Maximum likelihood geometry

Jan Draisma
Universität Bern

High school probability

n-sided die, probabilities $P = (p_1, \ldots, p_n)$, thrown N times
$\rightsquigarrow$ prob of $U = (u_1, \ldots, u_n) \in \mathbb{N}^n$ with $u_1 + \ldots + u_n = N$ is

multinomial$\cdot p_1^{u_1} \cdots p_n^{u_n}$

High school probability

n-sided die, probabilities $P = (p_1, \ldots, p_n)$, thrown N times
$\rightsquigarrow$ prob of $U = (u_1, \ldots, u_n) \in \mathbb{N}^n$ with $u_1 + \ldots + u_n = N$ is

multinomial$\cdot p_1^{u_1} \cdots p_n^{u_n}$ $\quad$ $=: \ell_U(P)$ *likelihood of P given U*

High school probability

n-sided die, probabilities $P = (p_1, \ldots, p_n)$, thrown N times
$\rightsquigarrow$ prob of $U = (u_1, \ldots, u_n) \in \mathbb{N}^n$ with $u_1 + \ldots + u_n = N$ is

multinomial$\cdot p_1^{u_1} \cdots p_n^{u_n}$ $=: \ell_U(P)$ *likelihood of P given U*

Basic statistical problem
Given U, *maximise* $\ell_U(P)$ subject to constraints on P.

High school probability

n-sided die, probabilities $P = (p_1, \ldots, p_n)$, thrown N times
$\rightsquigarrow$ prob of $U = (u_1, \ldots, u_n) \in \mathbb{N}^n$ with $u_1 + \ldots + u_n = N$ is

multinomial$\cdot p_1^{u_1} \cdots p_n^{u_n}$ $=: \ell_U(P)$ *likelihood of P given U*

Basic statistical problem
Given U, *maximise* $\ell_U(P)$ subject to constraints on P.

Example
If only constraints are $\sum_i p_i =: p_+ = 1$ and $p_i \geq 0$
$\rightsquigarrow$ maximum attained by $p_i := u_i/N$.

(Mixtures of) independence

$P = (p_{ij})_{ij} \in \mathbb{R}^{m \times n}$ joint distribution of *two* random variables
$\rightsquigarrow$ *independent* if p_{ij} can be written as $q_i t_j$, i.e., iff $\mathrm{rk}(P) = 1$

(Mixtures of) independence

$P = (p_{ij})_{ij} \in \mathbb{R}^{m \times n}$ joint distribution of *two* random variables
$\rightsquigarrow$ *independent* if p_{ij} can be written as $q_i t_j$, i.e., iff $\mathrm{rk}(P) = 1$

$U = (u_{ij})_{ij}$ data matrix, $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

(Mixtures of) independence

$P = (p_{ij})_{ij} \in \mathbb{R}^{m \times n}$ joint distribution of *two* random variables
$\rightsquigarrow$ *independent* if p_{ij} can be written as $q_i t_j$, i.e., iff $\mathrm{rk}(P) = 1$

$U = (u_{ij})_{ij}$ data matrix, $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

ML-problem for independence model
Maximise $\ell_U(P)$ subject to $p_{ij} \geq 0, p_{++} = 1, \mathrm{rk}(P) = 1.$

(Mixtures of) independence

$P = (p_{ij})_{ij} \in \mathbb{R}^{m \times n}$ joint distribution of *two* random variables
$\rightsquigarrow$ *independent* if p_{ij} can be written as $q_i t_j$, i.e., iff $\mathrm{rk}(P) = 1$

$U = (u_{ij})_{ij}$ data matrix, $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

ML-problem for independence model
Maximise $\ell_U(P)$ subject to $p_{ij} \geq 0, p_{++} = 1, \mathrm{rk}(P) = 1$.

Solution
$p_{ij} = u_{i+}u_{+j}/(u_{++}^2)$

(Mixtures of) independence

$P = (p_{ij})_{ij} \in \mathbb{R}^{m \times n}$ joint distribution of *two* random variables
$\rightsquigarrow$ *independent* if p_{ij} can be written as $q_i t_j$, i.e., iff $\mathrm{rk}(P) = 1$

$U = (u_{ij})_{ij}$ data matrix, $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

ML-problem for independence model
Maximise $\ell_U(P)$ subject to $p_{ij} \geq 0, p_{++} = 1, \mathrm{rk}(P) = 1$.

