The Squared Grassmannian

Hannah Friedman
University of Washington Combinatorics and Geometry Seminar
January 15, 2025

Outline

- Likelihood Geometry of Determinantal Point Processes (with Bernd Sturmfels and Maksym Zubkov)

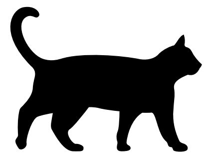
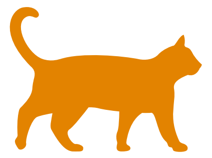
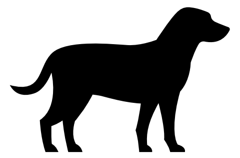
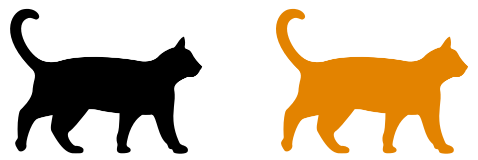
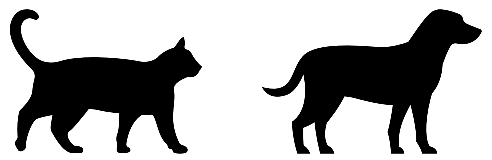
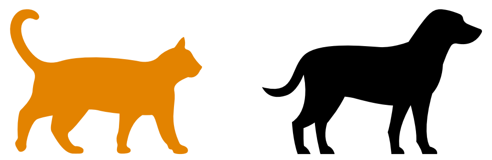
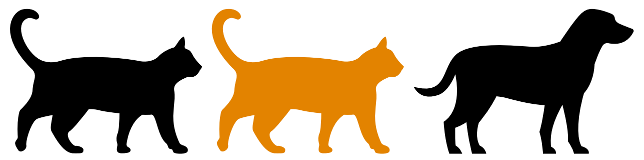

- The Two Lives of the Grassmannian (with Karel Devriendt, Bernhard Reinke, and Bernd Sturmfels)



- Likelihood Geometry of the Squared Grassmannian

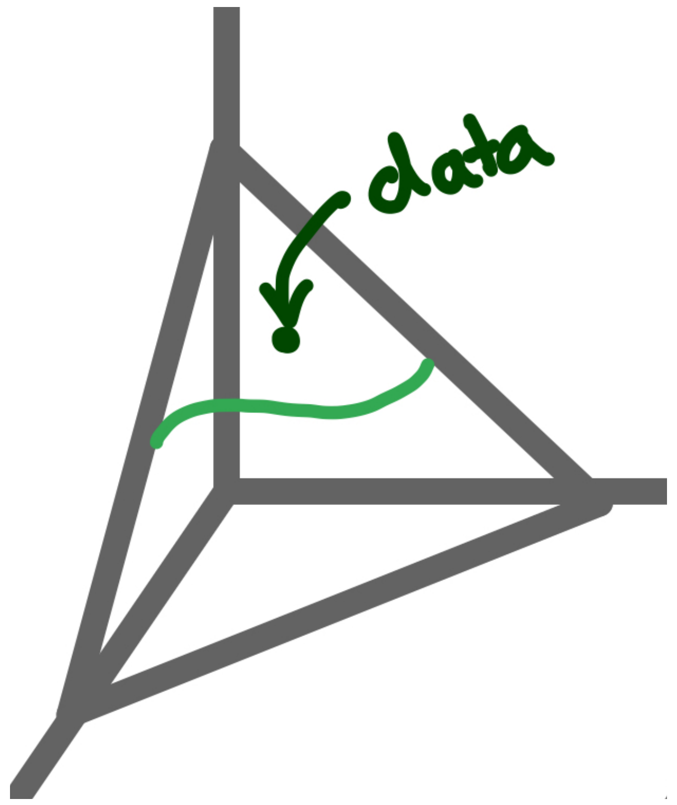
Jackie walks into an animal shelter and adopts some subset of animals at the shelter every day for 100 days. Every day, she decides which animals to take home by sampling from an unknown probability distribution.

Maximum Likelihood Estimation

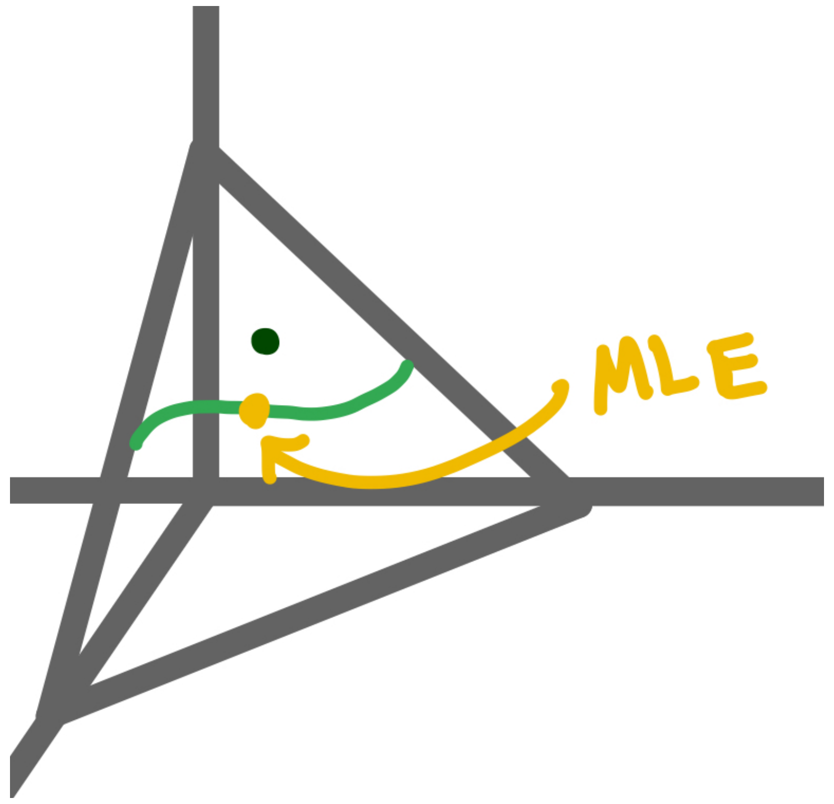
∅	5
	9
	7
	5
	9
	21
	36
	8

Since Jackie prefers to take home a pair of different animals, we assume that Jackie is sampling from a specific type of distribution called a **determinantal point process** (DPP).

Given:

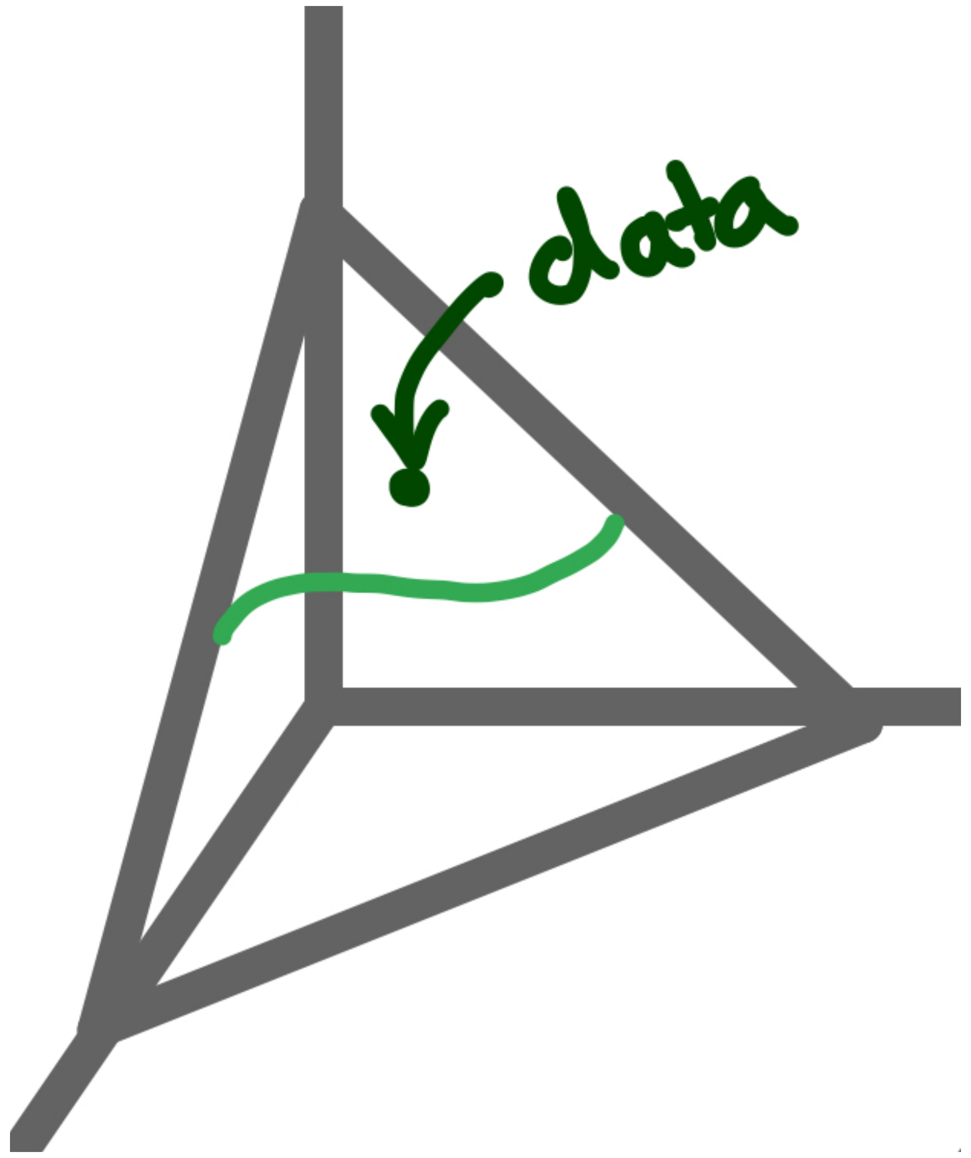


Find:

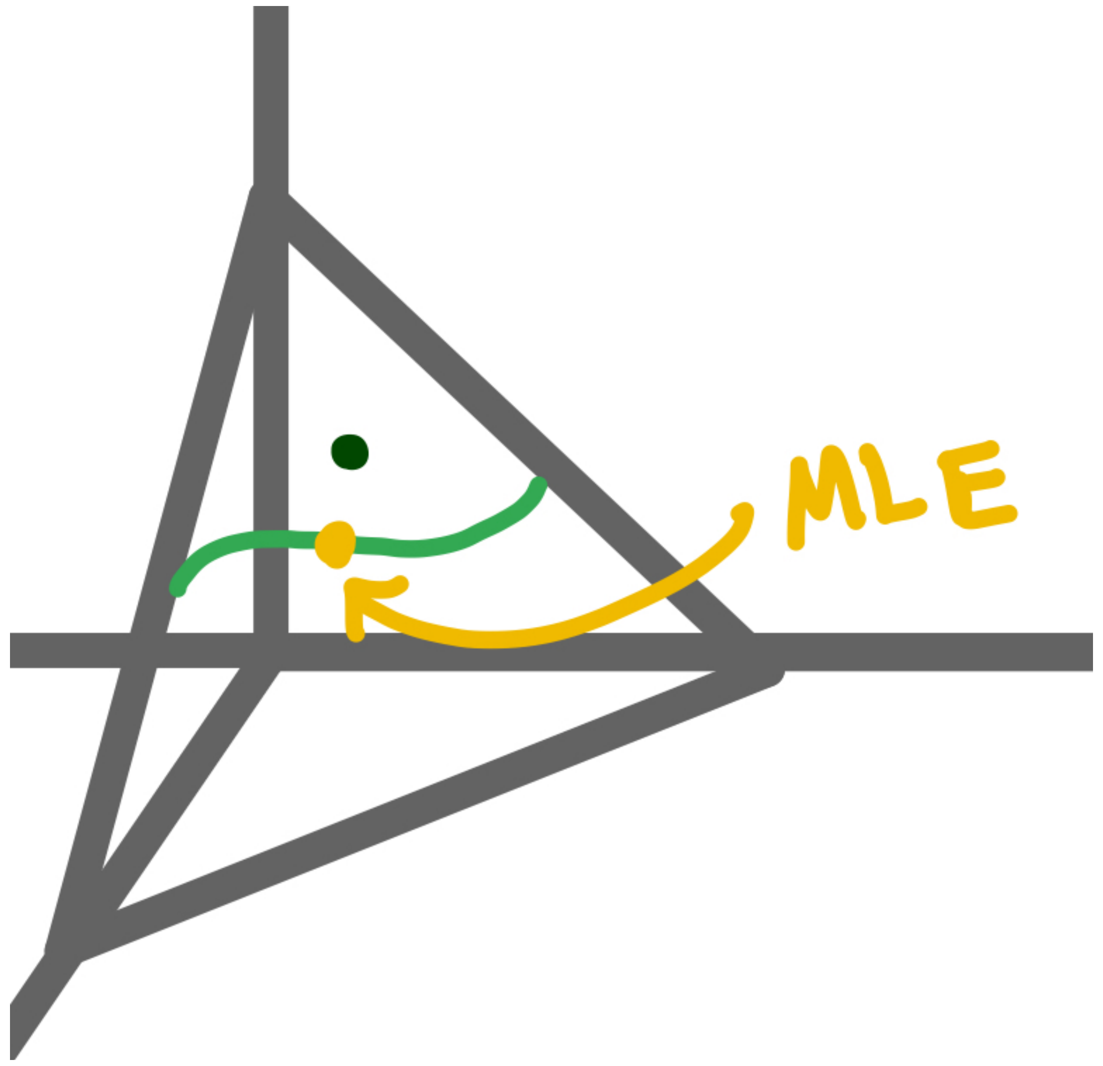


Maximum Likelihood Estimation

Given:



Find:



The maximum likelihood estimate is the point q which maximizes the log-likelihood function:

$$\underset{q \in V}{L_u(q)} = \sum_i u_i \log(q_i) - \left(\sum_i u_i\right) \log\left(\sum_i q_i\right).$$

Theorem (Huh-Sturmfels, 2014) The number of critical points of $L_u(q)$ is generically finite and does not depend on u. This number is called the **maximum likelihood degree** (ML degree) of V.

Motivating the Maximum Likelihood Degree

$$\max_{q \in V} L_u(q) = \sum_i u_i \log(q_i) - \left(\sum_i u_i\right) \log\left(\sum_i q_i\right).$$

Theorem (Huh-Sturmfels, 2014) The number of critical points of $L_u(q)$ is generically finite and does not depend on u. This number is called the **maximum likelihood degree** (ML degree) of V.

1. The more critical points there are, the harder the problem is to solve. The ML degree is an algebraic measure of the **difficulty of the problem**.
2. When numerically computing the solution to such an optimization problem, a heuristic stopping criterion is applied. Knowing the number of solutions a priori means that we don't need to wait until the criterion is met, so the **computation is much faster**.

Determinantal Point Processes

Let P be a real, symmetric matrix with eigenvalues in $[0,1]$. A ***determinantal point process*** with kernel P is a random variable Z with state space $2^{[n]}$ such that

$$\mathbb{P}[I \subseteq Z] = \det(P_I)$$

where P_I is the $d \times d$ principal submatrix of P obtained from the d rows and columns indexed by I.

Example (n = 3).

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix}$$

$$\mathbb{P}[\{2\} \subseteq Z] = p_{22}$$

$$\mathbb{P}[\{1,3\} \subseteq Z] = \det \begin{pmatrix} p_{11} & p_{13} \\ p_{13} & p_{33} \end{pmatrix} = p_{11}p_{33} - p_{13}^2$$

For maximum likelihood estimation, we need an explicit expression for the probability of observing a given set.

