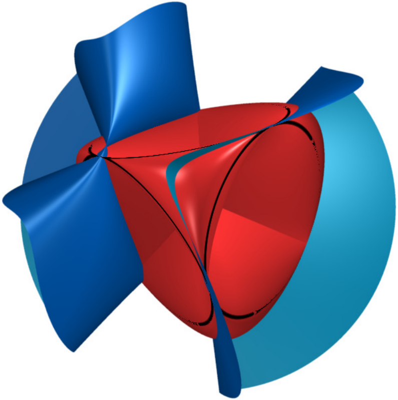
Exponential Varieties

Bernd Sturmfels
UC Berkeley



Joint paper with Mateusz Michałek,
Caroline Uhler, and Piotr Zwiernik

Motivation 1: Toric Geometry

A central theme in Algebraic Statistics is the connection between toric varieties and discrete exponential families. Binomial equations defining toric varieties are Markov bases.

[Diaconis-St 1998]

Example (Independence of binary random variables)

The *Segre variety* $\mathcal{V} = \mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$ is defined by

$$\det \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = 0.$$

The *moment map* takes $\mathcal{V}$ onto $K =$ the square $= \Delta_1 \times \Delta_1$.
It computes *sufficient statistics*:

$$\mathcal{V}_{\geq 0} \longrightarrow K$$

This is invertible. Its inverse is the *maximum likelihood estimator*.

Motivation 2: Gaussian Geometry

Let $\mathcal{L}$ be a linear space of real symmetric $m \times m$-matrices.

[St-Uhler 2010] studied the variety

$$\mathcal{L}^{-1} \;=\; \left\{ \sigma \in \mathrm{Sym}_2\mathbb{R}^m \,:\, \sigma^{-1} \in \mathcal{L} \right\}^{\mathrm{cl}}$$

The Gaussian model is the subset of covariance matrices

$$\mathcal{L}^{-1}_{\succ 0} \;=\; \left\{ \sigma \in \mathcal{L}^{-1} \,:\, \sigma \text{ positive definite} \right\}$$

Example (Graphical models)

$\mathcal{L}$ encodes sparsity of an undirected graph with m nodes.

The map dual to $\mathcal{L} \hookrightarrow \mathrm{Sym}_2\mathbb{R}^m$ computes *sufficient statistics*:

$$\mathcal{L}^{-1}_{\succ 0} \longrightarrow K \;=\; (\mathcal{L}_{\succ 0})^{\vee}.$$

This is invertible. Its inverse is the *maximum likelihood estimator*.

Exponential Families

An **exponential family** is a parametric statistical model

$$p_\theta(x) \;=\; \exp\big(-\langle\,\theta, T(x)\rangle \;-\; A(\theta)\,\big).$$

on a sample space (X, ν, T), with $T : X \to \mathbb{R}^d$ measurable.

Here $A(\theta)$ is the *log-partition function*.

Since $\int_X p_\theta(x)\nu(dx) = 1$,

$$A(\theta) \;\;=\;\; \log \int_X \exp\big(-\langle\theta, T(x)\rangle\big)\,\nu(dx).$$

Exponential Families

An **exponential family** is a parametric statistical model

$$p_\theta(x) \; = \; \exp\big(- \langle \theta, T(x) \rangle \; - \; A(\theta) \big).$$

on a sample space (X, ν, T), with $T : X \to \mathbb{R}^d$ measurable.

Here $A(\theta)$ is the *log-partition function*.

Since $\int_X p_\theta(x)\nu(dx) = 1$,

$$A(\theta) \;\; = \;\; \log \int_X \exp\big(-\langle \theta, T(x)\rangle\big)\, \nu(dx).$$

The following sets are **convex**:

Space of canonical parameters: $C \;=\; \big\{\theta \in \mathbb{R}^d : \; A(\theta) < +\infty\big\}$

Space of sufficient statistics: $K \;=\; \mathrm{conv}\big(T(X)\big) \,\subset\, \mathbb{R}^d$

Exponential Families

An **exponential family** is a parametric statistical model

$$p_\theta(x) \;=\; \exp\big(-\langle\,\theta, T(x)\rangle \;-\; A(\theta)\,\big).$$

on a sample space (X, ν, T), with $T : X \to \mathbb{R}^d$ measurable.

