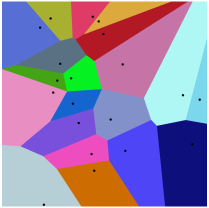
Combinatorics of logarithmic Voronoi cells

Yulia Alexandr (UC Berkeley)

Institute for Algebra and Geometry Seminar
Otto-von-Guericke-Universität Magdeburg, Germany

Voronoi cells in the Euclidean case

from Wikipedia:



Let X be a **finite** point configuration in $\mathbb{R}^n$.

- The *Voronoi cell* of $x \in X$ is the set of all points that are closer to x than any other $y \in X$, in the Euclidean metric.
- The subset of points that are equidistant from x and any other points in X is the *boundary* of the Voronoi cell of x.
- Voronoi cells partition $\mathbb{R}^n$ into convex polyhedra.

If X is a **variety**, each Voronoi cell is a convex semialgebraic set in the normal space of X at a point. The algebraic boundaries of these Voronoi cells were computed by Cifuentes, Ranestad, Sturmfels and Weinstein.

Logarithmic Voronoi cells for discrete models

- A *probability simplex* is defined as

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

- A *statistical model* is a subset of Δ_{n-1}. An *algebraic statistical model* is a subset $\mathcal{M} = \mathcal{V} \cap \Delta_{n-1}$ for some variety $\mathcal{V} \subseteq \mathbb{C}^n$.
- For an empirical data point $u = (u_1, ..., u_n) \in \Delta_{n-1}$, the *log-likelihood function* defined by u assuming distribution $p = (p_1, ..., p_n) \in \mathcal{M}$ is

$$\ell_u(p) = u_1 \log p_1 + u_2 \log p_2 + \cdots + u_n \log p_n.$$

Maximum likelihood estimation

Fix an algebraic statistical model $\mathcal{M} \subseteq \Delta_{n-1}$

1. The maximum likelihood estimation problem (MLE):

Given a sampled empirical distribution $u \in \Delta_{n-1}$, which point $p \in \mathcal{M}$ did it most likely come from? In other words, we wish to maximize $\ell_u(p)$ over all points $p \in \mathcal{M}$.

Maximum likelihood estimation

Fix an algebraic statistical model $\mathcal{M} \subseteq \Delta_{n-1}$

1. The maximum likelihood estimation problem (MLE):

Given a sampled empirical distribution $u \in \Delta_{n-1}$, which point $p \in \mathcal{M}$ did it most likely come from? In other words, we wish to maximize $\ell_u(p)$ over all points $p \in \mathcal{M}$.

2. Computing logarithmic Voronoi cells:

Given a point $q \in \mathcal{M}$, what is the set of all points $u \in \Delta_{n-1}$ that have q as a global maximum when optimizing the function $\ell_u(p)$ over $\mathcal{M}$?

The set of all such elements $u \in \Delta_{n-1}$ is the *logarithmic Voronoi cell* at q.

Proposition (A., Heaton)

Logarithmic Voronoi cells are convex sets.

The *log-normal space* at q is the space of possible data points $u \in \mathbb{R}^n$ for which q is a critical point of $\ell_u(p)$. It is a *linear* space.

Intersecting this space with the simplex Δ_{n-1}, we obtain a polytope, which we call the *log-normal polytope* at q.

The log-normal polytope at q contains the logarithmic Voronoi cell at q.

Proposition (A., Heaton)

Logarithmic Voronoi cells are convex sets.

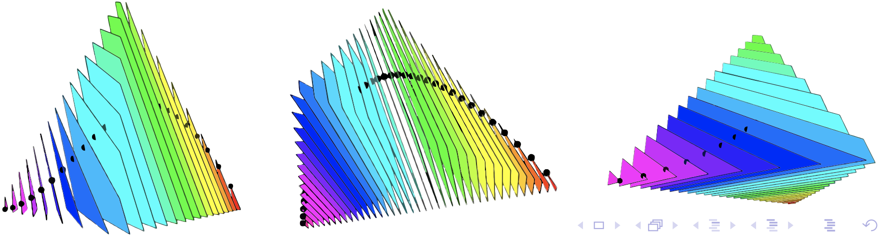
The *log-normal space* at q is the space of possible data points $u \in \mathbb{R}^n$ for which q is a critical point of $\ell_u(p)$. It is a *linear* space.

