Complete quadrics and Gaussian models

L. Manivel
+ M. Michalek, L. Monin, T. Seynnaeve, M. Vodicka

Toulouse Mathematics Institute

GO 60+1
June 25, 2021

Main problem: compute the ML-degree

Consider $V \subset M_n, SM_n, AM_{2n}$ a linear space of square complex matrices, not all singular, and define the reciprocal

$$V^{-1} := \text{closure of } \{M^{-1}, M \in V\}.$$

The ML-degree (maximum likelihood degree) can be defined as

$$ML(V) = \deg(\mathbb{P}V^{-1}).$$

PROBLEM: Compute it! In particular, compute

$$\phi(n, d) := ML(V)$$

when $V \subset SM_n$ is generic of dimension d.

Why is that difficult? The degree is computed by counting the intersection points with a generic linear space of codimension $d-1$ $\rightsquigarrow$ we impose relations between the maximal minors M_{ij} of $M \in V$:

$$\sum_{ij} a_{ij}^k M_{ij} = 0, \qquad k = 1, \ldots, d-1.$$

Easy fo small d:

$$\begin{aligned} \phi(n,1) &= 1, \\ \phi(n,2) &= n-1, \\ \phi(n,3) &= (n-1)^2. \end{aligned}$$

Problem: starting from $d-1=3$, the matrices of corank two contribute since they verify $M_{ij} = 0 \forall i,j$.
But we want to count only invertible solutions! So

$$\begin{aligned} \phi(n,4) &= (n-1)^3 - \tbinom{n+1}{3}, \\ \phi(n,5) &= \tfrac{1}{12}(n-1)(n-2)(7n^2-19n+6). \end{aligned}$$

The Sturmfels-Uhler conjecture

Conjecture (Sturmfels-Uhler 2009)

$\phi(n, d)$ is a polynomial function of n for each fixed d.

Main result of this talk:

Theorem (MMMSV 2020)

1. The Sturmfels-Uhler conjecture is true.
2. There is an explicit formula for $\phi(n, d)$.

$$\phi(n,10) = \frac{(n-1)(n-2)(n-3)}{362880}(8357n^6 - 114126n^5 + 629471n^4 - 1816902n^3 + 2911016n^2 - 2201088n + 60480).$$

A detour through projective duality

$X \subset \mathbb{P}^N \rightsquigarrow$ projectively dual variety $X^* \subset \hat{\mathbb{P}}^N$:

$$X^* := \{H \in \hat{\mathbb{P}}^N, \exists x \in X, H \supset T_x X\}.$$

There is an incidence correspondence

$$\begin{array}{ccc} & I_X = I_{X^*} & \\ \swarrow & & \searrow \\ \mathbb{P}^N \supset X & & X^* \subset \hat{\mathbb{P}}^N \end{array}$$

where the incidence variety I_X is the *conormal variety* $\mathbb{P}(N_X^*)$ (at least over the smooth locus).

$$[I_X] = \sum_{k=0}^{N+1} \delta_X(k) H_1^k H_2^{N+1-k} \in A^*(\mathbb{P}^N \times \hat{\mathbb{P}}^N).$$

The coefficients are the *multidegrees*.

Example (Plücker formula). Let $C \subset \mathbb{P}^2$ be a curve of degree d, with δ nodes and κ cusps, then $C^* \subset \hat{\mathbb{P}}^2$ is a curve of degree

$$d^* = d(d-1) - 2\delta - 3\kappa.$$

$\rightsquigarrow$ very sensitive to singularities.

Example (Determinantal varieties). Let $D_{n,r} \subset \mathbb{P}(SM_n)$ the variety of symmetric matrices of rank $\leq r$.
Consider a general point $[M]$, with M a matrix of rank r:

$$M = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0}_{n-r} \end{array}\right), \qquad \hat{T}_{[M]} D_{n,r} = \left(\begin{array}{c|c} * & * \\ \hline * & \mathbf{0}_{n-r} \end{array}\right).$$

So a hyperplane H is tangent to $D_{n,r}$ at $[M]$ iff it is defined by a matrix of the form

$$\left(\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & * \end{array}\right).$$

We can deduce the following facts:

- The projective dual of $D_{n,r}$ is $D_{n,n-r}$.
- The incidence variety is essentially

$$I_{D_{n,r}} \simeq \mathbb{P}(S^2U) \times_{Gr(r,n)} \mathbb{P}(S^2Q^*).$$

- The codegree of $D_{n,r}$ can be computed.

