The Degree of SO(r) and Low-rank SDP

Madeline Brandt, me, Taylor Brysiewicz, Robert Krone, Elina Robeva

March 31, 2017

The Main Question

Question (B. Sturmfels)

What is the degree of the special orthogonal group?

Set Up - SO(r)

- The **special orthogonal group** is the group:

$$\mathrm{SO}(r) := \left\{ M \in \mathrm{Mat}(r,r) \quad \middle| \det M = 1 \quad MM^t = \mathrm{Id} \right\}.$$

- The conditions for $M \in \mathrm{SO}(r)$ are polynomial in the entries of M.
- **Example:** If $r = 2$ then:

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} \middle| \begin{array}{c} x_{1,1}x_{2,2} - x_{1,2}x_{2,1} = 1, \\ x_{1,1}^2 + x_{1,2}^2 = 1, x_{2,1}^2 + x_{2,2}^2 = 1, \\ x_{1,1}x_{2,1} + x_{1,2}x_{2,2} = 0 \end{array} \right\} \subset \mathbb{C}^4$$

- This makes SO(r) an algebraic subvariety of $\mathbb{C}^{r^2}$.

Set Up - Degree

- The degree of a variety is a geometric invariant that intuitively measures the complexity of the variety.

Definition

If $X \subset \mathbb{C}^r$ is a subvariety of dimension k then the **degree** *of X is:*

$$\deg X := \#X \cap H,$$

where $H \subset \mathbb{C}^r$ is a generic plane of dimension $r - k$.

- Generic (essentially) means that the plane is transverse to X.
- More formally we are interested in the degree of the projective closure of $\mathrm{SO}(r)$ under the embedding given by the equations:

$$\det M = \mathrm{Id} \quad \text{and} \quad MM^t = \mathrm{Id}\,.$$

Some Answers

- Using symbolic and numerical techniques Macaulay2 / Bertini can answer this for small r.

r	Symbolic	Numerical
2	2	2
3	8	8
4	40	40
5	384	384
6	-	4768
7	-	111616
8	-	-

Table: The degree of SO(r) computed various ways.

Some Answers

- These computations quickly become difficult as the dimension (and degree) of $\mathrm{SO}(r)$ grow rapidly in r.
- Standard techniques were only effective for $r \leq 5$.
- A more efficient approach is to use monodromy methods.
  1. Pick random hyperplanes H and H' s.t. $H \cap \mathrm{SO}(r)$ contains Id.
  2. Pick a homotopy γ from $H \cap \mathrm{SO}(r)$ to $H' \cap \mathrm{SO}(r)$.
  3. Pick another homotopy γ' from $H' \cap \mathrm{SO}(r)$ to $H \cap \mathrm{SO}(r)$.
  4. Look at the point $\gamma'(\gamma(\mathrm{Id})) \in H \cap \mathrm{SO}(r)$.
- We were able to compute the degree of $\mathrm{SO}(7)$ in ≈ 12 hours, compared to the $30+$ days we expected with other methods.

The Answer

Theorem (B^3KR)

$$\deg \mathrm{SO}(r) = 2^{r-1} \det \left(\binom{2r-2i-2j}{r-2i} \right)_{1 \le i,j \le \lfloor \frac{r}{2} \rfloor}.$$

- We prove this by making explicit an earlier result of Kazarnovskij.

The Answer

Theorem (B^3KR)

$$\deg \mathrm{SO}(r) = 2^{r-1} \det \left(\binom{2r-2i-2j}{r-2i} \right)_{1 \le i,j \le \lfloor \frac{r}{2} \rfloor}.$$

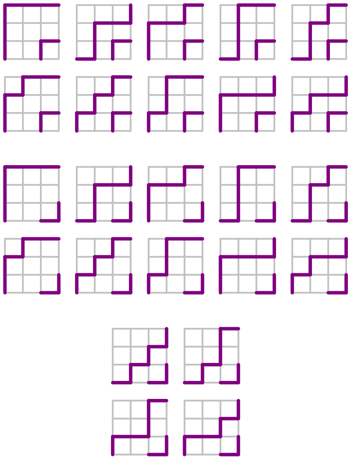
- We prove this by making explicit an earlier result of Kazarnovskij.
- The appearance of the binomial coefficents allows us to get by following corollary via the Gessel-Viennot Lemma.

