Invariant theory and optimization

Ankit Garg
Microsoft Research India
WACT 2019

Overview

- Natural *computational questions* in invariant theory.
- Fundamental problems phrased in this language.
- Surprising avenues for convexity: *geodesic convexity* and *moment polytopes*.
- Optimization/analytic algorithms.

Outline

- Invariant theory
- Null cone: expressive problem
- Orbit-closure intersection
- Open problems

Invariant theory

Linear actions of groups

Group G acts on vector space V $(= \boldsymbol{C}^d)$.

$M: G \to GL(V)$ $(d \times d$ matrices) group homomorphism.

$M_g: V \to V$ invertible linear map $\forall\, g \in G$.

$M_{g_1 g_2} = M_{g_1} M_{g_2}$ and $M_{id} = id$.

Example 1

$G = S_n$ acts on $V = \boldsymbol{C}^n$ by *permuting coordinates*.

$$M_\sigma\,(x_1, \ldots, x_n) = \left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right).$$

Example 2

$G = GL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})$ by *conjugation*.

$$M_A\, X = AXA^{-1}.$$

Objects of study

Group G acts on vector space V.

- Invariant polynomials: Polynomial functions on V invariant under action of G. p s.t. $p(M_g v) = p(v)$ for all $g \in G, v \in V$.
- Orbits: Orbit of vector v, $O_v = \{M_g v : g \in G\}$.
- Orbit-closures: Orbits may not be closed. Take their closures.

  Orbit-closure of vector v, $\overline{O_v} = \mathrm{cl}\,\{M_g v : g \in G\}$.

Example 1

$G = S_n$ acts on $V = \boldsymbol{C}^n$ by permuting coordinates.

$$M_\sigma\ (x_1, \ldots, x_n) \to \left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right).$$

- Invariants: symmetric polynomials.
- Orbits: x, y in same orbit iff they are of *same type*. $\forall c \in \boldsymbol{C}$, $|\{i: x_i = c\}| = |\{i: y_i = c\}|$.
- Orbit-closures: same as orbits.

Example 2

$G = GL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})$ by *conjugation*.

$$M_A\, X = AXA^{-1}.$$

- Invariants: trace of powers. $\mathrm{tr}(X^i)$.
- Orbits: Characterized by *Jordan normal form*.
- Orbit-closures: differ from orbits.

1. $\overline{O_X} \neq O_X$ iff X *not diagonalizable*.
2. $\overline{O_X}$ and $\overline{O_Y}$ intersect iff *same eigenvalues*.

Orbits and orbit-closures in algebraic complexity

Capture several interesting problems in theoretical computer science.

- *Graph isomorphism*: Whether orbits of two graphs the same. Group action: permuting the vertices.
- *Arithmetic circuits*: The VP vs VNP question. Whether permanent lies in the orbit-closure of the determinant. Group action: Action of $GL_{n^2}(\boldsymbol{C})$ on polynomials induced by action on variables.
- *Border rank*: Whether a tensor lies in the orbit-closure of the diagonal unit tensor. Group action: Natural action of $GL_n(\boldsymbol{C}) \times GL_n(\boldsymbol{C}) \times GL_n(\boldsymbol{C})$.

Invariant ring

Group G acts *linearly* on vector space V.

$\boldsymbol{C}[V]^G$: ring of invariant polynomials.

[Hilbert 1890, 93]: $\boldsymbol{C}[V]^G$ is *finitely generated* (quite generally).

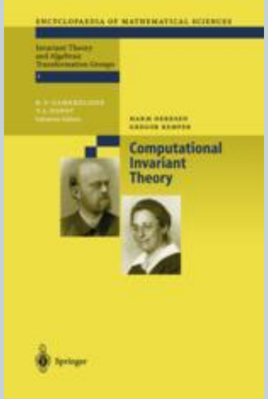
1. $G = S_n$ acts on $V = \boldsymbol{C}^n$ by permuting coordinates.
- $\boldsymbol{C}[V]^G$ generated by *elementary symmetric* polynomials.
2. $G = GL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})$ by conjugation.
- $\boldsymbol{C}[V]^G$ generated by $\mathrm{tr}(X^i)$, $1 \leq i \leq n$.