Solution
$p_{ij} = u_{i+}u_{+j}/(u_{++}^2)$

Mixture of r copies of independence
P convex combination of $P_1, \ldots, P_r$ as above
$\rightsquigarrow p_{++} = 1$ and $\mathrm{rk}(P) \leq r$

ML-problem much harder!

Critical points

ML-problem for manifold $M \subseteq (\mathbb{R}_{>0})^n$
Maximise $\ell_U(P) = \prod_i p_i^{u_i}$ subject to $P \in M$.

Critical points

ML-problem for manifold $M \subseteq (\mathbb{R}_{>0})^n$
Maximise $\ell_U(P) = \prod_i p_i^{u_i}$ subject to $P \in M$.

Derivative
$(d_P \ell_U)(X) = \ell_U(P) \sum_i \frac{x_i}{p_i} u_i$, $X \in T_P M$

Critical points

ML-problem for manifold $M \subseteq (\mathbb{R}_{>0})^n$
Maximise $\ell_U(P) = \prod_i p_i^{u_i}$ subject to $P \in M$.

Derivative
$(d_P \ell_U)(X) = \ell_U(P) \sum_i \frac{x_i}{p_i} u_i$, $X \in T_P M$

Necessary condition for P **to be the ML-estimate**
P *critical*: $d_P \ell_U$ vanishes identically on $T_P M$
$\Leftrightarrow \sum_i \frac{x_i}{p_i} u_i = 0$ for all $X \in T_P M \Leftrightarrow (p_1^{-1}, \ldots, p_n^{-1}) T_P M \subseteq U^\perp$

Critical points

ML-problem for manifold $M \subseteq (\mathbb{R}_{>0})^n$
Maximise $\ell_U(P) = \prod_i p_i^{u_i}$ subject to $P \in M$.

Derivative
$(d_P \ell_U)(X) = \ell_U(P) \sum_i \frac{x_i}{p_i} u_i,\ X \in T_P M$

Necessary condition for P **to be the ML-estimate**
P *critical*: $d_P \ell_U$ vanishes identically on $T_P M$
$\Leftrightarrow \sum_i \frac{x_i}{p_i} u_i = 0$ for all $X \in T_P M \Leftrightarrow (p_1^{-1}, \ldots, p_n^{-1}) T_P M \subseteq U^\perp$

Algebraic measure of complexity
Count number of critical points. . . easier over $\mathbb{C}$!

ML-degree

Setting

$M \subseteq (\mathbb{C}^*)^n$ smooth subvariety (locally closed) $\rightsquigarrow$ Crit(M):= $\{(P, U) \in M \times \mathbb{C}^n \mid P^{-1}T_P M \subseteq U^\perp\}$ *variety of critical points*

ML-degree

Setting

$M \subseteq (\mathbb{C}^*)^n$ smooth subvariety (locally closed) $\rightsquigarrow$ Crit(M):= $\{(P, U) \in M \times \mathbb{C}^n \mid P^{-1}T_P M \subseteq U^\perp\}$ *variety of critical points*

Fibre over $U \in \mathbb{N}^n$ is set of critical points for U.
Reasonable assumptions $\rightsquigarrow$ fibres finite of constant size for sufficiently general U.

ML-degree

Setting
$M \subseteq (\mathbb{C}^*)^n$ smooth subvariety (locally closed) $\rightsquigarrow$ Crit(M):=
$\{(P, U) \in M \times \mathbb{C}^n \mid P^{-1}T_P M \subseteq U^\perp\}$ *variety of critical points*

Fibre over $U \in \mathbb{N}^n$ is set of critical points for U.
Reasonable assumptions $\rightsquigarrow$ fibres finite of $\underbrace{\text{constant size}}_{\textit{ML-degree} \text{ of } M}$ for sufficiently general U.

ML-degree

Setting
$M \subsetneq (\mathbb{C}^*)^n$ smooth subvariety (locally closed) $\rightsquigarrow$ Crit(M):= $\{(P,U) \in M \times \mathbb{C}^n \mid P^{-1}T_P M \subseteq U^\perp\}$ *variety of critical points*

Fibre over $U \in \mathbb{N}^n$ is set of critical points for U.
Reasonable assumptions $\rightsquigarrow$ fibres finite of $\underbrace{\text{constant size}}_{\text{ML-degree of } M}$ for sufficiently general U.