Theorem (Harris-Tu 1984)

The degree of the determinantal variety $D_{n,n-r}$ is

$$\deg(D_{n,n-r}) = \prod_{k=0}^{r-1} \frac{\binom{n+k}{r-k}}{\binom{2k+1}{k}}.$$

Note: for r fixed, this is polynomial in n!

Two main ingredients in the proof.

1. Degeneracy loci formula for a symmetric morphism $\varphi : E^* \to E$, where E is some rank n vector bundle over a smooth variety X:

$$[D_r(\varphi)] = 2^{n-r} \det(c_{n-r+i-2j+1}(E))_{1\leq i,j<n-r}.$$

   (Also Jozefiak-Lascoux-Pragacz 1980.)

2. Application to the tautological morphism on $X = \mathbb{P}(SM_n)$

$$\begin{array}{ccc} \mathbb{C}^n \otimes \mathcal{O}_X(-1) & \longrightarrow & \mathbb{C}^n \otimes \mathcal{O}_X, \\ v \otimes M & \mapsto & M(v). \end{array}$$

   $\rightsquigarrow$ determinant of binomial coefficients which has to be evaluated (Giambelli formula vs Weyl dimension formula!).

Problem: Extend the codegree formula to general determinantal varieties $X = D_{n,r} \cap \mathbb{P}M$, at least for M generic of dimension $m+1$.

Beware! The dual of $X = D_{n,r} \cap \mathbb{P}M$ is a hypersurface only in the range given by the *Pataki inequalities*

$$\binom{n-r+1}{2} \leq m \leq \binom{n+1}{2} - \binom{r+1}{2}.$$

$\rightsquigarrow$ codegree/class $\delta(m,n,r)$ to be computed!

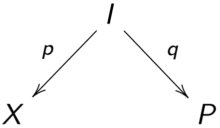
Some good news:

1. The codegree $\delta(m,n,r)$ has a name: the *algebraic degree of semidefinite programming*.
2. The codegree is a multidegree:

$$[I_{D_{n,r}}] = \sum_{m=0}^{\binom{n+1}{2}} \delta(m,n,r) H_1^m H_2^{\binom{n+1}{2}-m}.$$

3. The degeneracy loci formula was extended to Euler topological characteristics $e(D_r(\varphi))$ by Pragacz (1988).

This is classically related to the codegree.
Indeed, consider a general pencil P of sections by hyperplanes H, with intersection L:



Then p is the blow-up of $X \cap L$, while q has *codegree*(X) singular fibers. Computing $e(I)$ in two ways, one obtains the *class formula*

$$(-1)^{dim(X)} codeg(X) = e(X) - 2e(X \cap H) + e(X \cap L).$$

Obviously, X determinantal implies $X \cap H$, $X \cap L$ determinantal. So the Pragacz formula for the Euler topological characteristics yields an explicit formula for the codegree.

OBJECTION: The Pragacz formula applies **only** when X is smooth, which almost never happens if X is determinantal!

Conjecture (Nie-Ranestad-Sturmfels 2006)

The formula is true, notwithstanding singularities.

A bold and brilliant guess! A priori, we know the codegree is very sensitive to singularities. So the conjecture should not be true.

Theorem (MMMSV 2020)

The conjecture is true!

The explicit formula

$$\delta(m, n, n-s) = \sum_{\substack{I=(0\leq i_1<\cdots<i_s<n)\\ \Sigma I\leq m-s}} (-1)^{m-s-\Sigma I} \binom{m-1}{m-s-\Sigma I} \psi_I b_I(n).$$

Ingredients:

1. The ψ_I's are the Lascoux coefficients (more details later).
2. The $b_I(n)$'s are polynomials in n defined by:

$$\sum_{i\geq 0} b_i(n)h^i = \left(\frac{1+h/2}{1-h/2}\right)^n,$$

$$b_{i,j}(n) = b_i(n)b_j(n) + 2\sum_{k>0}(-1)^k b_{i-k}(n)b_{j+k}(n),$$

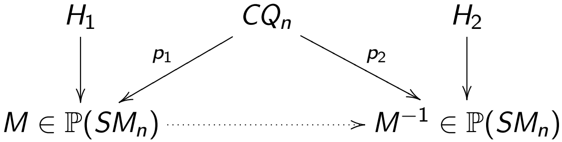
$$b_I(n) = Pf(b_{i_p,i_q}(n)).$$

Back to the ML-degree: complete quadrics

Recall the ML-degree

$$\phi(n,d) = ML(V) = \deg(\mathbb{P}V^{-1}) = \deg(\mathbb{P}\{M^{-1}, M \in V\}).$$