Corollary (B^3KR)

$$\deg \mathrm{SO}(r) = 2^{r-1} \cdot \# \left\{ \begin{matrix} \textit{Certain Non-intersecting} \\ \textit{Planar Lattice Paths} \end{matrix} \right\}.$$

The Answer

- **Ex:** The degree of $SO(5)$ is $384 = 2^4 \cdot 24$, and here are 24 non-intersecting planar lattice paths:



The Real Main Question

Question (B. Recht)

Why does the augmented Lagrangian algorithm for solving low-rank semidefinite programing problems work?

Linear Programing

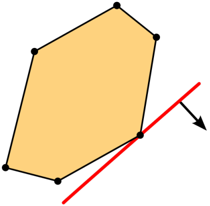
Linear Programing

Given a cost function $\vec{c} \in \mathbb{R}^n$ *and constraints* $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^m$ *the linear programing question* $(c; a_1, \ldots, a_m; b)$ *is to*

$$\min_{x \in \mathbb{R}^n} \langle \vec{c}, x \rangle$$
$$\textit{subject to}\ \langle \vec{a}_i, x \rangle = b_i \quad \textit{and}\ x \geq 0.$$

Linear Programing

- **Idea:** Optimize some function $\langle \vec{c}, x \rangle$ over a convex polytope given by the intersection of the positive orthant with the region defined by the hyperplanes $\langle \vec{a}_i, x \rangle = b_i$.



- There are many algorithms (ellipsoid, simplex, etc.) for efficiently solving linear programing problems.

Theorem (Khachiyan)

LP problems can be solved in polynomial time in n, the # of unknowns.

Semidefinite Programing

- Let $\mathbb{S}^n$ be the vector space of $n \times n$ real symmetric matrices.
- We can equip $\mathbb{S}^n$ with an inner product by setting:

$$\langle A, B \rangle := \operatorname{tr}\left(A^t B\right).$$

Semidefinite Programing

- Let $\mathbb{S}^n$ be the vector space of $n \times n$ real symmetric matrices.
- We can equip $\mathbb{S}^n$ with an inner product by setting:

$$\langle A, B \rangle := \operatorname{tr}\left(A^t B\right).$$

- A matrix M is **positive semidefinite** if and only if for every $\vec{x} \in \mathbb{R}^n$:

$$\vec{x}^t M \vec{x} \geq 0.$$

Lemma

The set of positive semidefinite matrices forms a convex cone inside of $\mathbb{S}^n$.

Semidefinite Programing

- **Idea:** Semidefinite programing generalizes linear programing by replacing $\mathbb{R}^n$ with $\mathbb{S}^n$ and the positive orthant with the p.s.d cone.

Semidefinite Programing

Given a cost function $C \in \mathbb{S}^n$ and constraints $A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $\vec{b} \in \mathbb{R}^m$ the semidefinite programing question $(C; A_1, \ldots, A_m; \vec{b})$ is to

$$\min_{X \in \mathbb{S}^n} \langle C, X \rangle$$

$$\textit{subject to} \langle A_i, X \rangle = b_i$$

$$\textit{and } X \textit{ is p.s.d.}$$

Solving SDP Problems

- Notice the region we are optimizing over is still convex, and so methods of convex optimization can be applied.

Theorem

A SDP problem can be solved in polynomial time in n, the # of unknowns.

- In practice, for many applications n is huge, which makes many of the standard algorithms costly, especially when the problem is sparse.

Solving SDP Problems

- Notice the region we are optimizing over is still convex, and so methods of convex optimization can be applied.

Theorem

A SDP problem can be solved in polynomial time in n, the # of unknowns.

- In practice, for many applications n is huge, which makes many of the standard algorithms costly, especially when the problem is sparse.
- **However**, the rank of the solution is often quite small.