Theorem (Huh, 2012; related Franecki-Kapranov, 2000)
M closed in addition to smooth (*very affine*)
$\rightsquigarrow$ ML-degree is $(-1)^{\dim_{\mathbb{C}} M}\chi(M)$, where χ is the Euler characteristic of M.

ML-degrees for matrices of given rank

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

ML-degrees for matrices of given rank

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \le n$

Theorem (Hauenstein-Rodriguez-Sturmfels, 2012)
For small $r \le m \le n$ ML-degree of M is as follows:

		(m,n)						
		$(3,3)$	$(3,4)$	$(3,5)$	$(4,4)$	$(4,5)$	$(4,6)$	$(5,5)$
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

ML-degrees for matrices of given rank

$M_r := \{P \in (\mathbb{C}^*)^{m \times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

Theorem (Hauenstein-Rodriguez-Sturmfels, 2012)
For small $r \leq m \leq n$ ML-degree of M is as follows:

		(m,n)						
		(3,3)	(3,4)	(3,5)	(4,4)	(4,5)	(4,6)	(5,5)
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

Bertini

ML-degrees for matrices of given rank

$M_r := \{P \in (\mathbb{C}^*)^{m \times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

Theorem (Hauenstein-Rodriguez-Sturmfels, 2012)
For small $r \leq m \leq n$ ML-degree of M is as follows:

		(m,n)						
		(3,3)	(3,4)	(3,5)	(4,4)	(4,5)	(4,6)	(5,5)
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

Bertini

Conjecture (HRS)
ML-degree (M_r) = ML-degree (M_{m-r+1})

ML-duality

$$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$$

ML-duality

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

Theorem (D-Rodriguez, 2012)
$U \in \mathbb{N}^{m\times n}$ sufficiently general $\rightsquigarrow$ the map $P \mapsto Q'$ defined by $p_{ij}q'_{ij} = u_{i+}u_{ij}u_{+j}/(u_{++}^3)$ is a bijection between critical points of ℓ_U on M_r and those on M_{m-r+1}.

ML-duality

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

Theorem (D-Rodriguez, 2012)
$U \in \mathbb{N}^{m\times n}$ sufficiently general $\rightsquigarrow$ the map $P \mapsto Q'$ defined by $p_{ij}q'_{ij} = u_{i+}u_{ij}u_{+j}/(u_{++}^3)$ is a bijection between critical points of ℓ_U on M_r and those on M_{m-r+1}.

Remark

- $\ell_U(P)\ell_U(Q')$ independent of P
- P positive real $\Leftrightarrow$ so is Q'
- $\ell_U(Q')$ decreases with increasing $\ell_U(P)$

ML-duality

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$

Theorem (D-Rodriguez, 2012)
$U \in \mathbb{N}^{m\times n}$ sufficiently general $\rightsquigarrow$ the map $P \mapsto Q'$ defined by $p_{ij}q'_{ij} = u_{i+}u_{ij}u_{+j}/(u^3_{++})$ is a bijection between critical points of ℓ_U on M_r and those on M_{m-r+1}.

Remark

- $\ell_U(P)\ell_U(Q')$ independent of P
- P positive real $\Leftrightarrow$ so is Q'
- $\ell_U(Q')$ decreases with increasing $\ell_U(P)$
- Rodriguez established a general ML-duality theory.

"M_r and M_{m-r+1} are ML-dual"

Rank-2 case

		(m,n)						
		$(3,3)$	$(3,4)$	$(3,5)$	$(4,4)$	$(4,5)$	$(4,6)$	$(5,5)$
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

Rank-2 case

		(m,n)						
		$(3,3)$	$(3,4)$	$(3,5)$	$(4,4)$	$(4,5)$	$(4,6)$	$(5,5)$
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

Theorem (Botong-Rodriguez 2015)
For $r = 2$ and $m = 3 \leq n$ the ML-degree is equal to $2^{n+1} - 6$.

Rank-2 case

		(m,n)						
		(3,3)	(3,4)	(3,5)	(4,4)	(4,5)	(4,6)	(5,5)
	1	1	1	1	1	1	1	1
	2	10	26	58	191	843	3119	6776
r	3	1	1	1	191	843	3119	61326
	4				1	1	1	6776
	5							1

Theorem (Botong-Rodriguez 2015)
For $r = 2$ and $m = 3 \leq n$ the ML-degree is equal to $2^{n+1} - 6$.

