Likelihood Geometry of the Squared Grassmannian

Hannah Friedman (UC Berkeley)
AlgStat 2025 Munich
March 27, 2025

Main Results

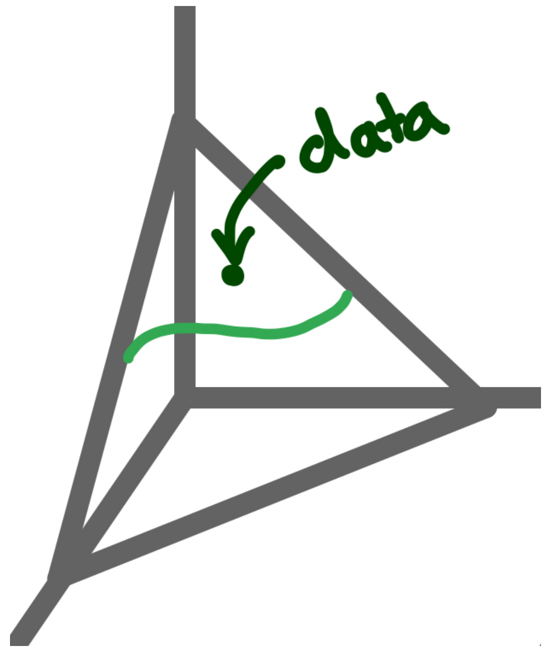
1. The ML degree of the squared Grassmannian $\mathrm{sGr}(2,n)$ is $\dfrac{(n-1)!}{2}$.

2. All critical points of the likelihood function are real and positive. Every critical point is a local maximum.

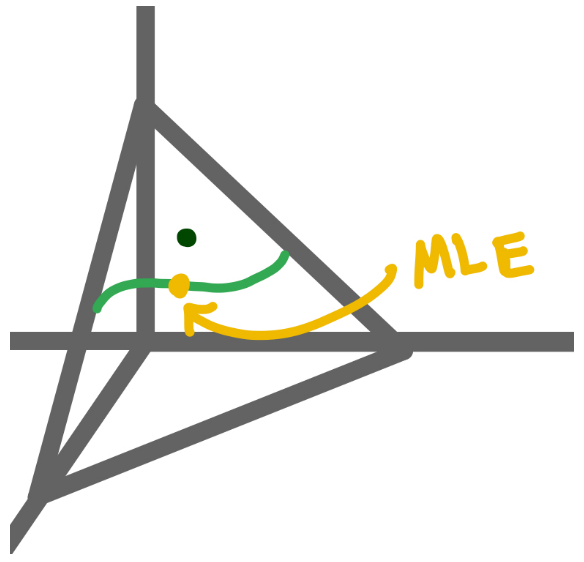
- What is the ML degree of a variety?
- What is the relevant statistical model?
- What is the squared Grassmannian?
- Proof?

Maximum Likelihood Estimation

Given:



Find:



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

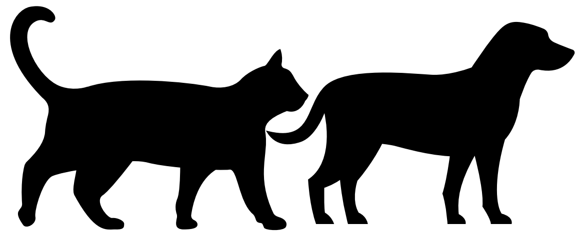
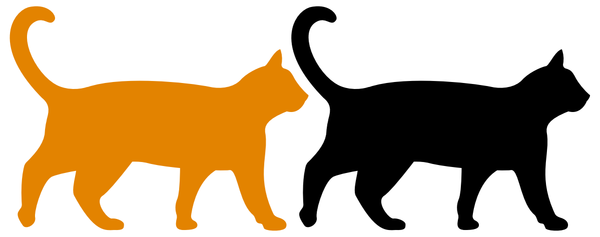
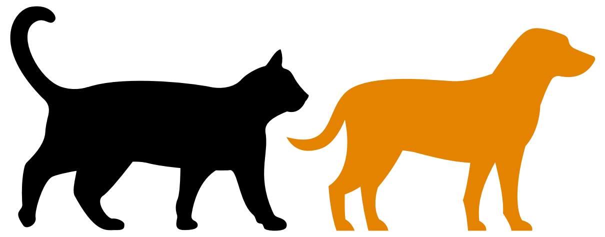
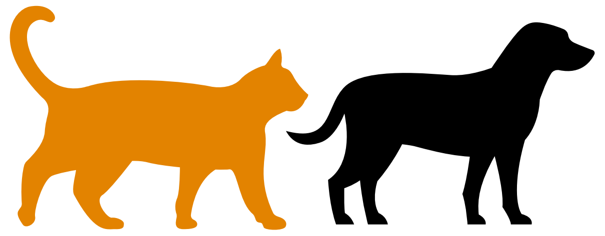
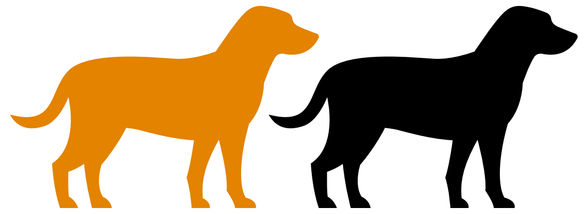
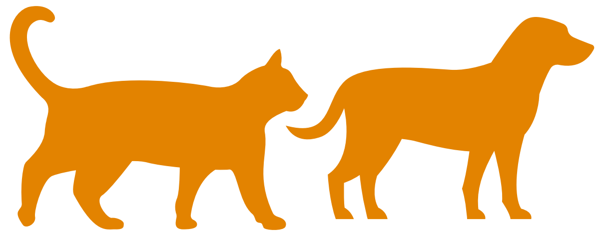
$$\max_{q \in V} L_u(q) = \sum_i u_i \log(q_i) - \left(\sum_i u_i\right) \log\left(\sum_i q_i\right).$$

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**.

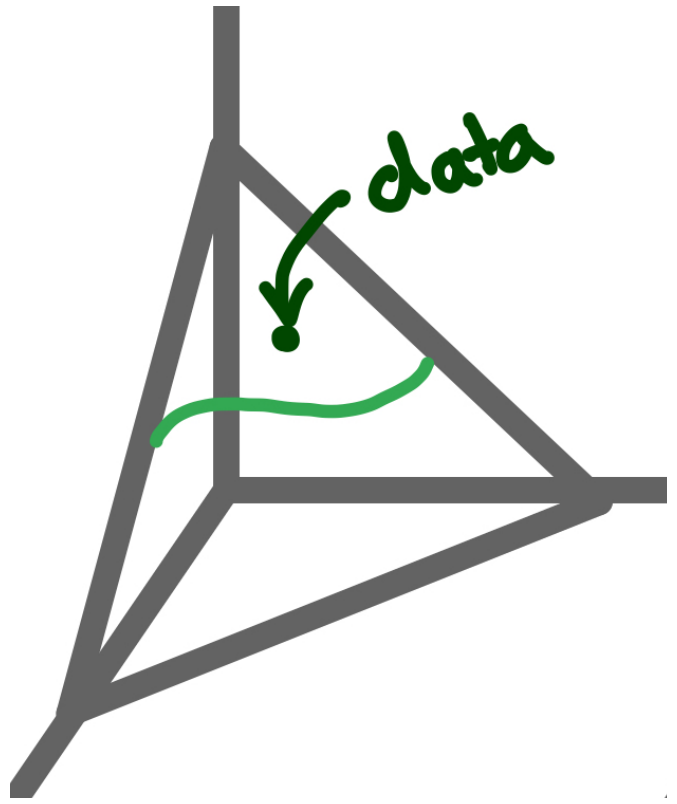
Jackie walks into an animal shelter and adopts 2 of the 4 animals at the shelter every day for 100 days. Every day, she decides which animals to take home by sampling from an unknown probability distribution.

