

$$L_u(t) = L_u(0), \qquad L_u'(t) = 0.$$

Together with $u_1 + u_2 + u_3 = 1$, these are three linear equations in the data vector u.

Solving them gives a one-parameter branch of data points $u = u(t)$.

A logarithm appears



Let r be the parameter value where the cusp reaches the boundary of the simplex:

$$y(r) = 0.$$

We follow the branch of tied data points $u(t)$ as $t \longrightarrow r^-$.

$$y(t) \sim c(r-t) \qquad \Longrightarrow \qquad \log y(t) = \log(r-t) + O(1).$$

The tie and criticality equations therefore give

$$u_2 \sim c_1(r-t), \qquad u_1 - u_1^* \sim c_2(r-t)\log(r-t).$$

Eliminating t, the data branch satisfies

$$u_1 - u_1^* = c\, u_2 \log u_2 + O(u_2), \qquad c \neq 0.$$

The term $u_2 \log u_2$ makes this branch non-algebraic.

From the logarithm to the boundary

Why is the branch non-algebraic?

Near a point, an algebraic curve has a Puiseux expansion:

$$u_1 - u_1^* = \sum_{m \geq m_0} a_m u_2^{m/N}.$$

It may contain fractional powers, but not logarithms.
Our branch satisfies

$$u_1 - u_1^* = c\, u_2 \log u_2 + O(u_2).$$

Hence the branch is not algebraic.

Why is it the actual boundary?

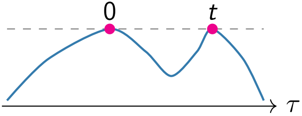
For data $u(t)$ on this branch, let

$$H_t(\tau) = \ell_{u(t)}(p(\tau)) - \ell_{u(t)}(p_0).$$

A global analysis shows that

$$H_t(\tau) \leq 0,$$

with equality only at $\tau = 0$ and $\tau = t$.



Hence, $\partial\, \mathrm{logVor}_{\mathcal{M}}(p_0)$ contains a non-algebraic real-analytic branch.

Thank you!