The proof involves subtle topological counting. More generally, they prove a recurrence relation for the rank-two case with which a closed formula for any fixed m and running n can be found.

Algebraic varieties with ML-degree 1

Recall

For the model of independence, the ML-degree is 1 and the ML estimate is $(u_{i+}u_{+j})/(u_{++}^2)$.

Algebraic varieties with ML-degree 1

Recall

For the model of independence, the ML-degree is 1 and the ML estimate is $(u_{i+}u_{+j})/(u_{++}^2)$.

In general, consider a rational map $\Psi : \mathbb{C}^n \to (\mathbb{C}^*)^n$ which is a composition of a linear map $\mathbb{C}^n \to \mathbb{C}^r$ followed by a monomial map $\mathbb{C}^r \to (\mathbb{C}^*)^n$ which is homogeneous of degree zero. Define M as the smooth locus of the closure of $\mathrm{im}\Psi$.

Algebraic varieties with ML-degree 1

Recall
For the model of independence, the ML-degree is 1 and the ML estimate is $(u_{i+}u_{+j})/(u_{++}^2)$.

In general, consider a rational map $\Psi : \mathbb{C}^n \to (\mathbb{C}^*)^n$ which is a composition of a linear map $\mathbb{C}^n \to \mathbb{C}^r$ followed by a monomial map $\mathbb{C}^r \to (\mathbb{C}^*)^n$ which is homogeneous of degree zero. Define M as the smooth locus of the closure of $\mathrm{im}\Psi$.

Theorem (Huh 2013)
The variety M has ML-degree 1, and Ψ is its ML-estimator. Moreover, every variety of ML-degree 1 arises like this.

Reintroducing inequalities

Forget about ML-degree.

$M_1 := \{P \in \mathbb{R}^{m\times n}_{\geq 0} \mid \mathrm{rk}(P) = 1, \sum_{ij} p_{ij} = 1\}$ independence

$M_r := \{c_1 P_1 + \ldots + c_r P_r \mid P_i \in M_1, c_i \in \mathbb{R}_{\geq 0}, c_1 + \ldots + c_r = 1\}$
mixture of r copies of M_1

Reintroducing inequalities

Forget about ML-degree.

$M_1 := \{P \in \mathbb{R}^{m\times n}_{\geq 0} \mid \mathrm{rk}(P) = 1, \sum_{ij} p_{ij} = 1\}$ independence

$M_r := \{c_1 P_1 + \ldots + c_r P_r \mid P_i \in M_1, c_i \in \mathbb{R}_{\geq 0}, c_1 + \ldots + c_r = 1\}$ mixture of r copies of M_1

For a nonnegative matrix A with $\sum_{ij} a_{ij} = 1$ the smallest r with $A \in M_r$ is the *nonnegative rank* of A. It is $\geq \mathrm{rk}(A)$.

Reintroducing inequalities

Forget about ML-degree.

$M_1 := \{P \in \mathbb{R}^{m \times n}_{\geq 0} \mid \mathrm{rk}(P) = 1, \sum_{ij} p_{ij} = 1\}$ independence

$M_r := \{c_1 P_1 + \ldots + c_r P_r \mid P_i \in M_1, c_i \in \mathbb{R}_{\geq 0}, c_1 + \ldots + c_r = 1\}$
mixture of r copies of M_1

For a nonnegative matrix A with $\sum_{ij} a_{ij} = 1$ the smallest r with $A \in M_r$ is the *nonnegative rank* of A. It is $\geq \mathrm{rk}(A)$.

Exercise
If a nonnegative matrix A has rank ≤ 2, then $\mathrm{rk}A$ equals its nonnegative rank.

Reintroducing inequalities

Forget about ML-degree.

$M_1 := \{P \in \mathbb{R}^{m\times n}_{\geq 0} \mid \mathrm{rk}(P) = 1, \sum_{ij} p_{ij} = 1\}$ independence

$M_r := \{c_1P_1 + \ldots + c_rP_r \mid P_i \in M_1, c_i \in \mathbb{R}_{\geq 0}, c_1 + \ldots + c_r = 1\}$
mixture of r copies of M_1

For a nonnegative matrix A with $\sum_{ij} a_{ij} = 1$ the smallest r with $A \in M_r$ is the *nonnegative rank* of A. It is $\geq \mathrm{rk}(A)$.