Maximum Likelihood Estimation

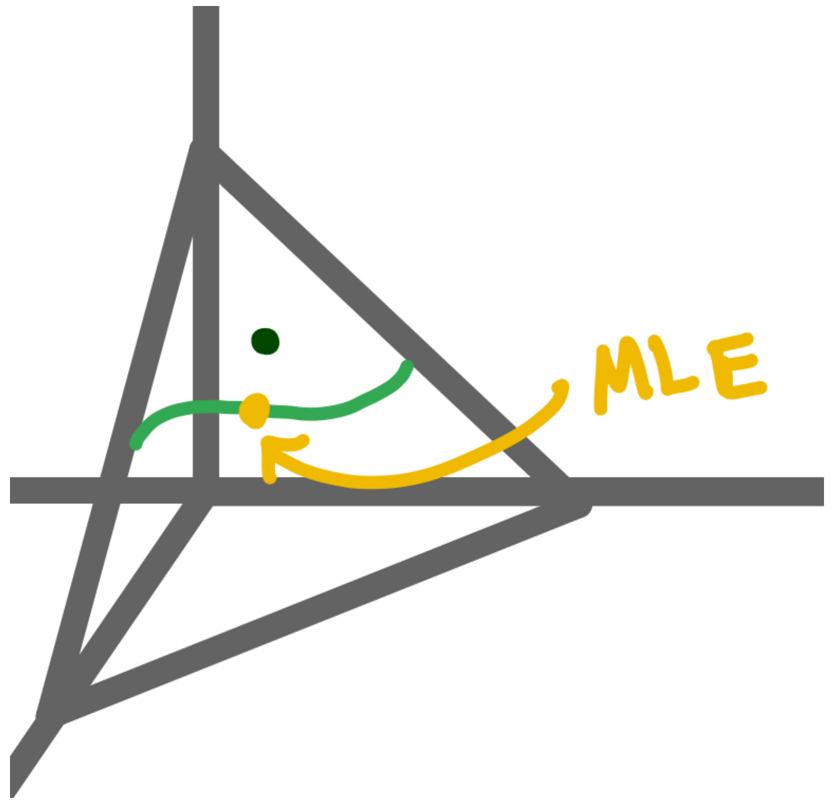
	14
	11
	26
	24
	9
	16

Since Jackie prefers to adopt "diverse" pairs of animals, she samples from a specific type of distribution called a **projection determinantal point process** (projection DPP).

Given:



Find:



Projection Determinantal Point Processes

Example Projection DPPs with state space $\binom{[4]}{2}$ are parameterized by symmetric matrices

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{12} & p_{22} & p_{23} & p_{24} \\ p_{13} & p_{23} & p_{33} & p_{34} \\ p_{14} & p_{24} & p_{34} & p_{44} \end{pmatrix}$$

satisfying

$$P^2 = P$$
$$\mathrm{trace}(P) = 2$$

and the distribution is defined by

$$\mathbb{P}_{ij} = \det(P_{ij}) = p_{ii}p_{jj} - p_{ij}^2 .$$

For projection DPPs with state space $\binom{[n]}{d}$

- P is $n \times n$.
- Probailities are $d \times d$ principal minors.

Two Lives of the Grassmannian

Definition The **Grassmannian** $\mathrm{Gr}(d, n)$ is the set of d-dimensional subspaces of $\mathbb{R}^n$.

Every point in $\mathrm{Gr}(2,n)$ is the row span of some $A \in \mathbb{R}^{2\times n}$, but this representation is not unique.

Orthogonal Projection Matrices

$$P = A^T(AA^T)^{-1}A$$

Plücker Coordinates

$$x = (\det(A_{ij}))_{1\leq i<j\leq n}$$

Lemma (Devriendt-F-Reinke-Sturmfels, 2024)

$$\mathbb{P}_{ij} = \det(P_{ij}) = \frac{x_{ij}^2}{\sum_{1\leq k<\ell\leq n} x_{k\ell}^2}.$$

The Squared Grassmannian

Every 2-dimensional subspace of $\mathbb{R}^n$ determines a projection DPP by

$$\mathbb{P}_{ij} = \det(P_{ij}) = \frac{x_{ij}^2}{\sum_{1\le k<\ell\le n} x_{k\ell}^2} = \frac{\det(A_{ij})^2}{\sum_{1\le k<\ell\le n} (A_{k\ell})^2}. \qquad A = \begin{pmatrix} 1 & 0 & a_{13} & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2n} \end{pmatrix}$$