Möbius Inversion & L-Ensembles

If P is the kernel of a DPP whose eigenvalues are in $(0,1)$, then we define $\Theta = P(\mathrm{Id}_n - P)$ so that

$$\mathbb{P}[I \subseteq Z] = \det(P_I) \qquad\qquad \mathbb{P}[I = Z] = \frac{\det(\Theta_I)}{\det(\Theta + \mathrm{Id}_n)}.$$

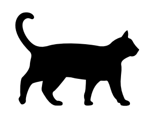
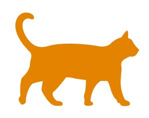
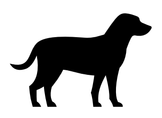
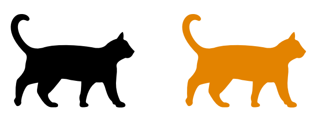
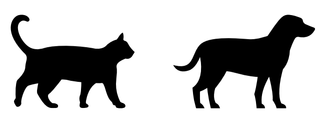
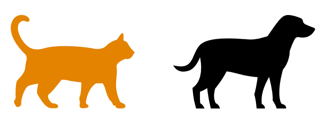
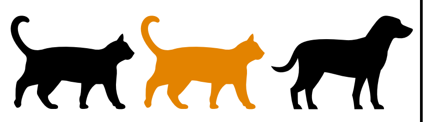
Implicit: $$L_u(q) = \sum_i u_i \log(q_i) - \left(\sum_i u_i\right) \log\left(\sum_i q_i\right) \qquad q \in V_n$$

V_n is the ***hyperdeterminantal variety*** (Oeding, 2011) and (Al Ahmadieh-Vinzant, 2024)

Parametric: $$L_u(\Theta) = \sum_{I \subseteq [n]} u_I \log(\det(\Theta_I)) - \left(\sum_{I \subseteq [n]} u_I\right) \log(\det(\Theta + \mathrm{Id}_n))$$

Example. $u = [5,9,7,5,9,21,36,8]$ $$\Theta = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

$$L_u(\Theta) = 5\log(1) + 9\log(a) + 7\log(d) + 5\log(f)$$
$$+9\log(ad - b^2) + 21\log(af - c^2) + 36\log(df - e^2) + 8\log(\det(\Theta))$$

∅	5
	9
	7
	5
	9
	21
	36
	8

Möbius Inversion & L-Ensembles

n	ML Degree(V_n)
1	1
2	1
3	13

$$L_u(\Theta) = 5\log(1) + 9\log(a) + 7\log(d) + 5\log(f)$$
$$+9\log(ad - b^2) + 21\log(af - c^2) + 36\log(df - e^2) + 8\log(\det(\Theta))$$

59 critical points:

$13 \cdot 2^2$ $\quad (1 \cdot 2^1)(1 \cdot 2^0)$ $\quad (1 \cdot 2^1)(1 \cdot 2^0)$ $\quad (1 \cdot 2^1)(1 \cdot 2^0)$ $\quad 1 \cdot 2^0$

Theorem (F-Sturmfels-Zubkov, 2023)

The critical points $\hat{\Theta}$ of the parametric log-likelihood function are found by solving various likelihood equations on submodels. If u is generic, their count is

$$\sum_{\pi \in \mathcal{P}_n} \prod_{i=1}^{k} (2^{|\pi_i|-1} \text{ML Degree}(\text{V}_{|\pi_i|})) .$$

Maximum Likelihood Estimation for DPPs

	L-Ensemble	Projection DPP
Name:	L-Ensemble	Projection DPP
Eigenvalues of P:	in (0,1)	in {0,1}
Model (variety):	Hyperdeterminantal variety	Squared Grassmannian
Parametric critical points:	$\sum_{\pi \in \mathscr{P}_n} \prod_{i=1}^{k} (2^{\lvert\pi_i\rvert - 1} \mathrm{ML\,Degree}(V_{\lvert\pi_i\rvert}))$.	$2^{n-1}\mathrm{ML\,Deg}(\mathrm{sGr}(2,n))$
ML Degrees:	1, 13, 3526, >29.5 million,…	d=2: 3, 12, 60, 360, 2520, … d=3: 12, 552, 73440, …

The Two Lives of the Grassmannian

The Grassmannian $\mathrm{Gr}(d, n)$ is the space of d-subspaces of n-space.

What's the best way to work with $\mathrm{Gr}(d, n)$ as an algebraic variety?

Plücker Coordinates

Pure Math
Projective Variety
Algebraic Combinatorics
Particle Physics
⋮

Orthogonal Projection Matrices

Applied Math
Affine Variety
Numerics and Statistics
Data Science
⋮

Plücker Coordinates

L : d-dimensional subspace of $\mathbb{R}^n$ $\quad$ A : $d \times n$ matrix whose rows span L

The ***Plücker coordinates*** for L are $x_I = \det(A_I)$ for $I \subseteq [n], |I| = d$, where A_I is the $d \times d$ submatrix of A formed by taking the columns indexed by I.

Example (d = 2, n = 5).

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \end{pmatrix}$$

$$x_{ij} = \begin{vmatrix} a_{i1} a_{i2} \\ a_{j1} a_{j2} \end{vmatrix} = a_{i1}a_{j2} - a_{j1}a_{i2}$$

Relations:

$$x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23} = 0$$
$$x_{12}x_{35} - x_{13}x_{25} + x_{15}x_{23} = 0$$
$$x_{12}x_{45} - x_{14}x_{25} + x_{15}x_{24} = 0$$
$$x_{13}x_{45} - x_{14}x_{35} + x_{15}x_{34} = 0$$
$$x_{23}x_{45} - x_{24}x_{35} + x_{25}x_{34} = 0$$

Orthogonal Projection Matrices

L : d-dimensional subspace of $\mathbb{R}^n$ $\quad$ A : $n \times d$ matrix whose columns span L

The $n \times n$ matrix $P = A(A^T A)^{-1} A^T$ is the unique orthogonal projection matrix onto L.

The matrix P satisfies

$$P^T = P,\ P^2 = P \text{ and } \operatorname{trace}(P) = d\,.$$

Theorem (Devriendt, F., Reinke, Sturmfels 2024).

$$\mathcal{I}(\mathrm{pGr}(d,n)) = \langle P^2 - P, \operatorname{trace}(P) - d \rangle.$$

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{pmatrix}$$

Moving Between the Two Lives

Take maximal minors of d linearly independent rows of P

Projection matrix P ⟶ Plücker coordinates $\mathbf{x}$

$$p_{ij} = \frac{\sum_{K \in \binom{[n]}{k-1}} x_{iK} x_{jK}}{\sum_{I \in \binom{[n]}{k}} x_I^2} \quad \text{(Bloch-Karp, 2023)}$$

Corollary (Devriendt-F-Reinke-Sturmfels, 2024). $\det(P_I) = \dfrac{x_I^2}{\sum_{J \in \binom{[n]}{d}} x_{k\ell}^2}.$

The Squared Grassmannian

Definition.

The ***squared Grassmannian*** $\mathrm{sGr}(d,n)$ is the image of the Grassmannian $\mathrm{Gr}(d,n) \subset \mathbb{P}^{\binom{n}{d}-1}$ in its Plücker embedding under the map

$$\mathrm{Gr}(d,n) \to \mathbb{P}^{\binom{n}{d}-1}$$
$$(x_I)_{I\in\binom{[n]}{d}} \mapsto (x_I^2)_{I\in\binom{[n]}{d}}$$

The squared Grassmannian satisfies

$$\dim(\mathrm{sGr}(d,n)) = d(n-d), \qquad \mathrm{degree}(\mathrm{sGr}(d,n)) = 2^{(d-1)(n-d-1)}\mathrm{degree}(\mathrm{Gr}(d,n))\,.$$

Theorem (Devriendt-F-Reinke-Sturmfels, 2024).

The prime ideal $\mathcal{I}(\mathrm{sGr}(2,n))$ is generated by 4-minors of

$$\begin{pmatrix} 0 & q_{12} & \cdots & q_{1n} \\ q_{12} & 0 & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1n} & q_{2n} & \cdots & 0 \end{pmatrix}.$$