Exercise
If a nonnegative matrix A has rank ≤ 2, then $\mathrm{rk}A$ equals its nonnegative rank.

Theorem (Vavasis 2009, Shitov 2015/2016)
nonnegative rank is NP-hard and may depend on the field.

Nonnegative rank three

Consider M_3. Given a data matrix $U \in \mathbb{Z}_{\geq 0}^{m \times n}$, the ML-estimate in M_3 is in practice approximated via the ME-algorithm.

Nonnegative rank three

Consider M_3. Given a data matrix $U \in \mathbb{Z}_{\geq 0}^{m \times n}$, the ML-estimate in M_3 is in practice approximated via the ME-algorithm.

Kubjas-Robeva-Sturmfels 2013
This often converges to the *boundary* of M_3 inside the variety of matrices of rank ≤ 3; find explicit description of fixpoints.

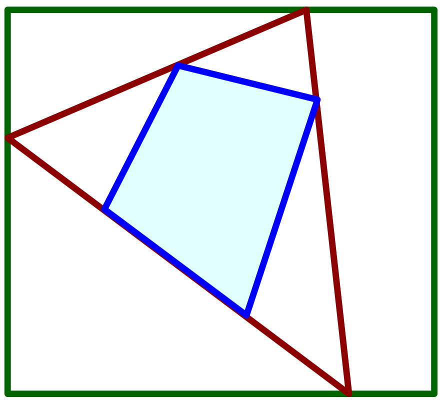
Nonnegative rank three

Consider M_3. Given a data matrix $U \in \mathbb{Z}_{\geq 0}^{m \times n}$, the ML-estimate in M_3 is in practice approximated via the ME-algorithm.

Kubjas-Robeva-Sturmfels 2013
This often converges to the *boundary* of M_3 inside the variety of matrices of rank ≤ 3; find explicit description of fixpoints.

Theorem (K-R-S)
This boundary has three orbits of irreducible components under row and column permutations:

- one orbit where an entry of U is zero
- one orbit corresponding to the picture:
- and its transpose



Margins

$M_r := \{P \in (\mathbb{C}^*)^{m \times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

Tangent space
$T_P M_r = \{X \in \mathbb{C}^{m \times n} \mid x_{++} = 0, X \ker P \subseteq \mathrm{im} P\}$

Margins

$M_r := \{P \in (\mathbb{C}^*)^{m \times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

Tangent space
$T_P M_r = \{X \in \mathbb{C}^{m \times n} \mid x_{++} = 0, X \ker P \subseteq \mathrm{im} P\}$

$\mathbf{1} := (1, \ldots, 1) \in \mathbb{C}^m$ or $\mathbb{C}^n$

Margins

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

Tangent space
$T_P M_r = \{X \in \mathbb{C}^{m\times n} \mid x_{++} = 0, X \ker P \subseteq \mathrm{im} P\}$

$\mathbf{1} := (1, \dots, 1) \in \mathbb{C}^m$ or $\mathbb{C}^n$

Lemma
$P\mathbf{1}$ is proportional to $U\mathbf{1}$

Margins

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

Tangent space
$T_P M_r = \{X \in \mathbb{C}^{m\times n} \mid x_{++} = 0, X \ker P \subseteq \mathrm{im} P\}$

$\mathbf{1} := (1, \ldots, 1) \in \mathbb{C}^m$ or $\mathbb{C}^n$

Lemma
$P\mathbf{1}$ is proportional to $U\mathbf{1}$

$$X = \begin{bmatrix} p_{2+} \cdot p_{11} & \cdots & p_{2+} \cdot p_{1n} \\ -p_{1+} \cdot p_{21} & \cdots & -p_{1+} \cdot p_{2n} \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in T_P M_r \rightsquigarrow 0 = \sum_{ij} \frac{x_{ij}}{p_{ij}} u_{ij}$$

□

Margins

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \le n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$

Tangent space
$T_P M_r = \{X \in \mathbb{C}^{m\times n} \mid x_{++} = 0, X \ker P \subseteq \mathrm{im} P\}$