Here $A(\theta)$ is the *log-partition function*.

Since $\int_X p_\theta(x)\nu(dx) = 1$,

$$A(\theta) \;=\; \log \int_X \exp\big(-\langle\theta, T(x)\rangle\big)\,\nu(dx).$$

The following sets are **convex**:

Space of canonical parameters: $C \;=\; \big\{\theta \in \mathbb{R}^d : \; A(\theta) < +\infty\big\}$

Space of sufficient statistics: $K \;=\; \mathrm{conv}\big(T(X)\big) \;\subset\; \mathbb{R}^d$

Theorem

Suppose C is open and K spans $\mathbb{R}^d$. The gradient map

$$F \;:\; \mathbb{R}^d \;\to\; \mathbb{R}^d,\;\; \theta \;\mapsto\; -\nabla A(\theta)$$

defines an analytic bijection between C and $\mathrm{int}(K)$.

From Analysis to Algebra

Our exponential families satisfy

$$A(\theta) \;=\; -\alpha \cdot \log(f(\theta)),$$

where $f(\theta)$ is a homogeneous polynomial and $\alpha > 0$.

The gradient of the log-partition function is the rational function

$$F \;:\quad \mathbb{R}^d \dashrightarrow \mathbb{R}^d \;:\; \theta \;\mapsto\; \frac{\alpha}{f(\theta)} \cdot \Big(\frac{\partial f}{\partial \theta_1}, \frac{\partial f}{\partial \theta_2}, \ldots, \frac{\partial f}{\partial \theta_d}\Big).$$

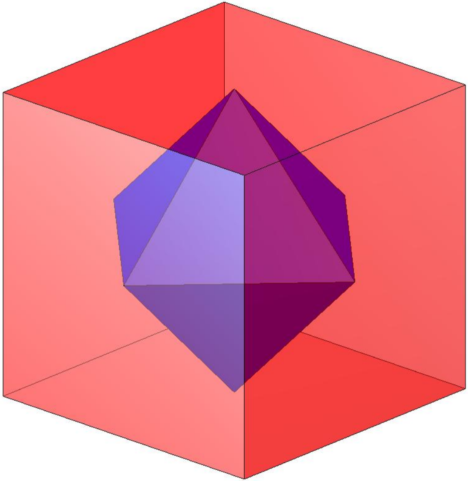
Algebraic geometers prefer

$$F \;:\; \mathbb{CP}^{d-1} \dashrightarrow \mathbb{CP}^{d-1} \;:\; \theta \;\mapsto\; \Big(\frac{\partial f}{\partial \theta_1} : \frac{\partial f}{\partial \theta_2} : \cdots : \frac{\partial f}{\partial \theta_d}\Big).$$

The partition function $f(\theta)^\alpha$ admits a nice integral representation. Which polynomials $f(\theta)$ and convex sets $C, K \subset \mathbb{R}^d$ are possible?

Duality of Polytopes

Example (How to morph a cube into an octahedron?)



[St-Uhler 2010, Example 3.5]

Duality of Polytopes

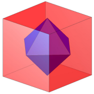
Example (Exponential family for cube $\to$ octahedron)

Fix the product of linear forms

$$f(\theta) \;=\; (\theta_1^2 - \theta_4^2)(\theta_2^2 - \theta_4^2)(\theta_3^2 - \theta_4^2)$$

The *space of canonical parameters* is

$$C \;=\; \text{cone over the 3-cube } \big\{\, |\theta_i| < 1 : \;\; i = 1,2,3 \,\big\}$$



The *space of sufficient statistics* is

$$K \;=\; \text{cone over the octahedron } \operatorname{conv}\{\pm\mathbf{e}_1, \pm\mathbf{e}_2, \pm\mathbf{e}_3\}$$

Gradient map $\nabla f : \mathbb{P}^3 \dashrightarrow \mathbb{P}^3$ gives bijection between C and $\operatorname{int}(K)$. Its inverse is an algebraic function of degree 7.

Question: What is (X, ν, T) in this case?