Intersecting this space with the simplex Δ_{n-1}, we obtain a polytope, which we call the *log-normal polytope* at q.

The log-normal polytope at q contains the logarithmic Voronoi cell at q.

Example (The twisted cubic.)

The curve is given by $p \mapsto \left(p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3\right)$.



The Hardy-Weinberg curve

Consider a model parametrized by

$$p \mapsto \left(p^2, 2p(1-p), (1-p)^2\right).$$

Performing implicitization, we find that the model $\mathcal{M} = \mathcal{V}(f)$ where

$$f = \begin{bmatrix} 4x_1x_3 - x_2^2 \\ x_1 + x_2 + x_3 - 1 \end{bmatrix}.$$

The augmented Jacobian is given by:

$$A = \begin{bmatrix} 4x_3 & -2x_2 & 4x_1 \\ 1 & 1 & 1 \\ u_1/x_1 & u_2/x_2 & u_3/x_3 \end{bmatrix}.$$

Fix a point $q \in \mathcal{M}$ and substitute x_i for q_i in A. All points $u \in \mathbb{R}^3$ at which the determinant vanishes define the log-normal space at q.

The Hardy-Weinberg curve

$$\det A = 4u_1 - 4u_3 - 4u_2 \cdot \frac{x_1}{x_2} + 2u_1 \cdot \frac{x_2}{x_1} - 2u_3 \cdot \frac{x_2}{x_3} + 4u_2 \cdot \frac{x_3}{x_2}$$

For example, at $p = 0.2$, we get a point $q = (0.04, 0.32, 0.64) \in \mathcal{M}$. The log-normal space at q is the plane

$$20u_1 + 7.5u_2 - 5u_3 = 0.$$

Sampling more points, we get the following pictures:

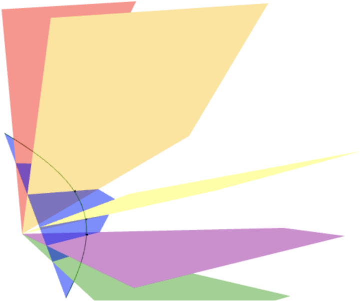
The Hardy-Weinberg curve

$$\det A = 4u_1 - 4u_3 - 4u_2 \cdot \frac{x_1}{x_2} + 2u_1 \cdot \frac{x_2}{x_1} - 2u_3 \cdot \frac{x_2}{x_3} + 4u_2 \cdot \frac{x_3}{x_2}$$

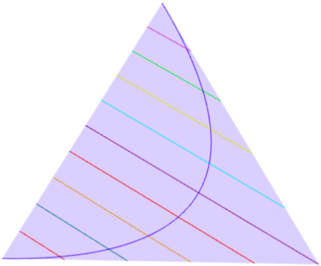
For example, at $p = 0.2$, we get a point $q = (0.04, 0.32, 0.64) \in \mathcal{M}$. The log-normal space at q is the plane

$$20u_1 + 7.5u_2 - 5u_3 = 0.$$

Sampling more points, we get the following pictures:



Log-normal spaces



Log-normal polytopes = Log-Voronoi cells

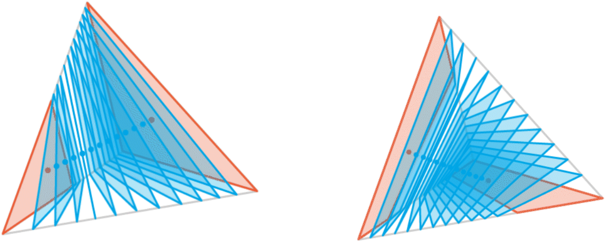
Polytopal cells

The *maximum likelihood degree* (ML degree) of $\mathcal{M}$ is the number of complex critical points when optimizing $\ell_u(x)$ over $\mathcal{M}$ for generic data u.

Theorem (A., Heaton)

If $\mathcal{M}$ is a finite model, a linear model, a toric model, or a model of ML degree 1, the logarithmic Voronoi cell at any point $p \in \mathcal{M}$ is equal to the log-normal polytope at p.

For linear models, logarithmic Voronoi cells at all interior points on the model have the same combinatorial type.