$\mathbf{1} := (1, \ldots, 1) \in \mathbb{C}^m$ or $\mathbb{C}^n$

Lemma
$P\mathbf{1}$ is proportional to $U\mathbf{1}$

$$X = \begin{bmatrix} p_{2+} \cdot p_{11} & \cdots & p_{2+} \cdot p_{1n} \\ -p_{1+} \cdot p_{21} & \cdots & -p_{1+} \cdot p_{2n} \\ 0 & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in T_P M_r \rightsquigarrow 0 = \sum_{ij} \frac{x_{ij}}{p_{ij}} u_{ij} = p_{2+} \cdot u_{1+} - p_{1+} \cdot u_{2+}$$

□

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m \times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Dual critical point
$q_{ij} := u_{i+} \frac{u_{ij}}{p_{ij}} u_{+j}, q'_{ij} := q_{ij}/(u_{++}^3)$

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Dual critical point
$q_{ij} := u_{i+} \frac{u_{ij}}{p_{ij}} u_{+j}, q'_{ij} := q_{ij}/(u_{++}^3)$

Proposition
$q_{++} = u_{++}^3$

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Dual critical point
$q_{ij} := u_{i+} \frac{u_{ij}}{p_{ij}} u_{+j}, q'_{ij} := q_{ij}/(u_{++}^3)$

Proposition
$q_{++} = u_{++}^3$

$Y := (u_{i+} \cdot u_{+j})_{ij}$ satisfies $Y \ker P \subseteq \mathrm{im} P$
P satisfies $P \ker P \subseteq \mathrm{im} P$

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \le n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Dual critical point
$q_{ij} := u_{i+} \frac{u_{ij}}{p_{ij}} u_{+j}, q'_{ij} := q_{ij}/(u_{++}^3)$

Proposition
$q_{++} = u_{++}^3$

$$\left.\begin{array}{l} Y := (u_{i+} \cdot u_{+j})_{ij} \text{ satisfies } Y \ker P \subseteq \mathrm{im} P \\ P \text{ satisfies } P \ker P \subseteq \mathrm{im} P \end{array}\right] \begin{array}{l} \rightsquigarrow Y - cP \in T_P M_r \\ c = y_{++}/p_{++} = u_{++}^2 \end{array}$$

Normalisation of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
U sufficiently general, P critical point for $\ell_U(P) = \prod_{ij} p_{ij}^{u_{ij}}$
$P\mathbf{1}, U\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U$)

Dual critical point
$q_{ij} := u_{i+} \frac{u_{ij}}{p_{ij}} u_{+j}, q'_{ij} := q_{ij}/(u_{++}^3)$

Proposition
$q_{++} = u_{++}^3$

$Y := (u_{i+} \cdot u_{+j})_{ij}$ satisfies $Y \ker P \subseteq \mathrm{im} P$
P satisfies $P \ker P \subseteq \mathrm{im} P$
$\rightsquigarrow Y - cP \in T_P M_r$
$c = y_{++}/p_{++} = u_{++}^2$

$$\sum_{ij} q_{ij} = \sum_{ij} y_{ij} \frac{u_{ij}}{p_{ij}} = \sum_{ij} c p_{ij} \frac{u_{ij}}{p_{ij}} = (u_{++})^3$$

□

Rank of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
$Q = R \cdot (\frac{U}{P}) \cdot K, R = \mathrm{diag}(U\mathbf{1}), K = \mathrm{diag}(\mathbf{1}^T U)$

Rank of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
$Q = R \cdot (\frac{U}{P}) \cdot K, R = \mathrm{diag}(U\mathbf{1}), K = \mathrm{diag}(\mathbf{1}^T U)$

Lemma
$T_P M_r$ is spanned by rank-one matrices vw^T with
$(v \in \mathrm{im} P$ or $w \perp \ker P)$ *and* $(v \perp \mathbf{1}$ or $w \perp \mathbf{1})$.

Rank of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
$Q = R \cdot (\frac{U}{P}) \cdot K, R = \mathrm{diag}(U\mathbf{1}), K = \mathrm{diag}(\mathbf{1}^T U)$

Lemma
$T_P M_r$ is spanned by rank-one matrices vw^T with
($v \in \mathrm{im}P$ or $w \perp \ker P$) *and* ($v \perp \mathbf{1}$ or $w \perp \mathbf{1}$).