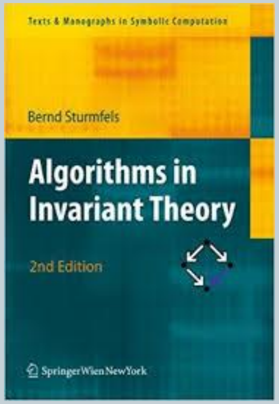
Computational invariant theory

Highly *algorithmic field*.
Algorithms sought and well developed.

Main problems:
- Describe all invariants (*generating set*).
- Simpler: *degree bounds* for generating set.
- *Isomorphism*/orbit problems: When are two objects the "same"?

1. Orbit intersection.
2. Orbit-closure intersection.
3. Orbit-closure containment.
4. Simpler: null cone. When is an object like 0? $0 \in \overline{O_v}$?





Null cone

Null cone

Group G acts on vector space V.

Null cone: Vectors v s.t. 0 lies in the orbit-closure of v.

$$\{v : 0 \in \overline{O_v}\}.$$

Sequence of group elements $g_1, \dots, g_k, \dots$ s.t. $\lim_{k \to \infty} M_{g_k} v = 0$.

Problem: Given $v \in V$, decide if it is in the null cone.

Captures many interesting questions.

[Hilbert 1893; Mumford 1965]: v in null cone iff $p(v) = 0$ for all homogeneous invariant polynomials p.

- One direction clear (polynomials are continuous).
- Other direction uses *Nullstellansatz* and some algebraic geometry.

Example 1

$G = S_n$ acts on $V = \boldsymbol{C}^n$ by permuting coordinates.

$$M_\sigma(x_1, \ldots, x_n) \to \left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right).$$

Null cone = $\{0\}$.

No closures.

Example 2

$G = GL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})$ by *conjugation*.

$$M_A\, X = AXA^{-1}.$$

- Invariants: generated by $\operatorname{tr}\left(X^i\right)$.
- Null cone: nilpotent matrices.

Example 3

$G = SL_n(\boldsymbol{C}) \times SL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})$ by left-right multiplication.

$$M_{(A,B)}\, X = AXB.$$

- Invariants: generated by $\mathrm{Det}(X)$.
- Null cone: Singular matrices.

Example 4

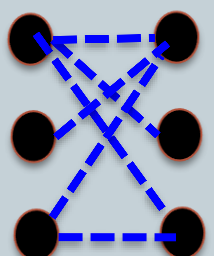
ST_n: group of $n \times n$ diagonal matrices with determinant 1.

$G = ST_n \times ST_n$ acts on $V = M_n(\boldsymbol{C})$ by left-right multiplication.

$$M_{(A,B)}\, X = AXB.$$

- Invariants: generated by $X_{1,\sigma(1)} \cdot X_{2,\sigma(2)} \cdots X_{n,\sigma(n)}$.
- Null cone: perfect matching.

  A_H is in null cone iff H has no perfect matching.



H

1	1	1
1	0	0
1	0	1

A_H

Example 5: Linear programming

T_n: (Abelian!) group of $n \times n$ diagonal matrices.

V: (Laurent) polynomials.

G acts on V by scaling variables. $\boldsymbol{t} \in T_n$, $\boldsymbol{t} = \mathrm{diag}(t_1, \dots, t_n)$.

$$M_{\boldsymbol{t}}\, q(x_1, \dots, x_n) = q(t_1 x_1, \dots, t_n x_n).$$

$q = \sum_{\alpha \in \Omega} c_\alpha x^\alpha$. $\mathrm{supp}(p) = \{\alpha \in \Omega : c_\alpha \neq 0\}$.