Gaussian models

Let X be an m-dimensional random vector, which has the density function

$$p_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{m/2}(\det \Sigma)^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}, \quad x \in \mathbb{R}^m$$

with respect to the parameters $\mu \in \mathbb{R}^m$ and $\Sigma \in \mathrm{PD}_m$.

Such X is distributed according to the *multivariate normal distribution*, also called the *Gaussian distribution* $\mathcal{N}(\mu, \Sigma)$.

For $\Theta \subseteq \mathbb{R}^m \times \mathrm{PD}_m$, the statistical model

$$\mathcal{P}_\Theta = \{\mathcal{N}(\mu, \Sigma) : \theta = (\mu, \Sigma) \in \Theta\}$$

is called a *Gaussian model*. We identify the Gaussian model $\mathcal{P}_\Theta$ with its parameter space Θ.

Gaussian models

For a sampled data consisting of n vectors $X^{(1)}, \cdots, X^{(n)} \in \mathbb{R}^m$, we define the *sample mean* and *sample covariance* as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X^{(i)} \qquad \text{and} \qquad S = \frac{1}{n}\sum_{i=1}^{n}(X^{(i)} - \bar{X})(X^{(i)} - \bar{X})^T,$$

respectively. The *log-likelihood function* is defined as

$$\ell_n(\mu, \Sigma) = -\frac{n}{2}\log\det\Sigma - \frac{1}{2}\operatorname{tr}\left(S\Sigma^{-1}\right) - \frac{n}{2}(\bar{X} - \mu)^T\Sigma^{-1}(\bar{X} - \mu).$$

In practice, we will only consider models given by parameter spaces of the form $\Theta = \mathbb{R}^m \times \Theta_2$ where $\Theta_2 \subseteq \mathrm{PD}_m$. **Thus, a Gaussian model is a subset of** PD_m. The log-likelihood function is then

$$\ell_n(\Sigma, S) = -\frac{n}{2}\log\det\Sigma - \frac{n}{2}\operatorname{tr}(S\Sigma^{-1}).$$

Algebraic models

All Gaussian models discussed in this talk are **algebraic**. In other words,

$$\Theta = \mathcal{V} \cap \mathrm{PD}_m,$$

where $\mathcal{V} \subseteq \mathbb{C}^m$ is a variety given by polynomials in the entries of $\Sigma = (\sigma_{ij})$.

Maximum likelihood estimation

Fix a Gaussian model $\Theta \subseteq \mathrm{PD}_m$.

1. The maximum likelihood estimation problem (MLE):

Given a sample covariance matrix $S \in \mathrm{PD}_m$, which matrix $\Sigma \in \Theta$ did it most likely come from? In other words, we wish to maximize $\ell_n(\Sigma, S)$ over all points $\Sigma \in \Theta$.

Maximum likelihood estimation

Fix a Gaussian model $\Theta \subseteq \mathrm{PD}_m$.

1. The maximum likelihood estimation problem (MLE):

Given a sample covariance matrix $S \in \mathrm{PD}_m$, which matrix $\Sigma \in \Theta$ did it most likely come from? In other words, we wish to maximize $\ell_n(\Sigma, S)$ over all points $\Sigma \in \Theta$.

2. Computing logarithmic Voronoi cells:

Given a matrix $\Sigma \in \Theta$, what is the set of all $S \in \mathrm{PD}_m$ that have Σ as a global maximum when optimizing the function $\ell_n(\Sigma, S)$ over Θ?

The set of all such matrices $S \in \mathrm{PD}_m$ is the *logarithmic Voronoi cell* at Σ.

Logarithmic Voronoi cells

Proposition (A., Hoșten)

Logarithmic Voronoi cells are ***still*** *convex sets.*

The *maximum likelihood degree* (ML degree) of Θ is the number of complex critical points in $\mathrm{Sym}_m(\mathbb{C})$ when optimizing $\ell_n(\Sigma, S)$ over Θ for a generic matrix S.

For $\Sigma \in \Theta$, the *log-normal matrix space* at Σ is the set of $S \in \mathrm{Sym}_m(\mathbb{R})$ such that Σ appears as a critical point of $\ell_n(\Sigma, S)$. The intersection of this space with PD_m is the *log-normal spectrahedron* $\mathcal{K}_\Theta\Sigma$ at Σ.

The logarithmic Voronoi cell at Σ is always contained in the log-normal spectrahedron at Σ.

