

The maximum likelihood degree of a very affine variety

June Huh

University of Michigan at Ann Arbor

October 21, 2012

The maximum likelihood degree

- \mathbb{P}^{n-1} , the projective space with coordinates p_1, \dots, p_n , where the p_i represents the probability of the i -th event.
- An implicit statistical model is a closed subvariety $V \subseteq \mathbb{P}^{n-1}$.
- The data comes in the form of integers u_1, \dots, u_n , where the u_i is the number of times the i -th event was observed.

The maximum likelihood degree

- In order to find the values of p_i on V which best explains the given data u_i , one finds critical points of the likelihood function

$$L(p_1, \dots, p_n) = p_1^{u_1} \cdots p_n^{u_n} / (p_1 + \cdots + p_n)^{u_1 + \cdots + u_n}.$$

- *The maximum likelihood degree* of the model is the number of critical points of $L|_V$, for sufficiently general u_1, \dots, u_n .
- This number is well-defined.

Log-concavity conjectures

- A sequence a_0, \dots, a_p is *log-concave* if for all i

$$a_{i-1} a_{i+1} \leq a_i^2.$$

- If there are *no internal zeroes*, log-concavity implies *unimodality*:

$$a_0 \leq \dots \leq a_{i-1} \leq a_i \geq a_{i+1} \geq \dots \geq a_p \quad \text{for some } i.$$

Chromatic polynomial of graphs

Let G be a graph.

- The chromatic polynomial of G is the function

$$\chi_G(q) = (\text{number of proper colorings of } G \text{ with } q \text{ colors}).$$

Example

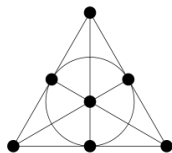


$$\begin{aligned}\chi_G(q) &= q(q-1)(q-2)(q-2) + q(q-1)(q-1) \\ &= 1q^4 - 4q^3 + 6q^2 - 3q\end{aligned}$$

Independent sets of vectors

- Let $\mathcal{A} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a subset of a vector space V .
- Define the f -vector of \mathcal{A} by

$f_i =$ (number of independent subsets of cardinality i in \mathcal{A}).



Example (Fano plane)

For $\mathcal{A} = \mathbb{F}_2^3 \setminus \{0\}$, we have

$$f_0 = 1, \quad f_1 = 7, \quad f_2 = 21, \quad f_3 = 28.$$

Conjecture (Read-Rota 68, Mason-Welsh 69)

- *The first sequence is log-concave for any G .*
- *The second sequence is log-concave for any \mathcal{A} .*

- These conjectures are proved (H, H-Katz, Lenz) using the topology of hypersurface complements, CSM class, tropical geometry, and matroid theory.
- The maximum likelihood framework provides solutions to stronger conjectures (Hoggar 74, Dawson 84, Colbourn 87) with simpler proofs, in characteristic zero.

Theorem

Let M be a matroid representable over a field of characteristic zero.

1. The h -vector of the matroid complex of M is log-concave.
2. The h -vector of the broken circuit complex of M is log-concave.

- “1” implies a conjecture of Colbourn on the reliability of a network.
- “1” was conjectured by Dawson in general.
- Any matroid complex is a broken circuit complex, but not conversely.
Therefore “2” is stronger than “1”.

Log-concavity conjectures

Corollary

Let M be a matroid representable over a field of characteristic zero.

- 1. The f -vector of the matroid complex of M is strictly log-concave.*
- 2. The f -vector of the broken circuit complex of M is strictly log-concave.*

This should be compared to f -vectors and h -vectors of

other “nice” shellable simplicial complexes, such as...

Björner's example (from Ziegler's polytope book)

Examples 8.40. The unimodality conjecture fails for a simplicial polytope of dimension $d = 20$ with the following f -vector, for which $f_{11} > f_{12} < f_{13}$.

$$\begin{array}{rcl} f_{-1} & = & 1 \\ f_0 & = & 4203045807626 \\ f_1 & = & 84060916163336 \\ f_2 & = & 798578704207074 \\ f_3 & = & 4791472253296106 \\ f_4 & = & 20363758019368323 \\ f_5 & = & 65164051780016980 \\ f_6 & = & 162910744316489788 \\ f_7 & = & 325834059588060117 \\ f_8 & = & 529707205213463823 \\ f_9 & = & 709935971390166248 \\ f_{10} & = & 805494832051588614 \\ f_{11} & = & 821976324224631043 \quad / \\ f_{12} & = & 821976324224611712 \quad \leftarrow \\ f_{13} & = & 822000129478641948 \quad \backslash \\ f_{14} & = & 747383755288236256 \\ f_{15} & = & 546761228419958342 \\ f_{16} & = & 293715859557026466 \\ f_{17} & = & 106920718330384544 \\ f_{18} & = & 23458617733909980 \\ f_{19} & = & 2345861773390998 \end{array}$$

How do you prove that a sequence is log-concave?