- Null cone $\leftrightarrow$ Linear Programming

q *not in* null cone $\leftrightarrow$ $0 \in \mathit{conv}\{\alpha : \alpha \in \mathrm{supp}(q)\}$.

- In non-Abelian groups, the null cone (membership) problem is a *non-commutative* analogue of *linear programming*.

Example 6

$G = SL_n(\boldsymbol{C}) \times SL_n(\boldsymbol{C})$ acts on $V = M_n(\boldsymbol{C})^{\oplus m}$ by *simultaneous left-right* multiplication.

$$M_{(B,C)}(X_1, \ldots, X_m) = (BX_1C, \ldots, BX_mC).$$

- Invariants [DW 00, DZ 01, SdB 01, ANS 10]: generated by $\mathrm{Det}(\sum_i D_i \otimes X_i)$.
- Null cone: Non-commutative singularity. Captures non-commutative rational identity testing.

[GGOW 16, DM 16, IQS 16]: Deterministic polynomial time algorithms.

Geometric Invariant Theory: computational perspective

What is *complexity* of *null cone* membership?

GIT puts it in $NP \cap coNP$ (morally).

- Hilbert-Mumford criterion: how to certify membership in null cone.
- Kempf-Ness theorem: how to certify non-membership in null cone.

Kempf-Ness

Group G acts on vector space V.

How to *certify* v not in null cone?

Exhibit *invariant* polynomial P s.t. $P(v) \neq 0$.

Not feasible in general.

Invariants hard to find, high degree, high complexity etc.

Kempf-Ness provides another (efficient) way.

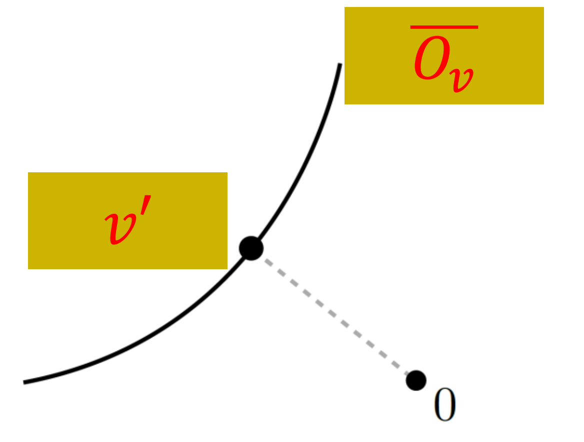
An optimization perspective

Finding *minimal norm* elements in orbit-closures!

Group G acts linearly on vector space V.

$\text{cap}(v) = \inf_{g \in G} \|M_g\, v\|_2^2$.

Null cone: v s.t. $\text{cap}(v) = 0$.



Moment map

Group G acts on vector space V.

$f_v(g) = \|M_g\, v\|_2^2$.

Moment map $\mu_G(v)$: gradient of $f_v(g)$ at $g = id$.

How much *norm* of v decreases by *infinitesimal action* around id.

Much more general.

Moment → *momentum*.

Fundamental in *symplectic geometry* and *physics*.

Kempf-Ness

Group G acts on vector space V.

[Kempf, Ness 79]: v not in null cone iff *non-zero* w in *orbit-closure* of v s.t. $\mu_G(w) = 0$.

w *certifies* v not in null cone.

One direction easy.

- v not in null cone. Take w vector of *minimal norm* in orbit-closure of v. w non-zero.
- w minimal norm in its orbit. ⇒ Norm does not decrease by infinitesimal action around id. ⇒ $\mu_G(w) = 0$.
- *Global* minimum ⇒ *local* minimum.

Kempf-Ness

Other direction: *local* minimum ⇒ *global*. Some "*convexity*".

- *Commutative* group actions – *Euclidean convexity* (change of variables) [exercise].
- *Non-commutative* group actions: *geodesic convexity*.