Discrete vs. Gaussian

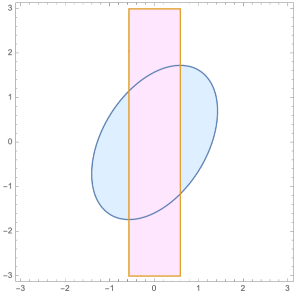
$$\begin{aligned}
\text{Simplex } \Delta_{n-1} &\longleftrightarrow \text{Cone } \mathrm{PD}_m \\
\text{Model } \mathcal{M} \subseteq \Delta_{n-1} &\longleftrightarrow \text{Model } \Theta \subseteq \mathrm{PD}_m \\
\sum_{i=1}^{n} u_i \log p_i &\longleftrightarrow \log\det\Sigma - \mathrm{tr}(S\Sigma^{-1}) \\
\text{Log-normal space } &\longleftrightarrow \text{Log-normal matrix space} \\
\text{Log-normal polytope} &\longleftrightarrow \text{Log-normal spectrahedron}
\end{aligned}$$

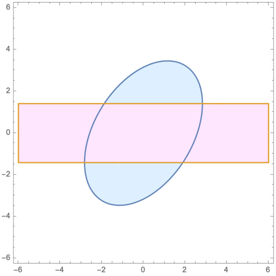
Example

Consider the model Θ given parametrically as

$$\Theta = \{\Sigma = (\sigma_{ij}) \in \mathrm{PD}_3 : \sigma_{13} = 0 \text{ and } \sigma_{12}\sigma_{23} - \sigma_{22}\sigma_{13} = 0\}.$$

This model is the union of two linear four-dimensional planes. It has ML degree 2. The log-normal spectrahedron of each point $\Sigma \in \Theta$ is an ellipse. Each log-Voronoi cell is given as:





Spectrahedral cells

When are logarithmic Voronoi cells equal to the log-normal spectrahedra?

Theorem (A., Hoșten)

If Θ *is a linear concentration model or a model of ML degree one, the logarithmic Voronoi cell at any* $\Sigma \in \Theta$ *equals the log-normal spectrahedron at* Σ*. In particular, this includes both undirected and directed graphical models.*

Spectrahedral cells

When are logarithmic Voronoi cells equal to the log-normal spectrahedra?

Theorem (A., Hoşten)

If Θ is a linear concentration model or a model of ML degree one, the logarithmic Voronoi cell at any $\Sigma \in \Theta$ equals the log-normal spectrahedron at Σ. In particular, this includes both undirected and directed graphical models.

Let $G = (V, E)$ be a simple undirected graph with $|V(G)| = m$. A *concentration model* of G is the model

$$\Theta = \{\Sigma \in \mathrm{PD}_m : (\Sigma)^{-1}_{ij} = 0 \text{ if } ij \notin E(G) \text{ and } i \neq j\}.$$

The logarithmic Voronoi cell at Σ is:

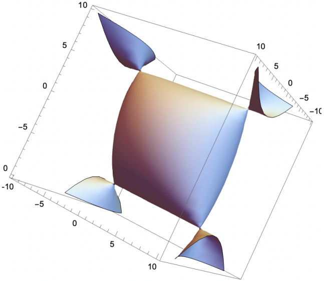
$$\log \mathrm{Vor}_{\Theta}(\Sigma) = \{S \in \mathrm{PD}_m : \Sigma_{ij} = S_{ij} \text{ for all } ij \in E(G) \text{ and } i = j\}.$$

Example

The concentration model of $\overset{1}{\bullet}\!-\!\!\!-\!\overset{2}{\bullet}\!-\!\!\!-\!\overset{3}{\bullet}\!-\!\!\!-\!\overset{4}{\bullet}$ is defined by

$$\Theta = \{\Sigma = (\sigma_{ij}) \in \mathrm{PD}_4 : (\Sigma^{-1})_{13} = (\Sigma^{-1})_{14} = (\Sigma^{-1})_{24} = 0\}.$$

Let $\Sigma = \begin{pmatrix} 6 & 1 & \frac{1}{7} & \frac{1}{28} \\ 1 & 7 & 1 & \frac{1}{4} \\ \frac{1}{7} & 1 & 8 & 2 \\ \frac{1}{28} & \frac{1}{4} & 2 & 9 \end{pmatrix}$.