Duality of Polytopes

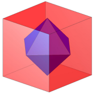
Example (Exponential family for cube $\rightarrow$ octahedron)

Fix the product of linear forms

$$f(\theta) \;=\; (\theta_1^2 - \theta_4^2)(\theta_2^2 - \theta_4^2)(\theta_3^2 - \theta_4^2)$$

The *space of canonical parameters* is

$$C \;=\; \text{cone over the 3-cube } \left\{\, |\theta_i| < 1 : \;\; i = 1, 2, 3 \,\right\}$$



The *space of sufficient statistics* is

$$K \;=\; \text{cone over the octahedron } \operatorname{conv}\{\pm\mathbf{e}_1, \pm\mathbf{e}_2, \pm\mathbf{e}_3\}$$

Gradient map $\nabla f : \mathbb{P}^3 \dashrightarrow \mathbb{P}^3$ gives bijection between C and $\operatorname{int}(K)$. Its inverse is an algebraic function of degree 7.

Question: What is (X, ν, T) in this case?
Answer: $X = K$, $T = \mathrm{id}$, and ν constructed via hypergeometric functions

Hyperbolic Polynomials

A homog. polynomial $f \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ of degree k is *hyperbolic* if, for some $\mathbf{t} \in \mathbb{R}^d$, every line through $\mathbf{t}$ intersects the complex hypersurface $\{f = 0\}$ in k real points. The connected component C of $\mathbf{t}$ in $\mathbb{R}^d \backslash \{f = 0\}$ is the *hyperbolicity cone*. It is convex.

Hyperbolic Polynomials

A homog. polynomial $f \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ of degree k is *hyperbolic* if, for some $\mathbf{t} \in \mathbb{R}^d$, every line through $\mathbf{t}$ intersects the complex hypersurface $\{f = 0\}$ in k real points. The connected component C of $\mathbf{t}$ in $\mathbb{R}^d \backslash \{f = 0\}$ is the *hyperbolicity cone*. It is convex.

Our integral representation lives on the *dual hyperbolicity cone*:

Theorem (Gårding 1951 ... Scott-Sokal 2015)

If $\alpha > d$, there exists a measure ν on the cone $K = C^\vee$ such that

$$f(\theta)^{-\alpha} = \int_K \exp(-\langle \theta, \sigma \rangle)\, \nu(d\sigma) \qquad \textit{for all } \theta \in C.$$

Furthermore, this property characterizes hyperbolic polynomials.

Hyperbolic Polynomials

A homog. polynomial $f \in \mathbb{R}[\theta_1, \ldots, \theta_d]$ of degree k is *hyperbolic* if, for some $\mathbf{t} \in \mathbb{R}^d$, every line through $\mathbf{t}$ intersects the complex hypersurface $\{f = 0\}$ in k real points. The connected component C of $\mathbf{t}$ in $\mathbb{R}^d \backslash \{f = 0\}$ is the *hyperbolicity cone*. It is convex.

Our integral representation lives on the *dual hyperbolicity cone*:

Theorem (Gårding 1951 ... Scott-Sokal 2015)

If $\alpha > d$, there exists a measure ν on the cone $K = C^\vee$ such that

$$f(\theta)^{-\alpha} \;=\; \int_K \exp(-\langle \theta, \sigma \rangle)\, \nu(d\sigma) \qquad \textit{for all } \theta \in C.$$

Furthermore, this property characterizes hyperbolic polynomials.

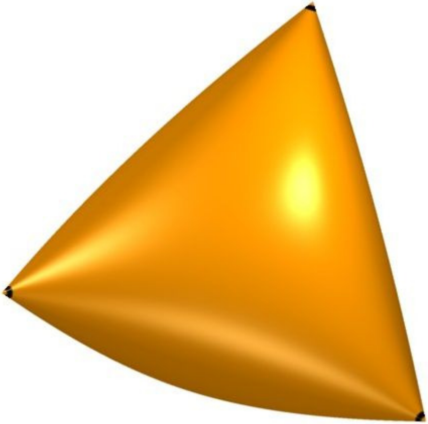
Proof: Riesz kernels and more. Lots of analysis.

The resulting statistical models are *hyperbolic exponential families*.
Related to *hyperbolic programming* in convex optimization [Güler].