Orbit-closure intersection

Orbit-closure intersection

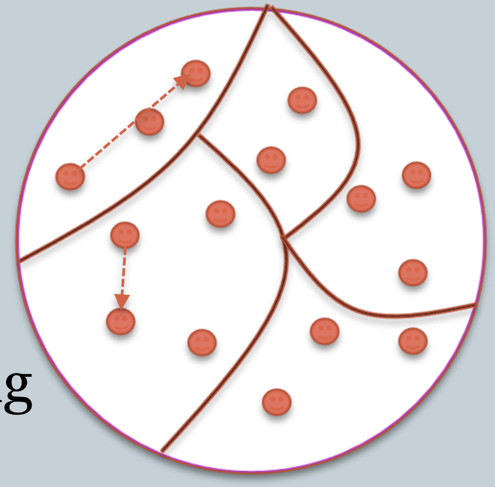
- Group G acts on vector space V.
- *Equivalence relation*: $v_1 \sim v_2$ if orbit-closures intersect.
- Problem: Given v_1, v_2 test equivalence.
- Could be useful for orbit equivalence for random orbits.

Polynomial equivalence

- Polynomials $p, q \in C[\underline{x}]^d$ *equivalent* if there exists invertible A s.t. $p(\underline{x}) = q(A\underline{x})$.
- Problem: Given equivalent p, q, find an equivalence.
- [Kayal 11, Kayal 12]: Algorithms for special cases.
- Open for $d = 3$ in average case.
- [Patarin 96]: Cryptosystem.
- New analytic approach.

Canonical elements

- Group G acts on vector space V. $G = GL_n$ for simplicity.
- *Equivalence relation*: $v_1 \sim v_2$ if orbit-closures intersect.
- Problem: Given v_1, v_2 test equivalence.
- Strategy for testing equivalence: find *canonical* elements and test if equal.



- How to define canonical elements so that finding them is *efficient*?
- Fundamental theorems in invariant theory: *minimal norm* elements canonical (*up to unitary* action).
- Reduce problem to simpler unitary subgroup.

Open problems

Open problems

- Polynomial time algorithms for

1. *Null cone* membership. Simpler: $NP \cap coNP$?
2. *Orbit-closure* intersection. $NP \cap coNP$?
3. *Polynomial equivalence* in average case.

- Algebraic algorithms for *linear programming*?

Thank You

- IAS workshop videos:

https://www.math.ias.edu/ocit2018

- Avi's CCC 2017 tutorial:

http://computationalcomplexity.org/Archive/2017/tutorial.php

Hilbert-Mumford

Group G acts linearly on vector space V.

How to *certify* $v \in N_G(V)$ (null cone)?

Sequence of group elements $g_1, \ldots, g_k, \ldots$ s.t. $\lim_{k\to\infty} M_{g_k} v = 0$.

Compact description of the sequence?

Given by *one-parameter subgroups*.

[Hilbert 1893; Mumford 1965]: $v \in N_G(V)$ iff $\exists$ one-parameter subgroup $\lambda: \boldsymbol{C}^* \to G$ s.t. $\lim_{t\to 0} M_{\lambda(t)} v = 0$.

One-parameter subgroups

One-parameter subgroup: *Group homomorphism* $\lambda: \boldsymbol{C}^* \to G$.
Also map *algebraic*.

- $G = \boldsymbol{C}^*$:

$$\lambda(t) = t^a,\ a \in \boldsymbol{Z}.$$

- $G = (\boldsymbol{C}^*)^{\times n}$:

$$\lambda(t) = (t^{a_1}, \dots, t^{a_n}),\ a_1, \dots, a_n \in \boldsymbol{Z}.$$

- $G = GL_n$:

$$\lambda(t) = S\, \mathrm{diag}(t^{a_1}, \dots, t^{a_n})\, S^{-1},\ S \in GL_n\,,\ a_1, \dots, a_n \in \boldsymbol{Z}.$$

Example

$G = ST_n \times ST_n$ acts on $V = M_n$.