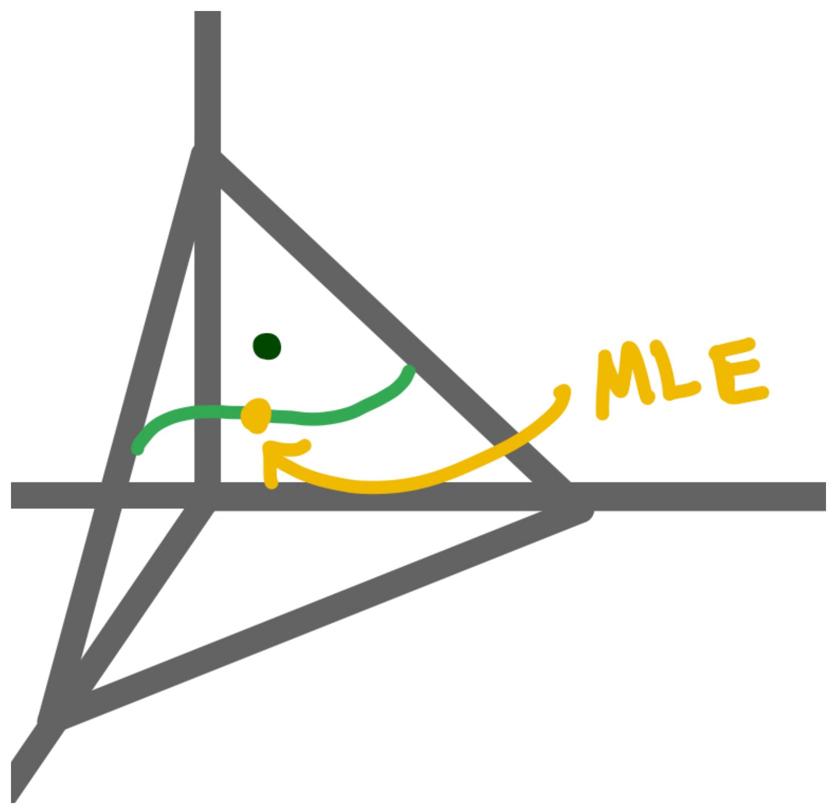
Theorem (Al Ahmadieh-Vinzant, 2024).

The squared Grassmannian $\mathrm{sGr}(d,n)$ is cut out by quartics derived from hyperdeterminants.

Projection Determinantal Point Processes

If P is an orthogonal projection matrix, i.e., $P \in \mathrm{pGr}(d,n)$, then P defines a special kind of determinantal point process, namely a ***projection determinantal point process***.

$$\mathbb{P}[I = Z] = \begin{cases} \det(P_I) = \dfrac{x_I^2}{\sum_{J \in \binom{[n]}{d}} x_J^2} & \text{if } |I| = d \\ 0 & \text{else} \end{cases}$$



Corollary (Devriendt-F-Reinke-Sturmfels, 2024).

The projection determinantal point process is the discrete statistical model on the state space $\binom{[n]}{d}$ whose underlying algebraic variety is the squared Grassmannian $\mathrm{sGr}(d,n)$.

Jackie walks into a new animal shelter and adopts 2 of the 4 animals at the shelter every day for 100 days. Every day, she decides which animals to take home by sampling from an unknown probability distribution.

Three Log-Likelihood Functions:

$$\mathbb{P}[Z = \{i,j\}] = \boxed{\det(P_{ij})} = \boxed{q_{ij}} = \boxed{\frac{x_{ij}^2}{\sum_{1 \le k \le \ell \le n} x_{k\ell}^2}}$$

$$L_u(P) = \sum_{i,j} u_{ij} \log(\det(P_{ij})) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} \det(P_{ij})\right) \qquad P \in \mathrm{pGr}(d,n)$$

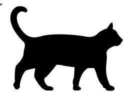
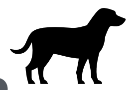
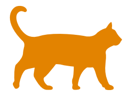
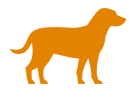
Implicit: $$L_u(q) = \sum_{i,j} u_{ij} \log(q_{ij}) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} q_{ij}\right) \qquad q \in \mathrm{sGr}(2,n)$$

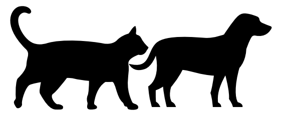
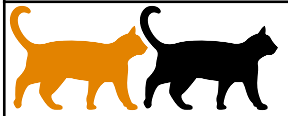
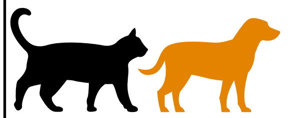
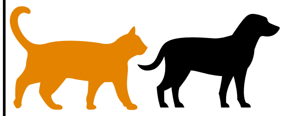
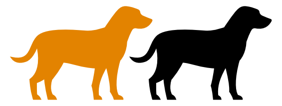
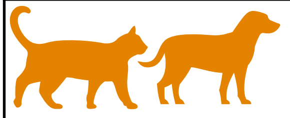
Parametric: $$L_u(A) = \sum_{i,j} u_{ij} \log(\det(A_{ij})^2) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} \det(A_{ij})^2\right) \qquad A = \begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2n} \end{pmatrix}$$

Computing the Maximum Likelihood Estimate

To compute the maximum likelihood estimate, we find the matrix A maximizing the log-likelihood function

$$L_u(A) = \sum_{i,j} u_{ij} \log(\det(A_{ij})^2) - \left(\sum_{i,j} u_{ij} \right) \log \left(\sum_{i,j} \det(A_{ij})^2 \right)$$

Example (n = 4). Sample 2-element subsets from {, , , }.

	14
	11
	26
	24
	9
	16

$$u = [14,11,26,24,9,16]$$

$$A = \begin{pmatrix} 1 & 0 & a_{13} & a_{14} \\ 0 & 1 & a_{23} & a_{24} \end{pmatrix}$$

$$L_u(A) = 14\log(1) + 11\log(a_{23}^2) + 26\log(a_{24}^2) + 24\log(a_{13}^2) + 9\log(a_{14}^2)$$
$$+16\log((a_{13}a_{24} - a_{14}a_{23})^2) - 100\log(1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2)$$

Computing the Maximum Likelihood Estimate

$$L_u(A) = 14\log(1) + 11\log(a_{23}^2) + 26\log(a_{24}^2) + 24\log(a_{13}^2) + 9\log(a_{14}^2) + 16\log((a_{13}a_{24} - a_{14}a_{23})^2)$$
$$-100\log(1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2)$$

1.

$$\frac{\partial L_u}{\partial a_{13}} = \frac{48}{a_{13}} + \frac{32a_{24}}{a_{13}a_{24} - a_{14}a_{23}} - 200\frac{a_{13} + a_{24}(a_{13}a_{12} - a_{14}a_{23})}{1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2} = 0$$

$$\frac{\partial L_u}{\partial a_{14}} = \frac{18}{a_{14}} - \frac{32a_{23}}{a_{13}a_{24} - a_{14}a_{23}} - 200\frac{a_{14} - a_{23}(a_{13}a_{12} - a_{14}a_{23})}{1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2} = 0$$

$$\frac{\partial L_u}{\partial a_{23}} = \frac{22}{a_{23}} - \frac{32a_{14}}{a_{13}a_{24} - a_{14}a_{23}} - 200\frac{a_{23} - a_{14}(a_{13}a_{12} - a_{14}a_{23})}{1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2} = 0$$

$$\frac{\partial L_u}{\partial a_{24}} = \frac{52}{a_{24}} + \frac{32a_{13}}{a_{13}a_{24} - a_{14}a_{23}} - 200\frac{a_{24} + a_{13}(a_{13}a_{12} - a_{14}a_{23})}{1 + a_{23}^2 + a_{24}^2 + a_{13}^2 + a_{14}^2 + (a_{13}a_{24} - a_{14}a_{23})^2} = 0$$

2. Apply **monodromy_solve** in **HomotopyContinuation.jl**.

$$\begin{pmatrix} 1 & 0 & 1.308 & 0.802 \\ 0 & 1 & 0.886 & 1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1.308 & -0.802 \\ 0 & 1 & -0.886 & 1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1.308 & -0.802 \\ 0 & 1 & 0.886 & 1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1.308 & -0.802 \\ 0 & 1 & 0.886 & -1.361 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1.308 & -0.802 \\ 0 & 1 & -0.886 & -1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1.308 & 0.802 \\ 0 & 1 & 0.886 & -1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1.308 & 0.802 \\ 0 & 1 & -0.886 & -1.361 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1.308 & 0.802 \\ 0 & 1 & -0.886 & 1.361 \end{pmatrix}$$