To compute it, need to regularize the rational map $M \mapsto M^{-1}$.
Classically done using the variety CQ_n of *complete quadrics*:



where p_i is the blowup of the determinantal loci $D_{n,1}, \ldots, D_{n,n-2}$
$\rightsquigarrow$ exceptional divisors $E_1, \ldots, E_{n-1}$,
$\rightsquigarrow$ hyperplane divisors $L_1 = p_1^* H_1, L_2 = p_2^* H_2$.

The ML-degree is now given by an intersection number:

$$\phi(n,d) = \int_{CQ_n} L_1^{\binom{n+1}{2}-d} L_2^{d-1}.$$

The intersection ring of CQ_n has been well-studied from Schubert 1879 (the 3264 conics) to De Concini-Procesi 1985 $\rightsquigarrow$ algorithm for a given n. Here we need to fix d, vary n.

Observation: $nL_1 = (n-1)E_1 + (n-2)E_2 + \cdots + E_{n-1}$, so

$$n\phi(n,d) = \sum_{s>0} s \int_{E_{n-s}} L_1^{\binom{n+1}{2}-d-1} L_2^{d-1}.$$

Moreover $E_{n-s} \simeq \mathbb{P}(S^2U) \times_{G(n-s,n)} \mathbb{P}(S^2Q^*)$ is essentially the conormal space to $D_{n,n-s}$, and the integral is a multidegree!

Fundamental identity

$$n\phi(n,d) = \sum_{\binom{s+1}{2}\leq d} s\delta(d,n,n-s).$$

Consequence: NRS conjecture $\Longrightarrow$ SU conjecture + explicit formula!

We focus on $\delta(d,n,n-s)$. Can use push-forward to compute

$$\delta(d,n,n-s) = \int_{E_{n-s}} L_1^{\binom{n+1}{2}-d-1} L_2^{d-1}.$$

Recall that for $\pi : \mathbb{P}(E) \to X$ and $\lambda = c_1\mathcal{O}_E(1)$, we have $\pi_*\lambda^k = s_{k-e+1}(E)$, a Segre class. We get

$$\delta(d,n,n-s) = \int_{G(n-s,n)} s_{\binom{n+1}{2}-\binom{n-s+1}{2}-d}(S^2U^*) s_{d-\binom{s+1}{2}}(S^2Q).$$

The Lascoux coefficients (Laksov Lascoux Thorup 1989)

$$s(S^2E) = \sum_{I=(0\leq i_1<\cdots<i_e)} \psi_I s_{\lambda(I)}(E).$$

Here $\lambda(I) = (i_e - e + 1, \ldots, i_2 - 1, i_1)$ is a partition and $s_\lambda(E)$ the characteristic class associated to the Schur function $s_\lambda(x_1, \ldots, x_e) = s_\lambda(x)$, so

$$\prod_{1\leq i,j\leq e} \frac{1}{1 - x_i - x_j} = \sum_I \psi_I s_{\lambda(I)}(x).$$

First values: $\psi_i = 2^i$ and $\psi_{ij} = \sum_{i<k\leq j} \binom{i+j}{k}$.

The Lascoux Pfaffian formula

$$\psi_I = Pf(\psi_{i_p,i_q}).$$

$$\begin{aligned}\psi_{1235} &= \psi_{12}\psi_{35} - \psi_{13}\psi_{25} + \psi_{15}\psi_{23} \\ &= 3 \times 126 - 10 \times 91 + 56 \times 10 = 28.\end{aligned}$$

Another formula: consider the infinite Pascal matrix $M_{i,j} = \binom{i}{j}$,

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 3 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 4 & 6 & 4 & 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 5 & 10 & 10 & 5 & 1 & 0 & 0 & 0 & \dots \end{pmatrix}$$

The Lascoux minor formula

$\psi_I = \sum_J M_{I,J}$, sum of minors with rows indexed by I.

$$\psi_{1235} = \sum 4 \times 4 \text{ minors of } \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 & 0 \\ 1 & 5 & 10 & 10 & 5 & 1 \end{pmatrix} = 28.$$

Plugging Lascoux coefficients into the formula for $\delta(d, n, n-s)$, we get

$$\delta(d, n, n-s) = \sum_{\substack{I=(i_1<\cdots<i_s)\subset[n] \\ \Sigma I = d-s}} \psi_I \psi_{[n]\setminus I},$$

where $[n] = \{0, \ldots, n-1\}$ (Ranestad & von Bothmer 2007).