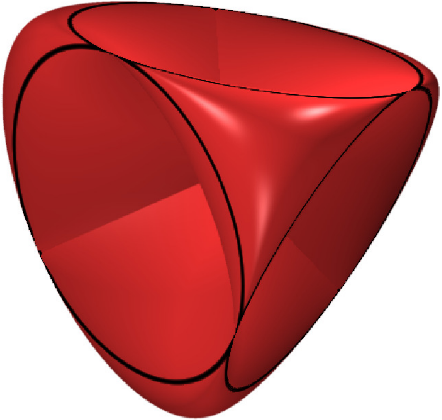
Hyperbolic Exponential Families: An Example

The space of canonical parameters C is the *hyperbolicity cone* of

$$f = \theta_1\theta_2\theta_3 + \theta_1\theta_2\theta_4 + \theta_1\theta_3\theta_4 + \theta_2\theta_3\theta_4.$$



Its dual $K = C^\vee$ is the space of sufficient statistics:

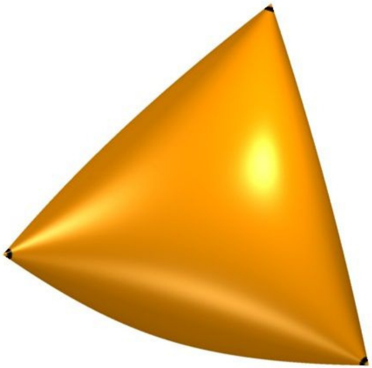


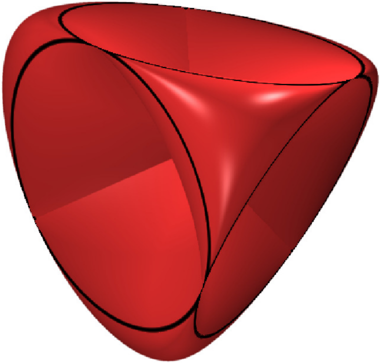
Steiner surface a.k.a Roman surface

$$\sum \sigma_i^4 - 4\sum \sigma_i^3\sigma_j + 6\sum \sigma_i^2\sigma_j^2 + 4\sum \sigma_i^2\sigma_j\sigma_k - 40\,\sigma_1\sigma_2\sigma_3\sigma_4.$$

Duality

Gradient map $\nabla f : \mathbb{P}^3 \to \mathbb{P}^3$ gives a bijection between C and K:





We shall be interested in the geometry its graph $X_f \subset \mathbb{P}^3 \times \mathbb{P}^3$.

Gaussian Family is Hyperbolic

Let $X = \mathbb{R}^m$, where ν is Lebesgue measure, and set

$$T(x) \;=\; \frac{1}{2} x \cdot x^T \;\in\; \mathrm{Sym}_2(\mathbb{R}^m) \;\simeq\; \mathbb{R}^d.$$

The symmetric determinant $f(\theta) = \det(\theta)$ is a hyperbolic polynomial in $d = \binom{m+1}{2}$ unknowns. Its hyperbolicity cone C consists of positive definite matrices. This cone is self-dual:

$$K \;=\; C^\vee \;=\; \mathrm{conv}(T(X)) \;\simeq\; C.$$

Gaussian Family is Hyperbolic

Let $X = \mathbb{R}^m$, where ν is Lebesgue measure, and set

$$T(x) \;=\; \frac{1}{2}x \cdot x^T \;\in\; \mathrm{Sym}_2(\mathbb{R}^m) \;\simeq\; \mathbb{R}^d.$$

The symmetric determinant $f(\theta) = \det(\theta)$ is a hyperbolic polynomial in $d = \binom{m+1}{2}$ unknowns. Its hyperbolicity cone C consists of positive definite matrices. This cone is self-dual:

$$K \;=\; C^\vee \;=\; \mathrm{conv}(T(X)) \;\simeq\; C.$$

Integral for $p_\theta(x)$ is the standard multivariate Gaussian, with

$$A(\theta) \;=\; -\frac{1}{2}\log\det(\theta) + \frac{m}{2}\log(2\pi).$$

The gradient map is matrix inversion $F : C \to K,\; \theta \mapsto \frac{1}{2}\theta^{-1}$.