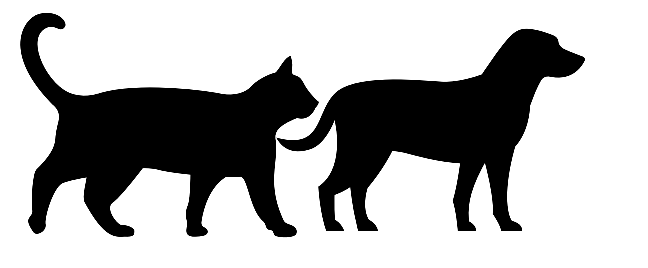
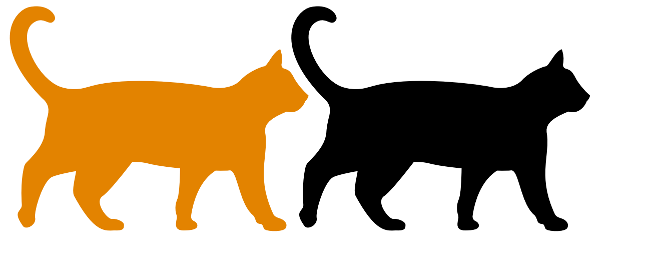
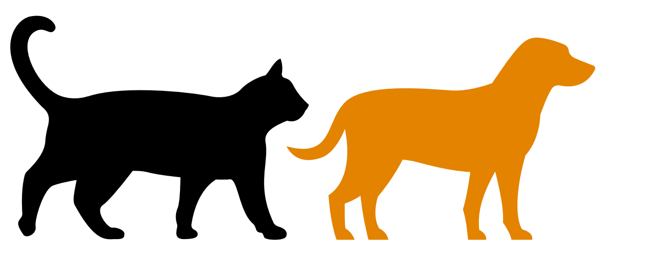
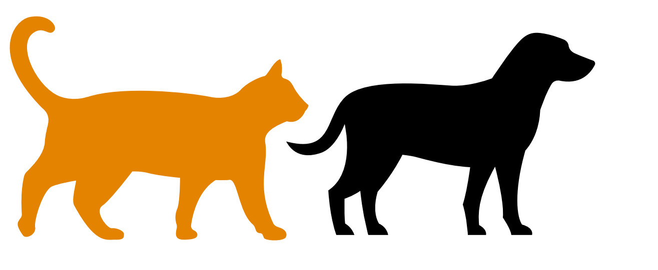
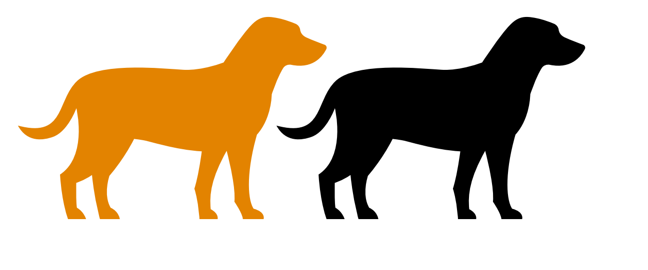
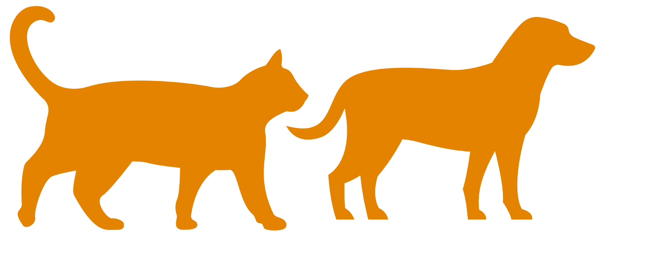
$$\begin{pmatrix} 1 & 0 & 0.839 & -0.507 \\ 0 & 1 & 0.584 & 0.888 \end{pmatrix} \times 8 \qquad \begin{pmatrix} 1 & 0 & 1.320 & 1.690 \\ 0 & 1 & 1.759 & 1.408 \end{pmatrix} \times 8$$

24 critical points

$\det(A_{ij})^2$ →

$$\begin{pmatrix} 1 \\ 0.786 \\ 1.852 \\ 1.710 \\ 0.643 \\ 1.143 \end{pmatrix}, \begin{pmatrix} 1 \\ 0.341 \\ 0.788 \\ 0.704 \\ 0.257 \\ 1.083 \end{pmatrix}, \begin{pmatrix} 1 \\ 3.093 \\ 1.982 \\ 1.744 \\ 2.855 \\ 1.238 \end{pmatrix}$$

Three Kinds of MLEs

	14
	11
	26
	24
	9
	16

$$A^* = \begin{pmatrix} 1 & 0 & 1.308 & 0.802 \\ 0 & 1 & 0.886 & 1.361 \end{pmatrix}$$

(unique up to flipping some signs)

$$P^* = \begin{pmatrix} 0.51 & -0.3154 & 0.3872 & -0.0204 \\ -0.3154 & 0.47 & 0.0041 & 0.3867 \\ 0.3872 & 0.0041 & 0.51 & 0.3161 \\ -0.0204 & 0.3867 & 0.3161 & 0.51 \end{pmatrix}$$

(unique up to flipping some signs)

$$q^* = \begin{pmatrix} 1 \\ 0.786 \\ 1.852 \\ 1.710 \\ 0.643 \\ 1.143 \end{pmatrix} \sim \begin{pmatrix} 0.14 \\ 0.110 \\ 0.259 \\ 0.239 \\ 0.090 \\ 0.160 \end{pmatrix}$$

(unique)

Likelihood Geometry of the Squared Grassmannian

Theorem (F, 2024).

The number of complex critical points of the parametric log-likelihood function

$$L_u(A) = \sum_{i,j} u_{ij} \log(\det(A_{ij})^2) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} \det(A_{ij})^2\right) \quad \text{is } 2^{n-2}(n-1)!.$$

Corollary (F, 2024).

The ML degree of the squared Grassmannian $\mathrm{sGr}(2,n)$ is $\dfrac{(n-1)!}{2}$.

proof of corollary.

The parameterization

$$\begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2(n-2)} & a_{2n} \end{pmatrix} \mapsto (1, a_{23}^2, a_{24}^2 \ldots, (a_{1(n-1)}a_{2n} - a_{2(n-2)}a_{1n})^2)$$

of the squared Grassmannian is 2^{n-1}-to-1. ∎

Likelihood Geometry of the Squared Grassmannian

Theorem (F, 2024).

The number of complex critical points of the parametric log-likelihood function

$$L_u(A) = \sum_{i,j} u_{ij} \log(\det(A_{ij})^2) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} \det(A_{ij})^2\right) \quad \text{is } 2^{n-2}(n-1)!.$$

proof of theorem.

Theorem (Huh, 2013).

If the very affine variety $X \backslash \mathscr{H}$ is smooth of dimension d, then the ML degree of X is the signed Euler characteristic $(-1)^d \chi(X \backslash \mathscr{H})$.

$$A_n = \begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2(n-2)} & a_{2n} \end{pmatrix}$$

$p_{ij} =$ ij-minor of A_n

$$Q_n = \sum_{1 \le i < j \le n} p_{ij}^2$$

$$X_n = \left\{ A_n \in \mathbb{C}^{2(n-2)} : Q_n\left(\prod_{1 \le i < j \le n} p_{ij} \right) \neq 0 \right\}$$

Need to show that $\chi(X_n) = 2^{n-2}(n-1)!$

Likelihood Geometry of the Squared Grassmannian

$A_n = \begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2(n-2)} & a_{2n} \end{pmatrix}$ $\quad p_{ij} = ij\text{-minor of } A_n \quad Q_n = \sum_{1\le i<j\le n} p_{ij}^2 \quad X_n = \left\{ A_n \in \mathbb{C}^{2(n-2)} : Q_n\left(\prod_{1\le i<j\le n} p_{ij}\right) \neq 0 \right\}$

Need to show that $\chi(X_n) = 2^{n-2}(n-1)!$

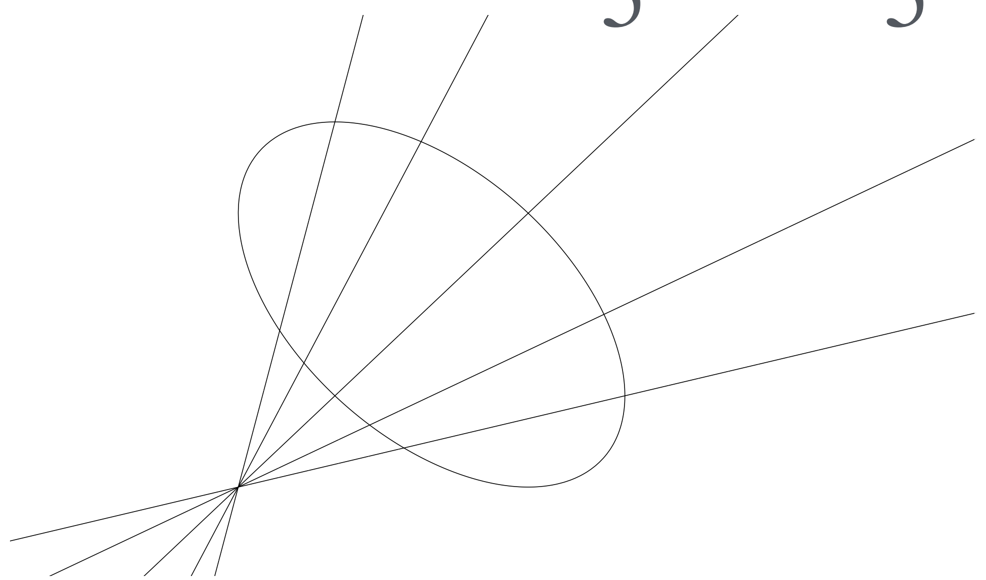
Use induction and the projection $\pi_{n+1}: X_{n+1} \to X_n$ to show that $\chi(X_{n+1}) = 2n\chi(X_n)$.