Definition The **squared Grassmannian** $\mathrm{sGr}(2,n)$ is the image of the Grassmannian $\mathrm{Gr}(2,n) \subset \mathbb{P}^{\binom{n}{2}-1}$ in its Plücker embedding under the map $\begin{array}{c}\mathrm{Gr}(2,n) \to \mathbb{P}^{\binom{n}{2}-1} \\ (x_{ij})_{1\le i<j\le n} \mapsto (x_{ij}^2)_{1\le i<j\le n}\end{array}$

Corollary (Devriendt-F-Reinke-Sturmfels, 2024) The projection determinantal point process is the discrete statistical model on the state space $\binom{[n]}{2}$ whose underlying algebraic variety is the squared Grassmannian $\mathrm{sGr}(2,n)$.

$$L_u(A) = \sum_{i,j} u_{ij} \log(\det(A_{ij})^2) - \left(\sum_{i,j} u_{ij}\right) \log\left(\sum_{i,j} \det(A_{ij})^2\right) \quad \text{vs.} \quad \underset{q \in \mathrm{sGr}(2,n)}{L_u(q)} = \sum_{i,j} u_{ij} \log(q_{ij}) - \left(\sum_{ij} u_{ij}\right) \log\left(\sum_{i,j} q_{ij}\right)$$

Computing the Maximum Likelihood Estimate

Example $A = \begin{pmatrix} 1 & 0 & a_{13} & a_{14} \\ 0 & 1 & a_{23} & a_{24} \end{pmatrix}$ $u = [14,11,26,24,9,16]$

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

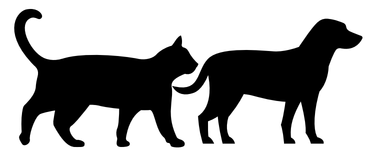
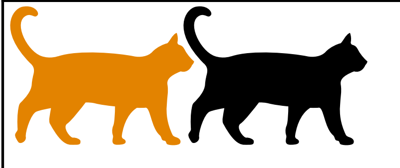
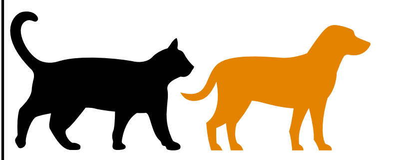
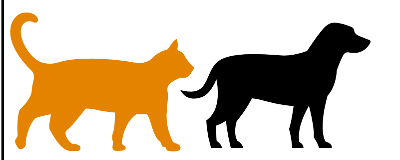
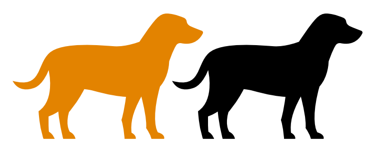
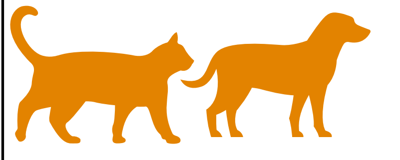
1.