Lemma

Let $M \in \mathbb{S}^n$ then:

$$M \text{ is p.s.d, } \operatorname{rank} M \leq r \iff M = RR^t, \text{ for some } R \in \operatorname{Mat}(n, r).$$

Solving SDP Problems

non-Linear Semidefinite Programing (Burer and Monteiro)

Given a cost function $C \in \mathbb{S}^n$ *and constraints* $A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $\vec{b} \in \mathbb{R}^m$ *the non-linear semidefinite programing question* $(C; A_1, \ldots, A_m; \vec{b})$ *is to*

$$\min_{R \in \mathrm{Mat}_{n,r}} \langle C, RR^t \rangle$$

$$\textit{subject to} \langle A_i, RR^t \rangle = b_i.$$

- Notice the condition that the solution RR^t be p.s.d is now implicit.

Solving SDP Problems

non-Linear Semidefinite Programing (Burer and Monteiro)

Given a cost function $C \in \mathbb{S}^n$ *and constraints* $A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $\vec{b} \in \mathbb{R}^m$ *the non-linear semidefinite programing question* $(C; A_1, \ldots, A_m; \vec{b})$ *is to*

$$\min_{R \in \mathrm{Mat}_{n,r}} \left\langle C, RR^t \right\rangle$$

$$\textit{subject to} \left\langle A_i, RR^t \right\rangle = b_i.$$

- Notice the condition that the solution RR^t be p.s.d is now implicit.

Advantages	**Disadvantages**
Fewer Unknowns	non-Linear in Constraints
	non-Linear in Cost Function
	non-Convex

non-Linear SDP Problems

- Burer and Monteiro proposed an algorithm that experimentally outperforms existing methods of solving SDP problems.
- Their method is essentially Lagrange multipliers.
- Despite its appealing experimental behavior it is not known to converge, and there are generally numerous local minima.

Question

How many critical points does the Lagrange equation arising from non-linear semidefinite programming (nLSDP) problem have?

Number of Critical Points

Theorem (B^3KR)

There is an explicit constant $\delta(n, m, r)$ such that:

$$\#\left\{\begin{matrix}\textit{critical points}\\ \textit{of}\\ \textit{nLSDP Algorithm}\end{matrix}\right\} = \delta(n, m, r)\deg \mathrm{SO}(r)$$

- Combining this with our other main result given an explicit formula.

Corollary

$$\#\left\{\begin{matrix}\textit{critical points}\\ \textit{of}\\ \textit{nLSDP Algorithm}\end{matrix}\right\} = \delta(n, m, r)2^{r-1}\det\left(\binom{2r-2i-2j}{r-2i}\right)_{1\le i,j\le\lfloor\frac{r}{2}\rfloor}.$$

More On Convergence

- Notice that that the number of critical points grows rapidly in r.
- The appealing behavior of this algorithm is unexplained by our result and remains mysterious.
- **But...** some of the critical points we counted are extraneous. In particular, when solving an SDP problem we are only interested in those critical points that:
  - are real (rational) valued points,
  - and satisfies additional linear constraints.
- It would be interesting to refine our theorem so that it only counts such critical points.

Real Locus of SO(r)

Question

Can the degree of SO(r) *always be realized over* $\mathbb{R}$*?*

Question

Does there exist a plane H *such that:*

$$\deg \mathrm{SO}(r) = \#(H \cap \mathrm{SO}(r) \cap \mathbb{R}^{r^2})?$$

- This is a question of real algebraic geometry, and seems quite difficult.
- **Example:** Both of these smooth surfaces have degree two:

$$X_1 := \left\{x^2 + y^2 = -1\right\}$$
$$X_2 := \left\{x^2 + y^2 = 1\right\},$$

but X_1 has no real points. The degree of X_2 can be realized over $\mathbb{R}$.

Real Locus of SO(r)

- We explored this question computationally using Macaulay2 / Bertini.
- We randomly generated generic hyperplanes and recorded the number of real points witnessed in the resulting intersection.

	Degree	Most $\mathbb{R}$-Points	# of Trials
SO(3)	8	8	1398000
SO(4)	40	30	1004100
SO(5)	384	76	48200

- We relied on polynomial homotopy continuation and monodromy techniques for these computations.