$$\begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_1 & a_{1(n+1)} \\ 0 & 1 & a_{23} & \cdots & a_2 & a_{2(n+1)} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2n} \end{pmatrix}$$

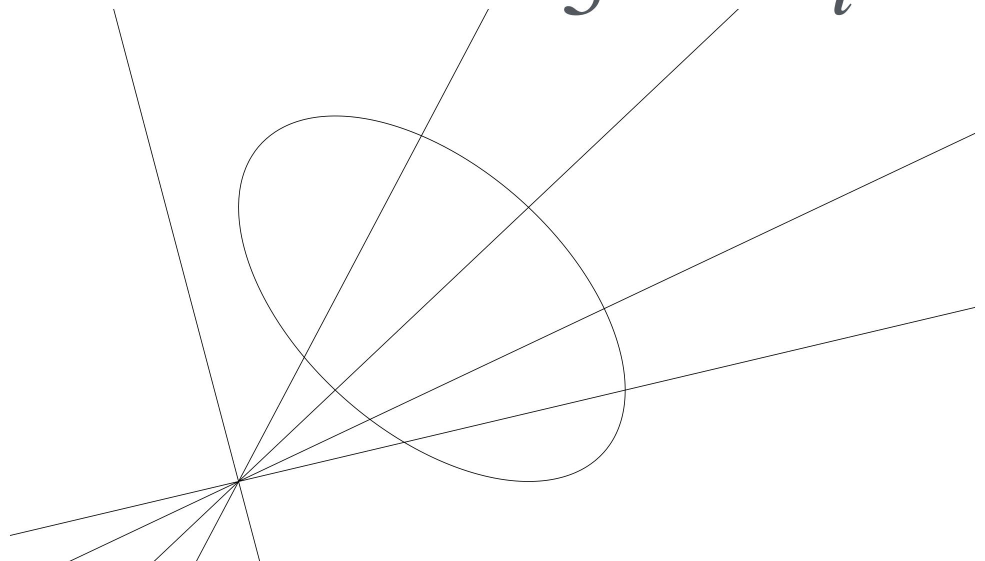
The map π_{n+1} is a stratified fibration with stratification

$$\mathcal{S} = \{X_n\} \cup \{S_i : i \in [n]\} \cup \{S_i \cap S_j : i,j \in [n]\} \quad \text{where} \quad S_i = \{A_n \in X_n \mid \sum_{j=1}^{n} p_{ij}^2 = 0\}.$$

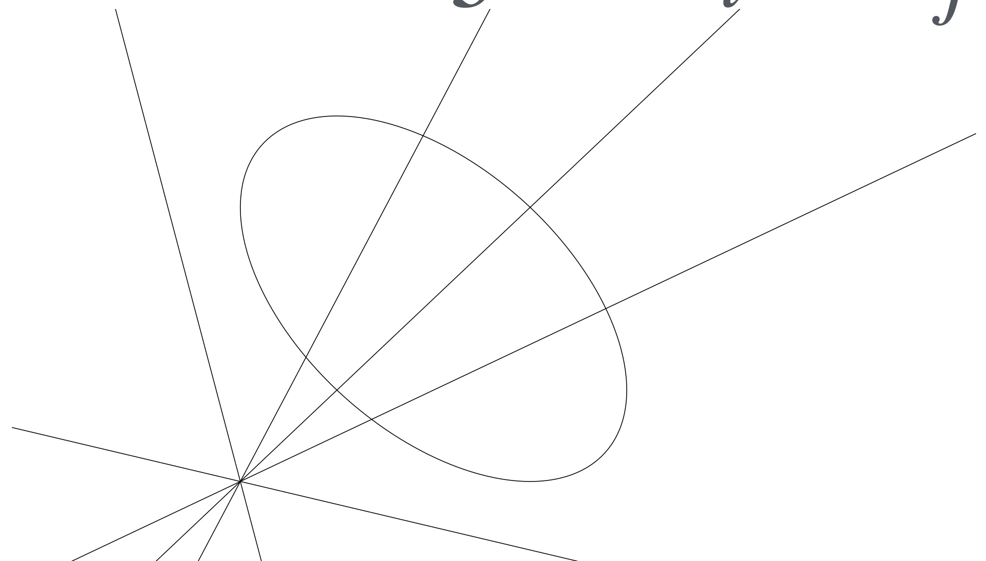
Fiber of $A_5 \in X_5$:



Fiber of $A_5 \in S_i$:

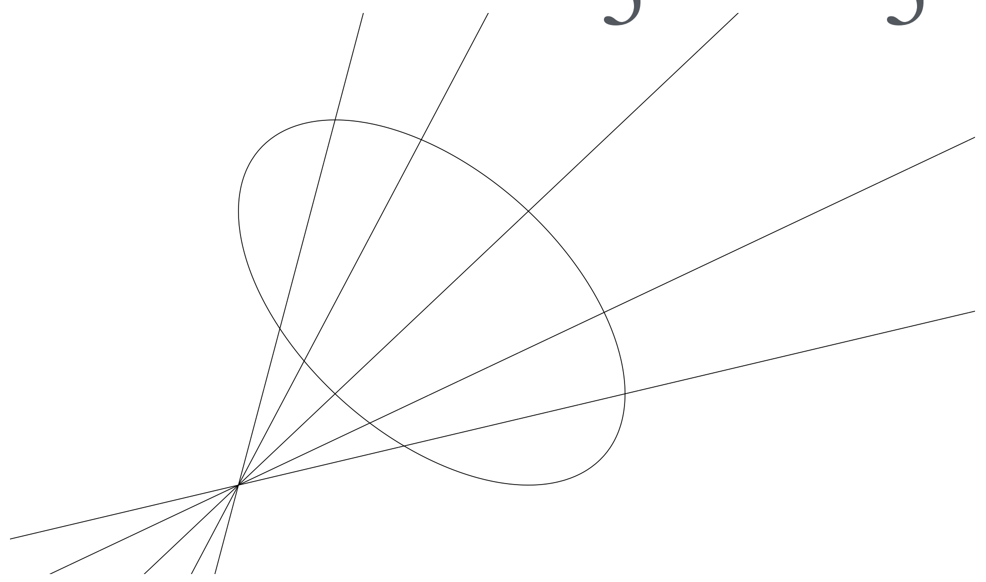


Fiber of $A_5 \in S_i \cap S_j$:



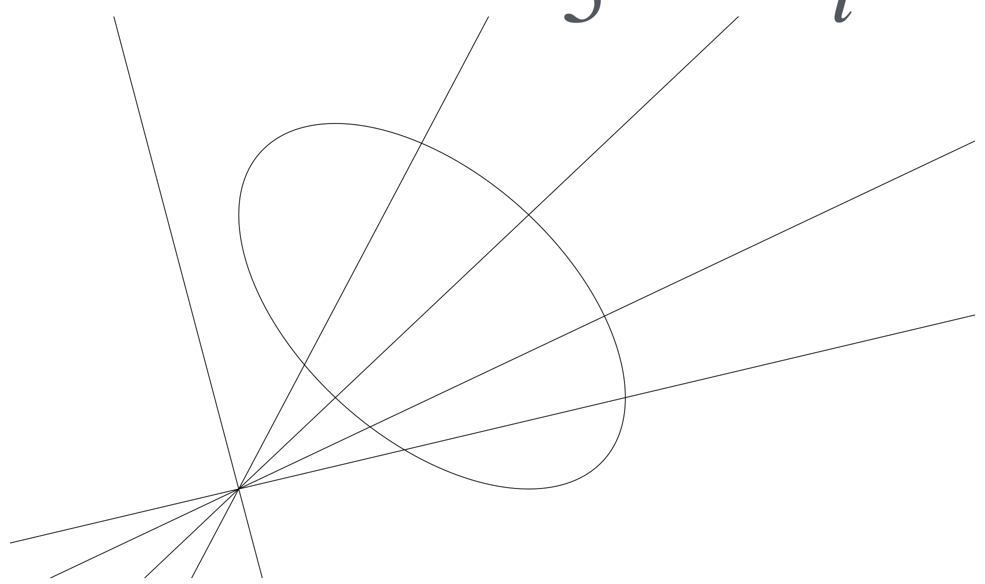
Likelihood Geometry of the Squared Grassmannian