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

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

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

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

	14
	11
	26
	24
	9
	16

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

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

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

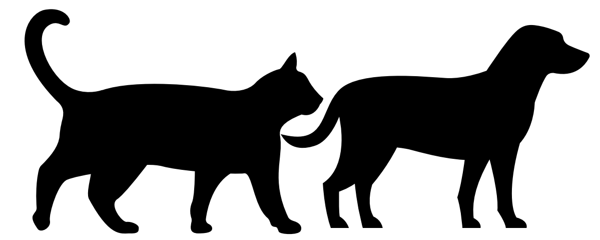
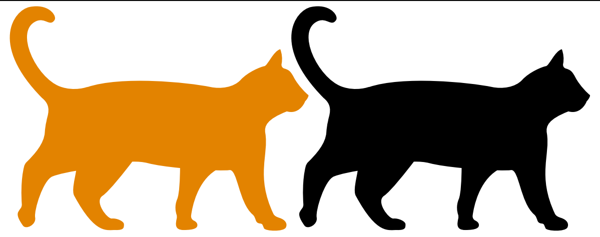
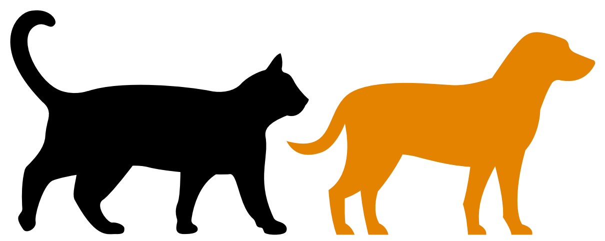
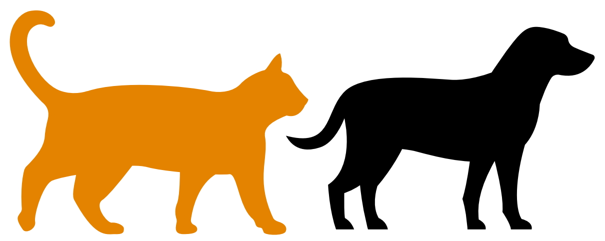
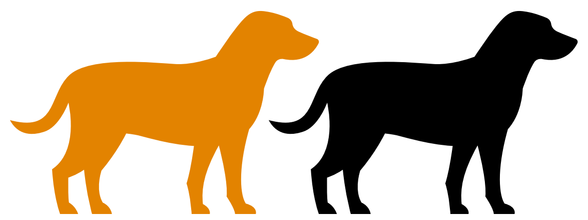
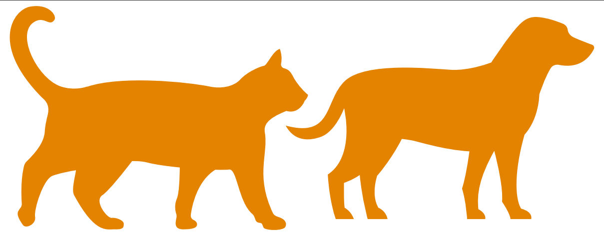
$\begin{pmatrix} 1 & 0 & 0.839 & -0.507 \\ 0 & 1 & 0.584 & 0.888 \end{pmatrix} \times 8$ $\begin{pmatrix} 1 & 0 & 1.320 & 1.690 \\ 0 & 1 & 1.759 & 1.408 \end{pmatrix} \times 8$

24 parametric 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

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

Corollary (F, 2024). The ML degree of the squared Grassmannian $\mathrm{sGr}(2,n)$ is $\dfrac{(n-1)!}{2}$.

proof idea: Apply the following 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 $(-1)^d \chi(X \backslash \mathscr{H})$.

and compute the Euler characteristic inductively using the deletion map

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

Real and Positive Solutions

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

Observe: Squaring means real parametric critical points imply positive critical points.

Proof outline:

All d!
1. Understand the real regions where the parametric log-likelihood function is defined.
2. Prove there is at least one parametric critical point per region.
3. Show that the number of regions equals the number of complex critical points.

1.

$$L_u(A) = \sum_I u_I \log(\det(A_I)^2) - \left(\sum_I u_I\right) \log\left(\sum_I \det(A_I)^2\right)$$

Sum of squares!

The log-likelihood function is defined where $\det(A_I) \neq 0$ and $\sum_I \det(A_I)^2 \neq 0$.

Each of these regions maps onto a region of the real open Grassmannian

$$\mathrm{Gr}_{\mathbb{R}}(d,n)^\circ = \{x \in \mathrm{Gr}_{\mathbb{R}}(d,n) : \prod_{I \in \binom{[n]}{d}} x_I \neq 0\}\,.$$