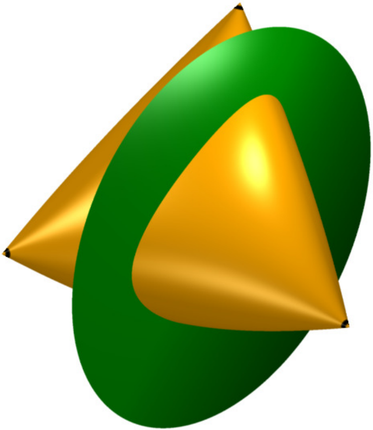
The measure that represents $f(\theta)^{-1/2}$ comes from the *Wishart distribution*, i.e. the distribution of the sample covariance matrix ...

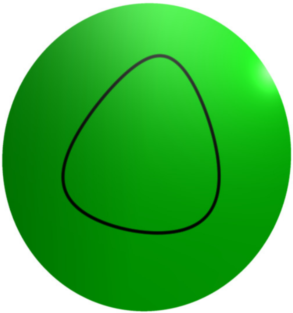
Intersecting with a Subspace

Fix exponential family with rational gradient map $F : C \to K$.

Main case: $F = \nabla f$ where f is hyperbolic

Consider a linear subspace $\mathcal{L} \subset \mathbb{R}^d$ with $C_{\mathcal{L}} := \mathcal{L} \cap C$ nonempty:

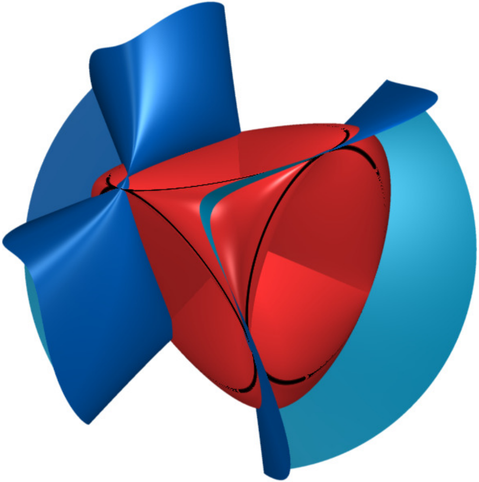


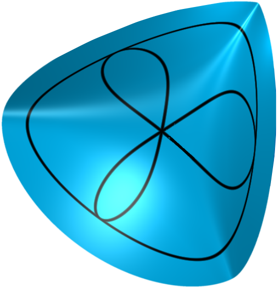


Exponential Varieties

The *exponential variety* is the image under the gradient map:

$$\mathcal{L}^F := \overline{F(\mathcal{L})} \quad \subset \mathbb{P}^{d-1}.$$





Its positive part $\mathcal{L}^F_{\succ 0}$ lives in K.

Convexity and Positivity

Theorem

Let (X, ν, T) be an exponential family with rational gradient map $F : \mathbb{R}^d \dashrightarrow \mathbb{R}^d$, and $\mathcal{L} \subset \mathbb{R}^d$ a linear subspace.

The restricted gradient map $F_{\mathcal{L}}$ is the composition

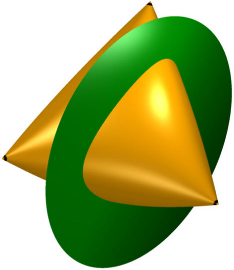
$$C_{\mathcal{L}} \ \subset \ C \ \xrightarrow{F} \ K \ \xrightarrow{\pi_{\mathcal{L}}} \ K_{\mathcal{L}}.$$

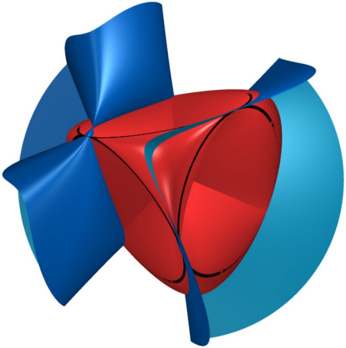
The convex set $C_{\mathcal{L}}$ of canonical parameters maps bijectively to the positive exponential variety $\mathcal{L}^F_{\succ 0}$, and $\mathcal{L}^F_{\succ 0}$ maps bijectively to the interior of the convex set $K_{\mathcal{L}}$ of sufficient statistics.

Maximum Likelihood Estimation for an exponential variety means inverting these two bijections, by solving polynomials.

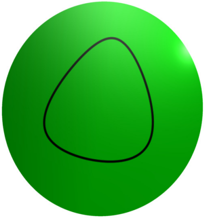
Math question: What is the algebraic degree of this inversion?