Derivative
$(d_P \ell_U)(vw^T) = \sum_{ij} v_i \frac{u_{ij}}{p_{ij}} w_j = v^t R^{-1} Q K^{-1} w$

Rank of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \leq n$
$Q = R \cdot (\frac{U}{P}) \cdot K, R = \mathrm{diag}(U\mathbf{1}), K = \mathrm{diag}(\mathbf{1}^T U)$

Lemma
$T_P M_r$ is spanned by rank-one matrices vw^T with
$(v \in \mathrm{im}P$ or $w \perp \ker P)$ *and* $(v \perp \mathbf{1}$ or $w \perp \mathbf{1})$.

Derivative
$(d_P \ell_U)(vw^T) = \sum_{ij} v_i \frac{u_{ij}}{p_{ij}} w_j = v^t R^{-1} Q K^{-1} w$

Proposition
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^{\perp}$ and $\mathrm{im}Q \subseteq (R^{-1}(\mathrm{im}P \cap \mathbf{1}^{\perp}))^{\perp}$

Rank of Q

$M_r := \{P \in (\mathbb{C}^*)^{m\times n} \mid p_{++} = 1, \mathrm{rk}(P) = r\}, m \le n$
$Q = R \cdot (\frac{U}{P}) \cdot K, R = \mathrm{diag}(U\mathbf{1}), K = \mathrm{diag}(\mathbf{1}^T U)$

Lemma
$T_P M_r$ is spanned by rank-one matrices vw^T with
($v \in \mathrm{im}P$ or $w \perp \ker P$) *and* ($v \perp \mathbf{1}$ or $w \perp \mathbf{1}$).

Derivative
$(d_P \ell_U)(vw^T) = \sum_{ij} v_i \frac{u_{ij}}{p_{ij}} w_j = v^t R^{-1} Q K^{-1} w$

Proposition
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^\perp$ and $\mathrm{im}Q \subseteq (R^{-1}(\mathrm{im}P \cap \mathbf{1}^\perp))^\perp$

$\rightsquigarrow \mathrm{rk}Q =: s \le m - r + 1$

Criticality of Q

$P\mathbf{1}, U\mathbf{1} = R\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U = \mathbf{1}^T K$)
$Q = R \cdot (\frac{U}{P}) \cdot K$, $\mathrm{rk} Q = s \leq m - r + 1$,
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^{\perp} =: W$
$\mathrm{im} Q \subseteq (R^{-1}(\mathrm{im} P \cap \mathbf{1}^{\perp}))^{\perp} =: V$

Criticality of Q

$P\mathbf{1}, U\mathbf{1} = R\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U = \mathbf{1}^T K$)
$Q = R \cdot (\frac{U}{P}) \cdot K$, $\mathrm{rk} Q = s \leq m - r + 1$,
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^{\perp} =: W$
$\mathrm{im} Q \subseteq (R^{-1}(\mathrm{im} P \cap \mathbf{1}^{\perp}))^{\perp} =: V$

Proposition
For all $x \in \mathbb{C}^m, y \in \mathbb{C}^n$ with ($x \in V$ or $y \perp W$) *and* ($x \perp \mathbf{1}$ or $y \perp \mathbf{1}$)
we have $x^T R^{-1} P K^{-1} y = 0$.

Criticality of Q

$P\mathbf{1}, U\mathbf{1} = R\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U = \mathbf{1}^T K$)
$Q = R \cdot (\frac{U}{P}) \cdot K$, $\mathrm{rk} Q = s \leq m - r + 1$,
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^\perp =: W$
$\mathrm{im} Q \subseteq (R^{-1}(\mathrm{im} P \cap \mathbf{1}^\perp))^\perp =: V$

Proposition
For all $x \in \mathbb{C}^m, y \in \mathbb{C}^n$ with ($x \in V$ or $y \perp W$) *and* ($x \perp \mathbf{1}$ or $y \perp \mathbf{1}$) we have $x^T R^{-1} P K^{-1} y = 0$.