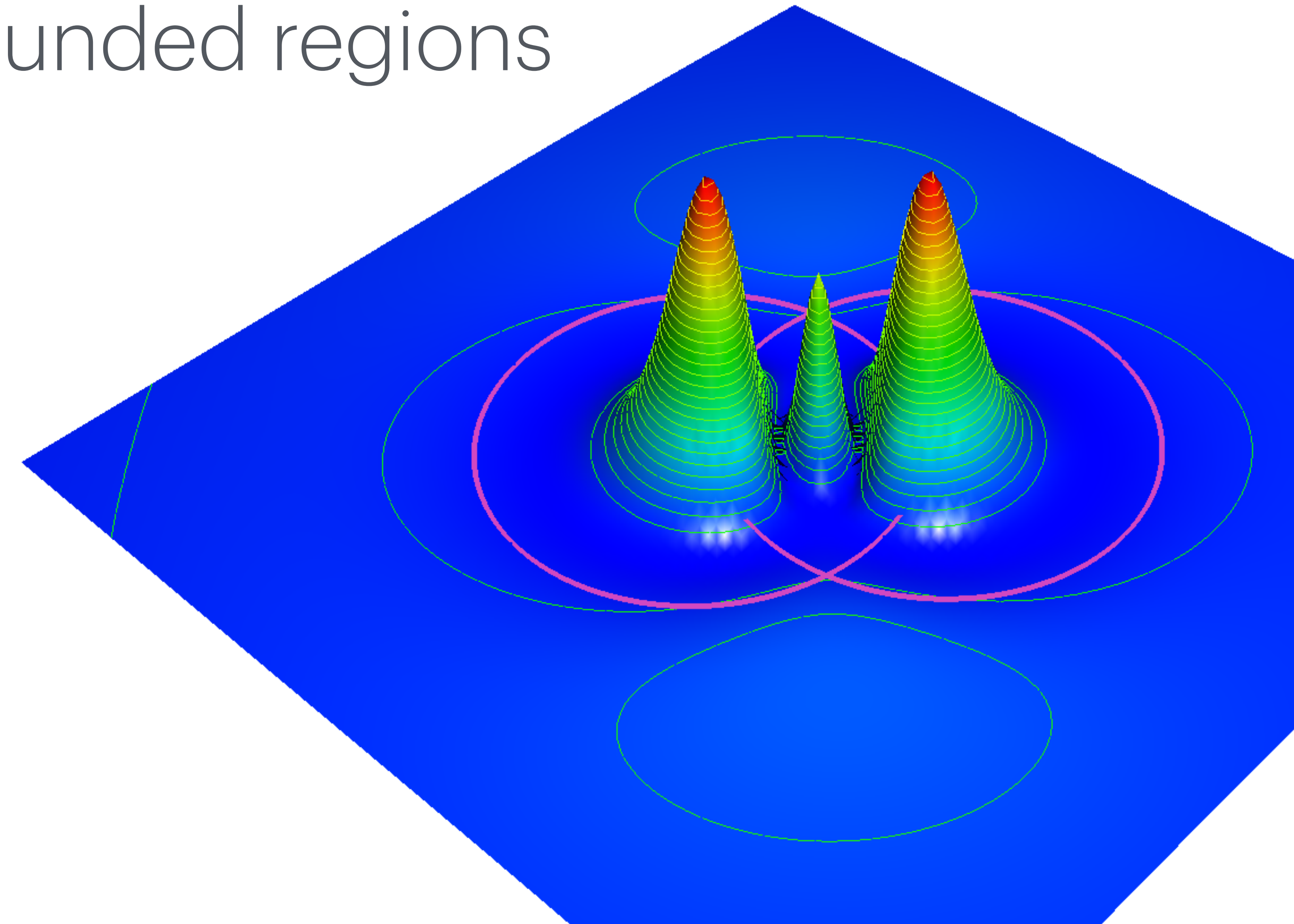
Real and Positive Solutions

2. There is one critical point of the log-likelihood function per region.

The likelihood function $\ell_u(A) = \dfrac{\prod_I \det(A_I)^{2u_I}}{\left(\sum_I \det(A_I)^2\right)^{\sum_I u_I}}$ shares critical points with the log-likelihood function.

$\ell_u(A)$ nonnegative $\Longrightarrow$ $\ell_u(A)$ is positive on every region and zero on the boundaries
$\Longrightarrow$ $\ell_u(A)$ has a local max on every *bounded* region

" $\lim_{A\to\infty} \ell_u(A) = 0$" $\Longrightarrow$ unbounded regions behave like bounded regions
$\Longrightarrow$ $\ell_u(A)$ has a local max on every *unbounded* region



$\#\{$regions of $\mathrm{Gr}_{\mathbb{R}}(d,n)^\circ\} \leq \#\{$parametric local maxima$\}$

Lower bound on the Number of Regions

3. $\#\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(d,n)) \leq \#\{$regions of $\mathrm{Gr}_{\mathbb{R}}(d,n)^\circ\} \leq \#\{$parametric local maxima$\}$ $\leq \#\{$real parametric critical points$\} \leq \#\{$complex parametric critical points$\}$

$$\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(d,n)) = \{\mathrm{sgn}(p_I)_{I \in \binom{[n]}{d}} : p \in \mathrm{Gr}(d,n)^\circ,\ p_{1\dots d} = 1\}$$

$$\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 4 \end{pmatrix} \to (+ \quad + \quad + \quad - \quad - \quad -)$$

Claim. $\#\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(2,n)) = 2^{n-2}(n-1)!$

Fix $a_{13}, \dots, a_{1n}, a_{23}, \dots a_{2n} > 0$.