Let's say that a homology class of an algebraic variety is *representable* if it is the class of a subvariety.

Theorem

Write an even dimensional homology class ξ as a linear combination

$$\xi = \sum_i e_i [\mathbb{P}^{k-i} \times \mathbb{P}^i] \in H_{2k}(\mathbb{P}^m \times \mathbb{P}^n; \mathbb{Z}).$$

Then some positive multiple of ξ is representable **if and only if** $\{e_i\}$ form a log-concave sequence of nonnegative integers with no internal zeros.

The maximum likelihood degree

- Therefore it suffices to show that there is a subvariety of $\mathbb{P}^m \times \mathbb{P}^n$ whose homology class is given by the coefficients of $\chi_M(q + 1)$.
- The idea for this comes from the *maximum likelihood estimation* in algebraic statistics.

Varchenko's conjecture

- \mathcal{A} = (arrangement in \mathbb{C}^r defined by linear functions f_1, \dots, f_n).
- $\varphi = \prod_{i=1}^n f_i^{u_i}$, $u_i \in \mathbb{Z}$.

Conjecture (Varchenko 95)

If the arrangement is essential and the u_i are general, then

1. φ has only finitely many critical points in $\mathbb{C}^r \setminus \mathcal{A}$.
2. All critical points of φ are nondegenerate.
3. The number of critical points is equal to $(-1)^r \chi(\mathbb{C}^r \setminus \mathcal{A})$ ($= \chi_{\mathcal{A}}(1)$).

Varchenko's conjecture is proved by Orlik and Terao.

- If \mathcal{A} is defined over the real numbers,

$$\begin{aligned}\chi_{\mathcal{A}}(1) &= (\text{the number of critical points of } \varphi) \\ &= (\text{the number of bounded regions of } \mathbb{R}^r \setminus \mathcal{A}).\end{aligned}$$

- \mathcal{A} is essential iff lowest dimensional flats are zero dimensional.
- \mathcal{A} is essential iff we have the embedding

$$\mathbb{C}^r \setminus \mathcal{A} \longrightarrow (\mathbb{C}^*)^n, \quad x \longmapsto (f_1, \dots, f_n).$$

- Any arrangement is of the form $\mathcal{A} = \mathbb{C}^k \times \mathcal{A}'$ with \mathcal{A}' essential.

- Let U be an r -dimensional smooth subvariety of $(\mathbb{C}^*)^n$.
- Let f_1, \dots, f_n be the coordinate functions on U .
- $\varphi = \prod_{i=1}^n f_i^{u_i}$, $u_i \in \mathbb{Z}$.

Theorem

If the u_i are sufficiently general, then

1. φ has only finitely many critical points in U .
2. All critical points of φ are nondegenerate.
3. The number of critical points is equal to $(-1)^r \chi(U)$.

- Next we consider the collection of all critical points of all possible master functions simultaneously:

$$\mathfrak{X}^o(U) = \left\{ \sum_{i=1}^n u_i \cdot \text{dlog}(p_i)(x) = 0 \right\} \subseteq U \times \mathbb{P}^{n-1},$$

where u_1, \dots, u_n are now homogeneous coordinates of \mathbb{P}^{n-1} .

- The closure $\mathfrak{X}(U) \subseteq \mathbb{P}^n \times \mathbb{P}^{n-1}$ gives the subvariety we want!
- In fact, this is truly a natural choice, because it is a geometric realization of the characteristic class of U .

Suppose U is r -dimensional and not isomorphic to a torus.

Theorem

If U is wonderful, then

$$[\mathfrak{X}(U)] = \sum_{i=0}^r v_i [\mathbb{P}^{r-i} \times \mathbb{P}^{n-1-r+i}] \in H_*(\mathbb{P}^n \times \mathbb{P}^{n-1}),$$

where

$$c_{SM}(\mathbf{1}_U) = \sum_{i=0}^r (-1)^i v_i [\mathbb{P}^{r-i}] \in H_*(\mathbb{P}^n).$$

If U is the complement $\mathbb{C}^r \setminus \mathcal{A}$, then (almost by definition)

$$\chi_{\mathcal{A}}(q+1) = \sum_{i=0}^r (-1)^i v_i q^{r-i}.$$

Theorem

Let U be the hypersurface $\{g = 0\} \subseteq (\mathbb{C}^*)^n$ with

$$c_{SM}(\mathbf{1}_U) = \sum_{i=0}^r (-1)^i v_i [\mathbb{P}^{r-i}] \in H_*(\mathbb{P}^n).$$

If g is general with respect to Δ_g , then U is wonderful, and

$$v_i = MV_n(\underbrace{\Delta_1, \dots, \Delta_r}_{r-i}, \underbrace{\Delta_g, \dots, \Delta_g}_{i+1}), \quad i = 0, \dots, r.$$

In particular, ML-degree of U is equal to the normalized volume

$$v_r = (-1)^r \int c_{SM}(\mathbf{1}_U) = \text{Volume}(\Delta_g).$$