Fiber of $A_5 \in X_5$:



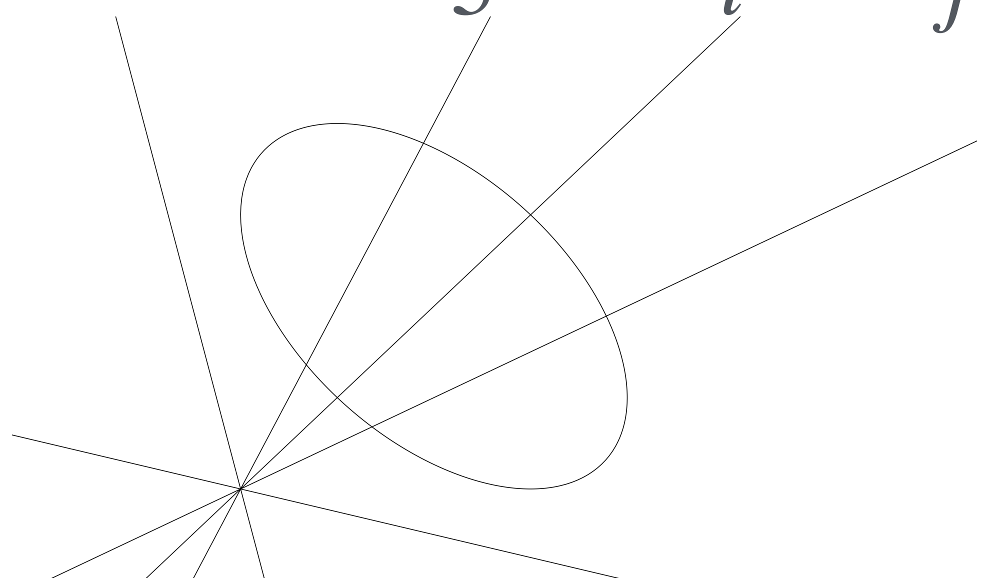
$\chi(F_{X_n}) = 2n$

Fiber of $A_5 \in S_i$:



$\chi(F_i) = 2n - 2$

Fiber of $A_5 \in S_i \cap S_j$:



$\chi(F_{ij}) = 2n - 4$

$$\chi(X_{n+1}) = \chi(F_{X_n})\chi(X_n) + \sum_{i=1}^{n} \chi(S_i) \sum_{S' \in \{S_i, X_n\}} \mu(S_i, S')(\chi(F_{S'}) - \chi(F_{X_n}))$$

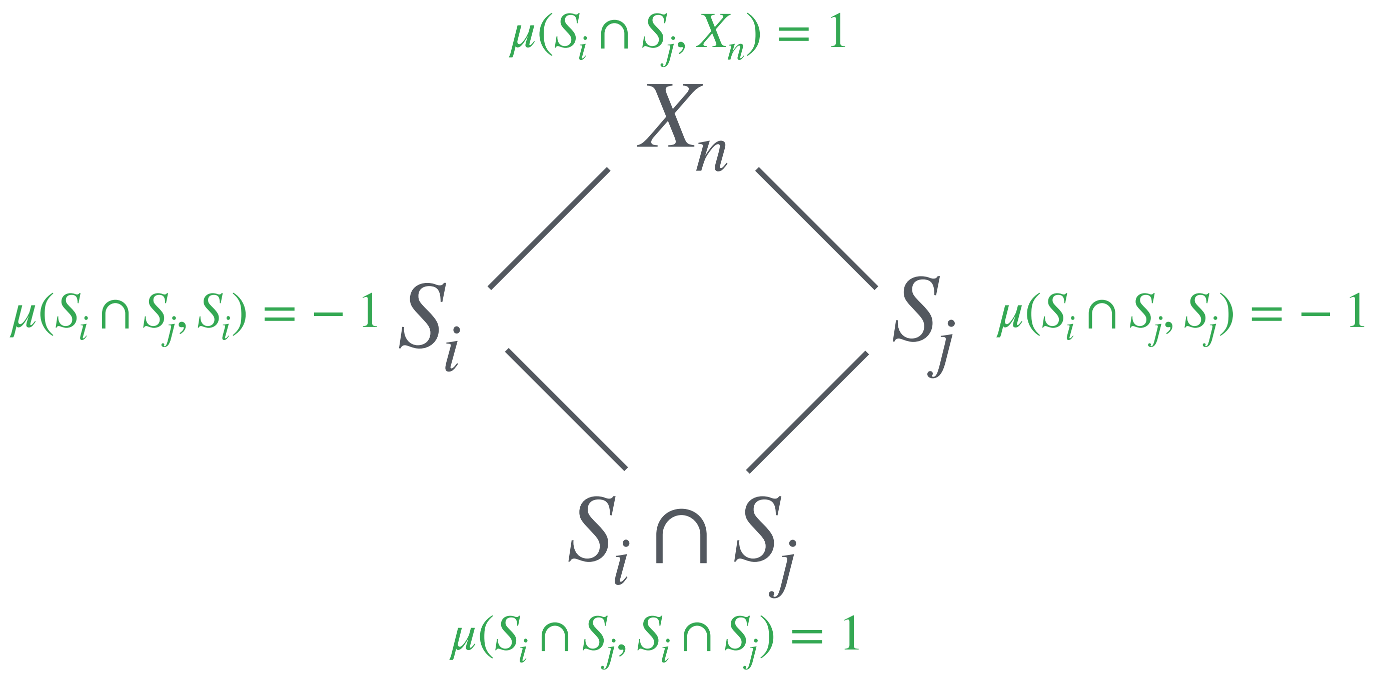
$$+ \sum_{1 \le i < j \le n} \chi(S_i \cap S_j) \sum_{S' \in \{S_i \cap S_j, S_i, S_j, X_n\}} \mu(S_i \cap S_j, S')(\chi(F_{S'}) - \chi(F_{X_n}))$$