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).

Real and Positive Solutions

Example

$$\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 4 \end{pmatrix} \to (+ \quad + \quad + \quad - \quad - \quad -)$$

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 3 & 4 \end{pmatrix} \to (+ \quad + \quad + \quad + \quad - \quad -)$$

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & 4 & 3 \end{pmatrix} \to (+ \quad + \quad + \quad - \quad + \quad +)$$

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 4 & -3 \end{pmatrix} \to (+ \quad + \quad - \quad - \quad - \quad -)$$

8 = 32 - 24 vectors not in $\mathbf{sgn}(\mathrm{Gr}_{\mathbb{R}}(2,4))$:

$$\begin{matrix} (+ & - & + & + & + & +) \\ (+ & - & - & - & + & +) \\ (+ & + & + & + & - & +) \\ (+ & + & - & - & - & +) \\ (+ & + & - & + & + & -) \\ (+ & + & + & - & + & -) \\ (+ & - & - & + & - & -) \\ (+ & - & + & - & - & -) \end{matrix}$$

$2^{n-2}(n-1)! = \#\{$regions of $\mathrm{Gr}_{\mathbb{R}}(2,n)^{\circ}\} \leq \#\{$parametric local maxima$\}$

$\leq \#\{$real parametric critical points$\} \leq \#\{$complex parametric critical points$\} = 2^{n-2}(n-1)!$ $\blacksquare$

Real and Positive Solutions

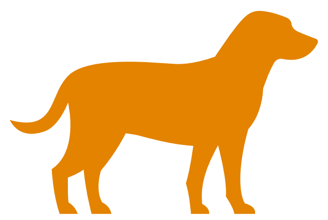
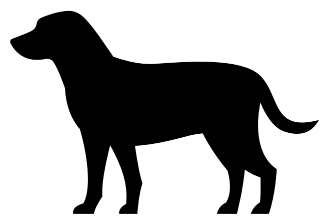
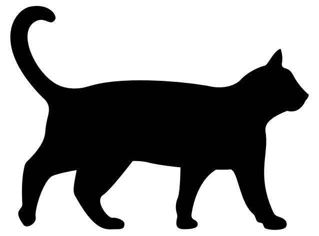
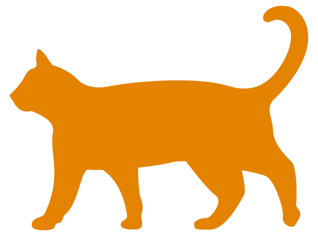
Example When $d = 3$ and $n = 6$, there are are **17664** parametric critical points.
For data vectors with entries sampled uniformly at random from $\{1,\ldots,1000\}$, there are **11904** real critical points, all of which are local maxima.
Numerical computations show that there are precisely **11904** different sign vectors of Plücker coordinates that can arise for points in $\mathrm{Gr}_{\mathbb{R}}(3,6)$.

In general, we have

$$\#\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(d,n)) \leq \#\{\text{regions of } \mathrm{Gr}_{\mathbb{R}}(d,n)^{\circ}\}$$
$$\leq \#\{\text{parametric local maxima}\} \leq \#\{\text{real parametric critical points}\}\,.$$

Conjecture The last two inequalities are equalities.

Thank you!

- Hannah Friedman, *Likelihood Geometry of the Squared Grassmannian*, to appear in *Proceedings of the American Mathematical Society*.
- Karel Devriendt, Hannah Friedman, Bernhard Reinke, and Bernd Sturmfels, *The Two Lives of the Grassmannian*, to appear in *Acta Universitatis Sapientiae, Mathematica*.
- 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.