Then $\log \mathrm{Vor}_\Theta(\Sigma) = \left\{ (x,y,z) : \begin{pmatrix} 6 & 1 & x & y \\ 1 & 7 & 1 & z \\ x & 1 & 8 & 2 \\ y & z & 2 & 9 \end{pmatrix} \succ 0 \right\}$.

Covariance models and the bivariate correlation model

Let $A \in \mathrm{PD}_m$ and let $\mathcal{L}$ be a linear subspace of $\mathrm{Sym}_m(\mathbb{R})$. Then $A + \mathcal{L}$ is an affine subspace of $\mathrm{Sym}(\mathbb{R}^m)$. Models defined by $\Theta = (A + \mathcal{L}) \cap \mathrm{PD}_m$ are called *covariance models*.

The *bivariate correlation model* is the covariance model

$$\Theta = \left\{ \Sigma_x = \begin{pmatrix} 1 & x \\ x & 1 \end{pmatrix} : x \in (-1, 1) \right\}.$$

This model has ML degree 3. For a general matrix $S = (S_{ij}) \in \mathrm{PD}_2$, the critical points are given by the roots of the polynomial

$$f(x) = x^3 - bx^2 - x(1 - 2a) - b,$$

where $b = S_{12}$ and $a = (S_{11} + S_{22})/2$ [Améndola and Zwiernik].

The bivariate correlation model

Fix $c \in (-1,1)$ so $\Sigma_c \in \Theta$. The log-normal spectrahedron of Σ_c is

$$\begin{aligned}\mathcal{K}_\Theta(\Sigma_c) &= \{S \in \mathrm{PD}_2 : f(c) = 0\} \\ &= \{S \in \mathrm{PD}_2 : a = (bc^2 - c^3 + b + c)/2c\} \\ &= \left\{ S_{b,k} = \begin{pmatrix} k & b \\ b & 2a-k \end{pmatrix} \succ 0 : \substack{0 \le k \le 2a, \\ a=(bc^2-c^3+b+c)/2c} \right\}.\end{aligned}$$

Theorem (A., Hoșten)

Let Θ be the bivariate correlation model and let $\Sigma_c \in \Theta$. If $c > 0$, then

$$\log \mathrm{Vor}_\Theta(\Sigma_c) = \{S_{b,k} \in \mathcal{K}_\Theta(\Sigma_c) : b \ge 0\}.$$

If $c < 0$, then

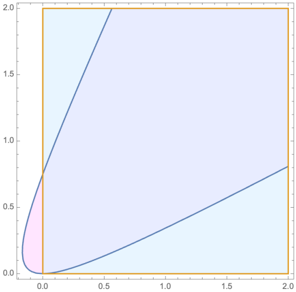
$$\log \mathrm{Vor}_\Theta(\Sigma_c) = \{S_{b,k} \in \mathcal{K}_\Theta(\Sigma_c) : b \le 0\}.$$

The bivariate correlation model

Important things to note:

- The log-Voronoi cell of Σ_c is strictly contained in the log-normal spectrahedron of Σ_c.
- Logarithmic Voronoi cells of Θ are semi-algebraic sets! **This is extremely surprising!**

The logarithmic Voronoi cell and the log-normal spectrahedron at $c = 1/2$:



The boundary: transcendental? algebraic?

Given a Gaussian model Θ and $\Sigma \in \Theta$, the matrix $S \in \mathrm{PD}_m$ is on the boundary of $\log \mathrm{Vor}_\Theta(\Sigma)$ if $S \in \log \mathrm{Vor}_\Theta(\Sigma)$ and there is some $\Sigma' \in \Theta$ such that $\ell(\Sigma, S) = \ell(\Sigma', S)$.

The bivariate correlation models fit into a larger class of models known as *unrestricted correlation models*. Such a model is given by the parameter space

$$\Theta = \{\Sigma \in \mathrm{Sym}(\mathbb{R}^m) : \Sigma_{ii} = 1,\ i \in [m]\} \cap \mathrm{PD}_m\,.$$

When $m = 3$, the model is a compact spectrahedron known as the elliptope. Its ML degree is 15.

Conjecture

The logarithmic Voronoi cells for general points on the elliptope are not semi-algebraic; in other words, their boundary is defined by transcendental functions.

Thanks!