E.g., $x \in V$ and $x \perp \mathbf{1} \rightsquigarrow PK^{-1}y = cR\mathbf{1} + v$ with $v \in \mathrm{im} P \cap \mathbf{1}^\perp$
$\rightsquigarrow x^T R^{-1} P K^{-1} y = c x^T \mathbf{1} + x^T R^{-1} v = 0 + 0 = 0$ $\square$

Criticality of Q

$P\mathbf{1}, U\mathbf{1} = R\mathbf{1}$ proportional (and so are $\mathbf{1}^T P, \mathbf{1}^T U = \mathbf{1}^T K$)
$Q = R \cdot (\frac{U}{P}) \cdot K$, $\mathrm{rk} Q = s \leq m - r + 1$,
$\ker Q \supseteq K^{-1}(\ker P + \mathbb{C}\mathbf{1})^{\perp} =: W$
$\mathrm{im} Q \subseteq (R^{-1}(\mathrm{im} P \cap \mathbf{1}^{\perp}))^{\perp} =: V$

Proposition
For all $x \in \mathbb{C}^m, y \in \mathbb{C}^n$ with ($x \in V$ or $y \perp W$) *and* ($x \perp \mathbf{1}$ or $y \perp \mathbf{1}$) we have $x^T R^{-1} P K^{-1} y = 0$.

E.g., $x \in V$ and $x \perp \mathbf{1} \rightsquigarrow PK^{-1}y = cR\mathbf{1} + v$ with $v \in \mathrm{im} P \cap \mathbf{1}^{\perp}$
$\rightsquigarrow x^T R^{-1} P K^{-1} y = c x^T \mathbf{1} + x^T R^{-1} v = 0 + 0 = 0$ $\square$

In particular for $x \in \mathrm{im} Q$ or $y \perp \ker Q \rightsquigarrow Q'$ critical in M_s!

Bijection

$$\mathrm{Crit}(M_r) \overset{\psi_r}{\dashrightarrow} \mathrm{Crit}(M_{f(r)}),\ f(r) = s \leq m - r + 1$$

$$(P, U) \dashrightarrow ((R \cdot \tfrac{U}{P} \cdot K)/(u^3_{++}), U)$$

Bijection

$$\mathrm{Crit}(M_r) \overset{\psi_r}{\dashrightarrow} \mathrm{Crit}(M_{f(r)}),\ f(r) = s \leq m - r + 1$$

$$(P, U) \dashrightarrow ((R \cdot \tfrac{U}{P} \cdot K)/(u^3_{++}), U)$$

Observation

- ψ_r injective
- ψ_r dominant (both spaces have dimension $mn - 1$)

Bijection

$$\begin{array}{ccc} \mathrm{Crit}(M_r) & \overset{\psi_r}{\dashrightarrow} & \mathrm{Crit}(M_{f(r)}),\ f(r) = s \leq m - r + 1 \\ (P, U) & \dashrightarrow & ((R \cdot \frac{U}{P} \cdot K)/(u_{++}^3), U) \end{array}$$

Observation

- ψ_r injective
- ψ_r dominant (both spaces have dimension $mn - 1$)

$\rightsquigarrow$ ψ_r birational, $f : \{1, \ldots, m\} \to \{1, \ldots, m\}$ bijection

$\rightsquigarrow$ $f(r) = m - r + 1$

Bijection

$$\mathrm{Crit}(M_r) \overset{\psi_r}{\dashrightarrow} \mathrm{Crit}(M_{f(r)}),\ f(r) = s \leq m - r + 1$$

$$(P, U) \dashrightarrow ((R \cdot \tfrac{U}{P} \cdot K)/(u^3_{++}), U)$$

Observation

- ψ_r injective
- ψ_r dominant (both spaces have dimension $mn - 1$)

$\rightsquigarrow \psi_r$ birational, $f : \{1, \ldots, m\} \to \{1, \ldots, m\}$ bijection

$\rightsquigarrow f(r) = m - r + 1$

Theorem (D-Rodriguez, 2012)

M_r and M_{m-r+1} are ML-dual. □

Bijection

$$\begin{array}{ccc} \mathrm{Crit}(M_r) & \overset{}{\underset{\psi_r}{\dashrightarrow}} & \mathrm{Crit}(M_{f(r)}),\ f(r) = s \leq m - r + 1 \\ (P, U) & \dashrightarrow & ((R \cdot \frac{U}{P} \cdot K)/(u_{++}^3), U) \end{array}$$

Observation

- ψ_r injective
- ψ_r dominant (both spaces have dimension $mn - 1$)

$\rightsquigarrow \psi_r$ birational, $f : \{1, \ldots, m\} \to \{1, \ldots, m\}$ bijection

$\rightsquigarrow f(r) = m - r + 1$

Theorem (D-Rodriguez, 2012)

M_r and M_{m-r+1} are ML-dual. $\square$

Further work

symmetric/alternating matrices!

tensors? other ML-dual pairs of varieties?