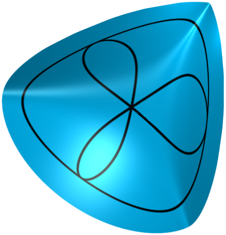
Bijections in Pictures

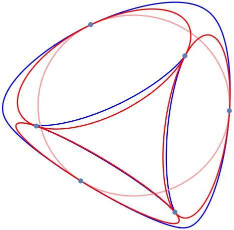




Green maps to *blue* maps to *green*$^\vee$. Inverting this map is MLE.







Graph of Gradient Map

Fix a hyperbolic polynomial $f(\theta)$, and let $X_f \subset \mathbb{P}^{d-1} \times \mathbb{P}^{d-1}$ be the graph of its gradient map ∇f, a variety of dimension $d-1$.

The *gradient multidegree* of f is its class $[X_f]$ in the cohomology

$$H^*\big(\mathbb{P}^{d-1} \times \mathbb{P}^{d-1}; \mathbb{Z}\big) \;=\; \mathbb{Z}[\,s,\,t\,]/\langle s^d, t^d\rangle.$$

Graph of Gradient Map

Fix a hyperbolic polynomial $f(\theta)$, and let $X_f \subset \mathbb{P}^{d-1} \times \mathbb{P}^{d-1}$ be the graph of its gradient map ∇f, a variety of dimension $d-1$.

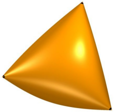
The *gradient multidegree* of f is its class $[X_f]$ in the cohomology

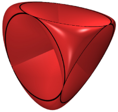
$$H^*\left(\mathbb{P}^{d-1} \times \mathbb{P}^{d-1}; \mathbb{Z}\right) \;=\; \mathbb{Z}[s, t]/\langle s^d, t^d \rangle.$$

If α_i is the cardinality of a linear section $X_f \cap (L_{i-1} \times M_{d-i})$ then

$$[X_f] \;=\; \alpha_d s^{d-1} + \alpha_{d-1} s^{d-2} t + \alpha_{d-2} s^{d-3} t^2 + \cdots + \alpha_2 s t^{d-2} + \alpha_1 t^{d-1}.$$

The leading coefficient α_d is the *gradient degree* of f.





Example: If $f \;=\; \theta_1\theta_2\theta_3 + \theta_1\theta_2\theta_4 + \theta_1\theta_3\theta_4 + \theta_2\theta_3\theta_4$
then $[X_f] \;=\; \mathbf{4}s^3 + 4s^2t + 2st^2 + 1t^3$.

Degrees

Fix a subspace $\mathcal{L} \subset \mathbb{R}^d$ of dimension c. Let $\pi_{\mathcal{L}} : \mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{c-1}$ be the projection with center $\mathcal{L}^{\perp}$. We define

$$\mathrm{MLdegree}(\mathcal{L}^{\nabla f}) \;:=\; \mathrm{degree}\big(\mathcal{L}^{\nabla f} \dashrightarrow \mathbb{P}^{c-1}\big).$$

The *ML degree* is the algebraic complexity of the function that maps sufficient statistics in $K_{\mathcal{L}}$ to the MLE in the model $\mathcal{L}^{\nabla f}_{\succeq 0}$.

Degrees

Fix a subspace $\mathcal{L} \subset \mathbb{R}^d$ of dimension c. Let $\pi_{\mathcal{L}} : \mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{c-1}$ be the projection with center $\mathcal{L}^{\perp}$. We define

$$\mathrm{MLdegree}(\mathcal{L}^{\nabla f}) \;:=\; \mathrm{degree}\big(\mathcal{L}^{\nabla f} \dashrightarrow \mathbb{P}^{c-1}\big).$$

The *ML degree* is the algebraic complexity of the function that maps sufficient statistics in $K_{\mathcal{L}}$ to the MLE in the model $\mathcal{L}^{\nabla f}_{\succeq 0}$.

Theorem

The following inequalities hold for all exponential varieties:

$$\mathrm{MLdegree}(\mathcal{L}^{\nabla f}) \;\leq\; \mathrm{degree}(\mathcal{L}^{\nabla f}) \;\leq\; \textit{the coefficient } \alpha_c \textit{ in } [X_f].$$

Right inequality is an equality for generic subspaces $\mathcal{L}$.
Left inequality is an equality if and only if $\mathcal{L}^{\nabla f} \cap \mathcal{L}^{\perp} = \emptyset$.