Real Locus of SO(r)

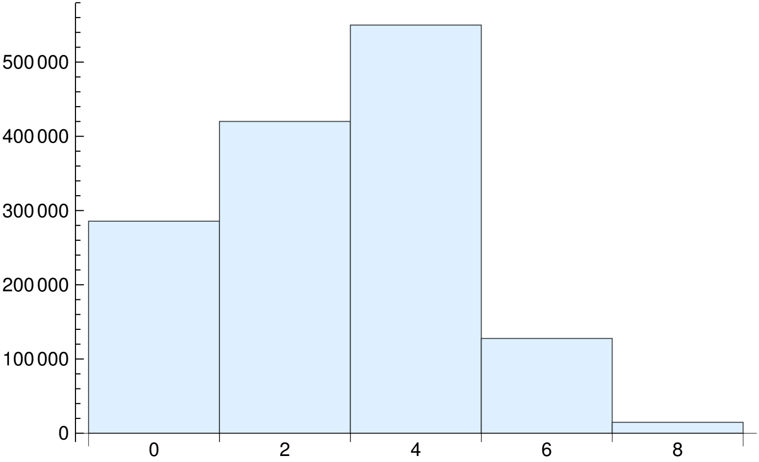


Figure: Number of Real Points in Witness Sets for SO(3)

Real Locus of SO(r)

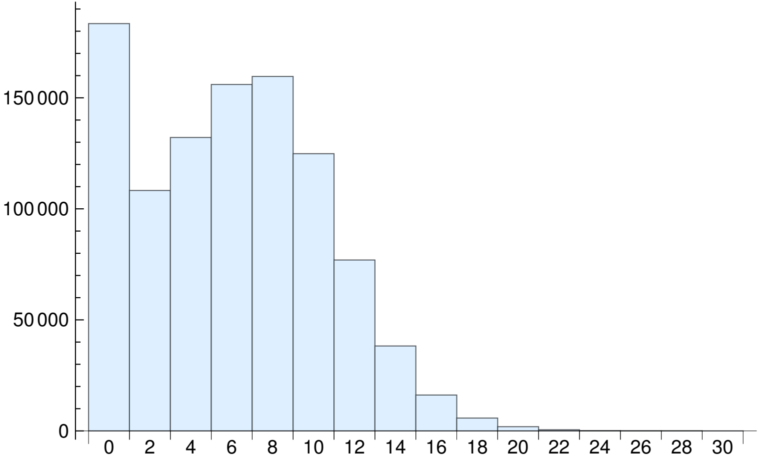


Figure: Number of Real Points in Witness Sets for SO(4)

Real Locus of SO(r)

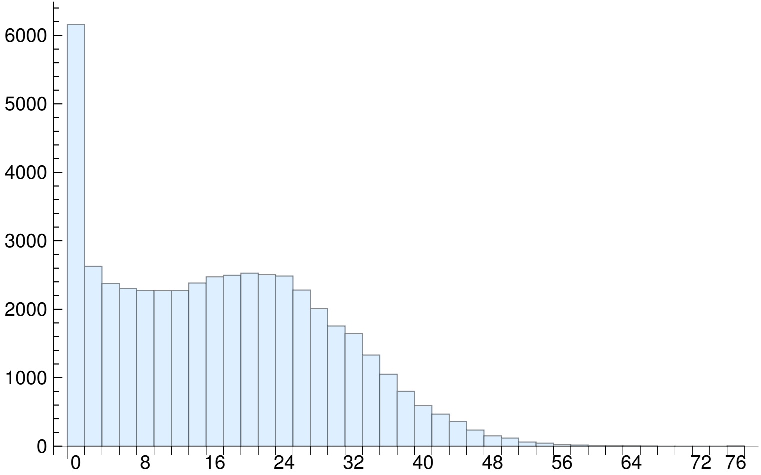


Figure: Number of Real Points in Witness Sets for SO(5)

Thanks

- I would like to thank the following individuals for their valuable advice, guidance, and helpful conversations throughout various parts of this:
  - Anton Leykin
  - Ben Recht
  - Bernd Sturmfels.
- I would also like to acknowledge the following organizations for their generous support:
  - Fields Institute,
  - Max Planck Institute for Mathematics in the Sciences - Leipzig
  - NSF GRFP - Grant No. DGE-1256259,
  - the University of Wisconsin - Madison,
  - Wisconsin Alumni Research Foundation.
- Finally, I want to thank David A. Cox, Megumi Harada, Diane Maclagan, Greg Smith, and Ravi Vakil for organizing the special semester on Combinatorial Algebraic Geometry.