0

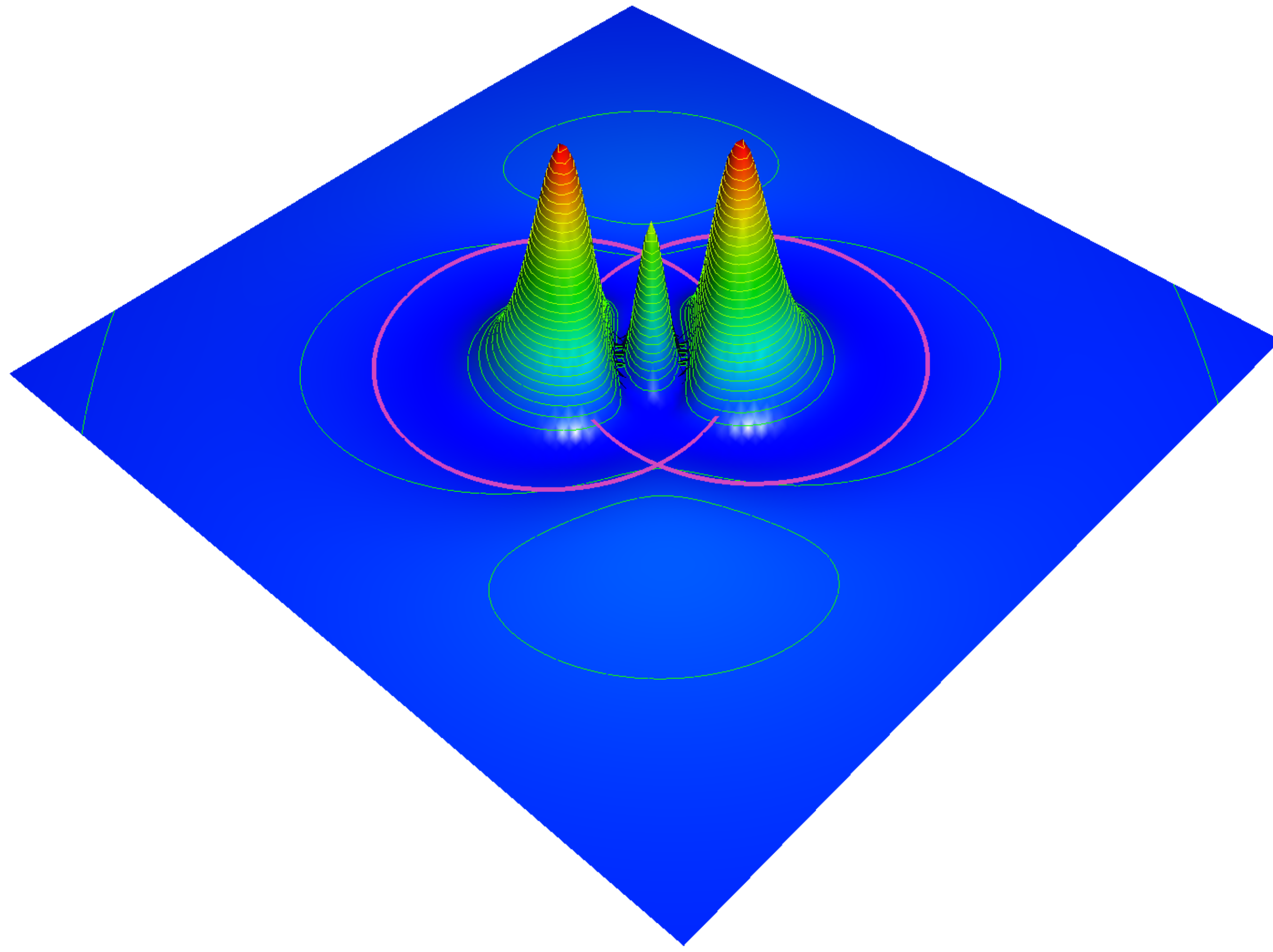
0

$(2n - 2n) - 2(2n - 2 - 2n) + (2n - 4 - 2n) = 0$

$$= \chi(F_{X_n})\chi(X_n) = 2n(\chi(X_n)) \quad \blacksquare$$



Real and Positive Solutions



Example

Parametric Critical Points

$$\begin{pmatrix} 1 & 0 & 1.308 & 0.802 \\ 0 & 1 & 0.886 & 1.361 \end{pmatrix} \times 8$$
$$\begin{pmatrix} 1 & 0 & 0.839 & -0.507 \\ 0 & 1 & 0.584 & 0.888 \end{pmatrix} \times 8$$
$$\begin{pmatrix} 1 & 0 & 1.320 & 1.690 \\ 0 & 1 & 1.759 & 1.408 \end{pmatrix} \times 8$$

Implicit Critical points

$$\begin{pmatrix} 1 \\ 0.786 \\ 1.852 \\ 1.710 \\ 0.643 \\ 1.143 \end{pmatrix}, \begin{pmatrix} 1 \\ 0.341 \\ 0.788 \\ 0.704 \\ 0.257 \\ 1.083 \end{pmatrix}, \begin{pmatrix} 1 \\ 3.093 \\ 1.982 \\ 1.744 \\ 2.855 \\ 1.238 \end{pmatrix}$$

Theorem (F, 2024). All critical points are real and positive. Every critical point is a local maximum of the likelihood function.

proof.

Squaring means real parametric critical points imply positive critical points.

The likelihood function $\ell_u(A) = \dfrac{\prod_{i,j} \det(A_{ij})^{2u_{ij}}}{\left(\sum_{i,j} \det(A_{ij})^2\right)^{\sum_{ij} u_{i,j}}}$ is nonnegative and so

has at least one local maximum in every region, bounded or unbounded, of $\mathbb{R}^{2(n-2)} \setminus \bigcup_{i,j} \{\det(A_{ij}) = 0\}$.

Real and Positive Solutions

Claim. The space $\mathbb{R}^{2(n-2)} \setminus \bigcup_{i,j} \{\det(A_{ij}) = 0\}$ has $2^{n-2}(n-1)!$ connected regions.

The regions are in bijection with the possible sign vectors that can arise from a vector of Plücker coordinates in $\mathrm{Gr}(2,n)$.

1. Choose how many columns have two different signs ($n-1$ choices)

$$A_n = \begin{pmatrix} 1 & 0 & -a_{13} & \cdots & -a_{1k} & a_{1(k+1)} & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2k} & a_{2(k+1)} & \cdots & a_{2n} \end{pmatrix}$$

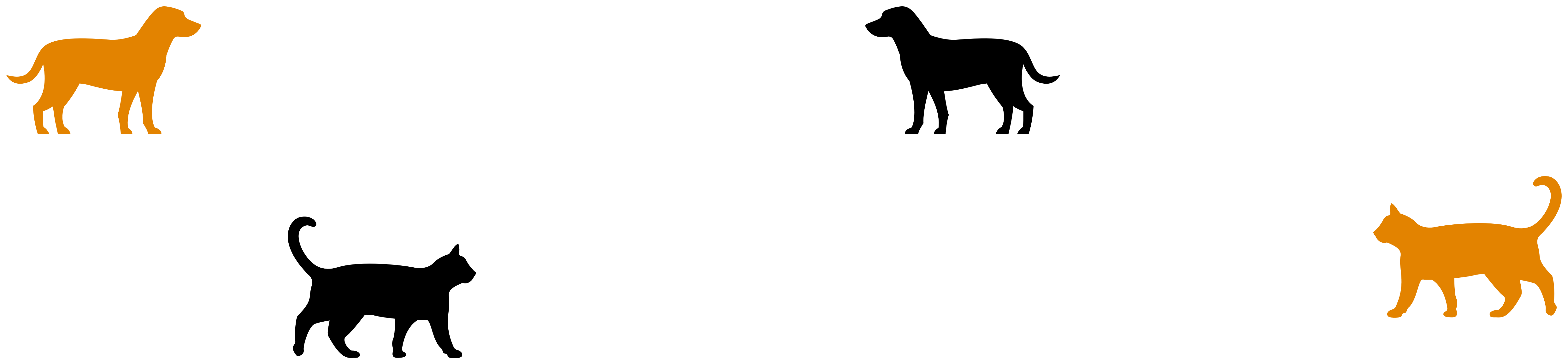
2. Permute the last $n-2$ columns ($(n-2)!$ choices).

3. Flip the signs of any of the last $n-2$ columns (2^{n-2} choices). ■

Recap

- Maximum likelihood estimation over DPPs is hard and there are many extraneous parametric critical points.
- The Grassmannian has two lives as an algebraic variety: one in applied settings and one in algebraic geometry.
- The squared Grassmannian is a model for projection determinantal point processes.
- The squared Grassmannian is one of the first examples of a model on which the likelihood function has the property that all of its critical points are local maxima.

Thank you!



References

- Hannah Friedman, *Likelihood Geometry of the Squared Grassmannian* (2024), `arXiv: 2409.03730.`
- Karel Devriendt, Hannah Friedman, Bernhard Reinke, and Bernd Sturmfels, *The Two Lives of the Grassmannian*, to appear in *Acta Universitatis Sapientiae, Mathematica* (2024).
- Hannah Friedman, Bernd Sturmfels, Maksym Zubkov, *Likelihood Geometry of Determinantal Point Processes*, Algebraic Statistics **15** (2024) 15-25.
- June Huh, *The Maximum Likelihood Degree of a Very Affine Variety*, Composito Mathematica **149** (2013), 1245-1266.
- June Huh and Bernd Sturmfels, *Likelihood Geometry*, Combinatorial Algebraic Geometry (eds. Aldo Conca et al.), Lecture Notes in Mathematics 2108, Springer, (2014) 63-117.
- Paul Breiding and Sascha Timme, *HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia*, Mathematical Software - ICMS 2018, Spring International Publishing (2018), 458-465.