We conjecture that $\mathcal{L}^{\nabla f} \cap \mathcal{L}^{\perp} = \emptyset$ holds for generic $\mathcal{L}$.
All four sign combinations occur even for Gaussian graphical models.

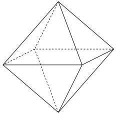
Elementary Symmetric Polynomials

We study the hyperbolic exponential family given

$$E_m(\theta) \quad = \sum_{1 \le i_1 < \cdots < i_m \le d} \theta_{i_1}\theta_{i_2}\cdots\theta_{i_m}$$

An explicit formula is given for the gradient multidegree $[X_{E_m}]$ in terms of *mixed Eulerian numbers*; for instance, for $d = 7$:

$$\begin{array}{rcl}
[X_{E_2}] & = & 1s^6 + 1s^5t + 1s^4t^2 + 1s^3t^3 + 1s^2t^4 + 1st^5 + 1t^6 \\
[X_{E_3}] & = & \mathbf{57}s^6 + 32s^5t + 16s^4t^2 + 8s^3t^3 + 4s^2t^4 + 2st^5 + 1t^6 \\
[X_{E_4}] & = & 302s^6 + \mathbf{222}s^5t + 81s^4t^2 + 27s^3t^3 + 9s^2t^4 + 3st^5 + 1t^6 \\
[X_{E_5}] & = & 302s^6 + 422s^5t + \mathbf{221}s^4t^2 + 64s^3t^3 + 16s^2t^4 + 4st^5 + t^6 \\
[X_{E_6}] & = & 57s^6 + 157s^5t + 170s^4t^2 + \mathbf{90}s^3t^3 + 25s^2t^4 + 5st^5 + 1t^6 \\
[X_{E_7}] & = & 1s^6 + 6s^5t + 15s^4t^2 + 20s^3t^3 + \mathbf{15}s^2t^4 + 6st^5 + 1t^6
\end{array}$$



Given any $\mathcal{L}$, this bounds the degree – and hence the ML degree – of the exponential variety $\mathcal{L}^{\nabla E_m}$. These models are not Gaussian.

Hankel Matrices

Fix the Gaussian family $f = \det(\theta)$. Let $\mathcal{L}$ be the space of $m \times m$ Hankel matrices, so $d = \binom{m+1}{2}$, $c = 2m - 1$.

$C_{\mathcal{L}}$  is the cone of positive definite Hankel matrices.

$$\begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_3 & \theta_4 & \theta_5 \\ \theta_3 & \theta_4 & \theta_5 & \theta_6 \\ \theta_4 & \theta_5 & \theta_6 & \theta_7 \end{pmatrix} \qquad m = 4, c = 7$$

Identify $\mathbb{P}^{c-1}$ with $\{\text{polynomials of degree } 2m{-}2 \text{ in } x\}$.

Hankel Matrices

Fix the Gaussian family $f = \det(\theta)$. Let $\mathcal{L}$ be the space of $m \times m$ Hankel matrices, so $d = \binom{m+1}{2}$, $c = 2m - 1$.

$C_{\mathcal{L}}$  is the cone of positive definite Hankel matrices.

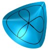
$$\begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_3 & \theta_4 & \theta_5 \\ \theta_3 & \theta_4 & \theta_5 & \theta_6 \\ \theta_4 & \theta_5 & \theta_6 & \theta_7 \end{pmatrix} \qquad m = 4, c = 7$$

Identify $\mathbb{P}^{c-1}$ with $\{$polynomials of degree $2m-2$ in $x\}$.

The map $\pi_{\mathcal{L}} : \mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{c-1}$ is

$$\sigma \mapsto (1, x, x^2, \ldots, x^{m-1}) \cdot \sigma \cdot (1, x, x^2, \ldots, x^{m-1})^T$$

The image $K_{\mathcal{L}}$  of the PSD cone $K = C^\vee$ under $\pi_{\mathcal{L}}$ is the cone of nonnegative polynomials.