Theorem (MMMSV 2020)

$\forall I$, the Lascoux coefficient $LP_I(n) = \psi_{[n]\setminus I}$ is polynomial in n.

Proof by induction formulas, or Pfaffian formulas:

Pfaffian formula

$$LP_I(n) = Pf(LP_{i_p,i_q}(n)).$$

$\rightsquigarrow$ Sturmfels-Uhler conjecture.

For the NRS conjecture, one needs a more precise statement. Define coefficients $s_{I,J}$, for I, J increasing sequences of length e, by the identity

$$s_{\lambda(I)}(E \otimes L) = \sum_{J \leq I} s_{I,J} s_{\lambda(J)}(E) c_1(L)^{\Sigma I - \Sigma J}.$$

Theorem (MMMSV 2020)

$$LP_I(n) = \sum_{J \leq I} \left(-\frac{1}{2}\right)^{\Sigma I - \Sigma J} s_{I,J} b_J(n).$$

And this implies the NRS conjecture.

We have two proofs of this statement:

1. By induction using our Pfaffian formulas for $LP_I(n)$.
2. By applying a projection formula for Schur Q-polynomials due to Pragacz.

Schur Q-polynomials

One defines polynomials Q_I in variables $x_1, \ldots, x_n$ as follows:

$$\sum_{i \geq 0} Q_i(x) h^i = \prod_{k=1}^{n} \frac{1 + hx_k}{1 - hx_k},$$

$$Q_{i,j}(x) = Q_i(x) Q_j(x) + 2 \sum_{k>0} (-1)^k Q_{i-k}(x) Q_{j+k}(x),$$

$$Q_I(x) = Pf(Q_{i_p, i_q}(x)).$$

$\rightsquigarrow$ Important in representation theory (of symmetric groups and Lie super-algebras) and Schubert calculus on classical Grassmannians.

Note that $Q_I(\frac{1}{2}, \cdots, \frac{1}{2}) = b_I(n)$.

By the usual splitting principle, we get characteristic classes $Q_I(\mathcal{E})$ for any complex vector bundle $\mathcal{E}$.

Projection formula (Pragacz 1994)

Let $\mathcal{E} \to X$ be a rank n vector bundle, $\pi : G^s(\mathcal{E}) \to X$ be the relative Grassmannian of rank s quotients, with $\mathcal{K}$ and $\mathcal{Q}$ the tautological bundles. Then

$$Q_{I+1^s}(\mathcal{E}) = \pi_*(c_{top}(\mathcal{K} \otimes \mathcal{Q}) Q_{I+1^s}(\mathcal{Q})).$$

Application: suppose that $\mathcal{E} = \mathcal{E}_0 \otimes L$ for a line bundle L and a trivial vector bundle $\mathcal{E}_0$. Then $G^s(\mathcal{E}) = G^s(\mathbb{C}^n) \times X$ and

$$\mathcal{K} = \mathcal{K}_0 \boxtimes L, \qquad \mathcal{Q} = \mathcal{Q}_0 \boxtimes L.$$

The projection formula gives exactly the desired identity!!

Final remarks

1. We have **similar results** for general square matrices and skew-symmetric matrices
$\rightsquigarrow$ analogues of the NRS conjecture!

   Same pattern by replacing complete quadrics by complete collineations and using projection formulas.

2. Recall that the NRS conjecture should not be true since it applied the class formula in a context where it should fail.
$\rightsquigarrow$ Why does it hold true?

   There is an extension of the class formula using *MacPherson's Euler obstruction classes* $\hat{e}(X)$ (Conan-Leung 2001). There should exist an extension of the Pragacz formula involving these obstruction classes!?

Thanks for your attention!

Happy Birthday Giorgio!!