Q: Who is the middleman  in these bijections:

$$\{\text{psd Hankel}\} = C_{\mathcal{L}} \xrightarrow{\nabla f} \mathcal{L}^{\nabla f}_{\succ 0} \xrightarrow{\pi_{\mathcal{L}}} K_{\mathcal{L}} = \{\text{nonnegative polynomials}\} \ ?$$

The Other Positive Grassmannian

Theorem

After a linear change of coordinates, the exponential variety $\mathcal{L}^{-1}$ of inverse Hankel matrices equals the Grassmannian $\mathrm{Gr}(2, m+1)$ in its Plücker embedding in $\mathbb{P}^{d-1}$. The ML degree of $\mathcal{L}^{-1}$ equals the degree of $\mathcal{L}^{-1}$, which is the Catalan number $\frac{1}{m}\binom{2m-2}{m-1}$.

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_3 & \theta_4 & \theta_5 \\ \theta_3 & \theta_4 & \theta_5 & \theta_6 \\ \theta_4 & \theta_5 & \theta_6 & \theta_7 \end{bmatrix}^{-1} = \begin{bmatrix} p_{12} & p_{13} & p_{14} & p_{15} \\ p_{13} & p_{14}+p_{23} & p_{15}+p_{24} & p_{25} \\ p_{14} & p_{15}+p_{24} & p_{25}+p_{34} & p_{35} \\ p_{15} & p_{25} & p_{35} & p_{45} \end{bmatrix} \quad p_{ij}p_{kl} - p_{ik}p_{jl} + p_{il}p_{jk} = 0$$

The Other Positive Grassmannian

Theorem

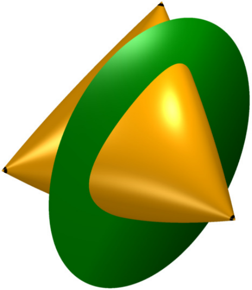
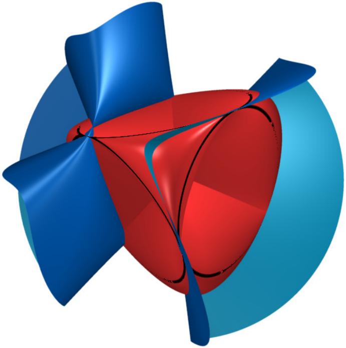
After a linear change of coordinates, the exponential variety $\mathcal{L}^{-1}$ of inverse Hankel matrices equals the Grassmannian $\mathrm{Gr}(2, m+1)$ in its Plücker embedding in $\mathbb{P}^{d-1}$. The ML degree of $\mathcal{L}^{-1}$ equals the degree of $\mathcal{L}^{-1}$, which is the Catalan number $\frac{1}{m}\binom{2m-2}{m-1}$.

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_3 & \theta_4 & \theta_5 \\ \theta_3 & \theta_4 & \theta_5 & \theta_6 \\ \theta_4 & \theta_5 & \theta_6 & \theta_7 \end{bmatrix}^{-1} = \begin{bmatrix} p_{12} & p_{13} & p_{14} & p_{15} \\ p_{13} & p_{14}+p_{23} & p_{15}+p_{24} & p_{25} \\ p_{14} & p_{15}+p_{24} & p_{25}+p_{34} & p_{35} \\ p_{15} & p_{25} & p_{35} & p_{45} \end{bmatrix} \quad p_{ij}p_{kl} - p_{ik}p_{jl} + p_{il}p_{jk} = 0$$

The positive Grassmannian $\mathrm{Gr}(2, m+1)_{\succ 0}$ consists of positive definite *Bézout matrices*. These represent pairs of polynomials in x of degree $m-1$ whose roots are all real and interlace.

Open Problems: *What about higher Grassmannians?*
... generalized Hankel matrices (catalecticants)?
... sum of square polynomials in more variables?

Invitation to Read

Abstract. Exponential varieties arise from exponential families in statistics. These real algebraic varieties have strong positivity and convexity properties, generalizing those of toric varieties and their moment maps. Another special class, including Gaussian graphical models, are varieties of inverses of symmetric matrices satisfying linear constraints. We develop a general theory of exponential varieties, with focus on those defined by hyperbolic polynomials and their integral representations on the hyperbolicity cone. We compare multidegrees and ML degrees of the gradient map for such polynomials.

The End