ST_n: $n \times n$ diagonal matrices with det 1.

$M_{(D_1,D_2)}\, X = D_1\, X\, D_2$.

One-parameter subgroups:

$$\lambda(t) = \left((t^{a_1}, \dots, t^{a_n}), (t^{b_1}, \dots, t^{b_n}) \right)$$

$$a_1, \dots, a_n, b_1, \dots, b_n \in \boldsymbol{Z}\colon \textstyle\sum_i a_i = \sum_j b_j = 0.$$

$\lambda(t)$ sends X to $0 \Leftrightarrow$

$$a_i + b_j > 0 \;\forall\, (i,j) \in \mathrm{supp}(X)$$

$$\mathrm{supp}(X) = \{(i,j) \in [n] \times [n] \colon X_{i,j} \neq 0\}$$

Example

X in *null cone* $\Leftrightarrow \exists\ a_1, \dots, a_n, b_1, \dots, b_n \in \boldsymbol{Z}$:

$$\sum_i a_i = \sum_j b_j = 0$$

s.t. $a_i + b_j > 0\ \forall\ (i, j) \in \mathrm{supp}(X)$.

[Exercise]: $\Leftrightarrow$ *bipartite* graph defined by $\mathrm{supp}(X)$ does not have a *perfect matching*.

[Hint]: Hall's theorem.

Moment polytopes

Non-uniform matrix scaling

(r, c): probability distributions over $\{1, \dots, n\}$.

Non-negative $n \times n$ matrix A.

Scaling of A with *row sums* $r_1, \dots, r_n$

and *column sums* $c_1, \dots, c_n$?

$P_A = \{\text{such } (r, c)\}$.

- [...; Rothblum, Schneider 89]: P_A *convex polytope*!
- $P_A = \{(r, c) \colon \exists\, Q, \operatorname{supp}(Q) \subseteq \operatorname{supp}(A), Q$ marginals $(r, c)\}$.

Commutative group actions: *classical marginal* problems.

Also related to *maximum entropy* distributions.

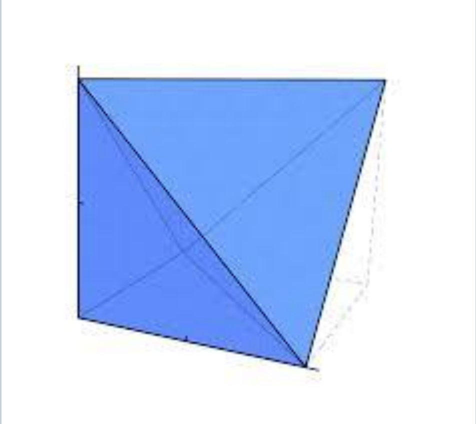
$$\begin{array}{c|c|} & c_1 \quad \dots \quad \dots \quad c_n \\ \hline r_1 & \\ \vdots & \\ & B = RAC \\ \vdots & \\ r_n & \\ \hline \end{array}$$

Quantum marginals

Pure quantum state $|\psi\rangle_{S_1,\ldots,S_d}$ (d quantum systems).

Characterize marginals $\rho_{S_1},\ldots,\rho_{S_d}$ (marginal states on systems)?

Only the spectra matter (local rotations for free).



- Collection of such spectra *convex polytope*!
- Follows from theory of *moment polytopes*.
- [BFGOWW 18]: Efficient algorithms via *non-uniform tensor scaling*.

Underlying group action: Products of GL's on *tensors*.

More examples

- Newton: $q \in \boldsymbol{C}[x_1, \dots, x_n]$, $q = \sum_{\alpha \in \Omega} c_\alpha x^\alpha$ homogeneous polynomial.
$$P_q = \text{conv}\{\alpha : \alpha \in \Omega\} \subseteq \boldsymbol{R}^n$$
- Schur-Horn: A $n \times n$ symmetric matrix.
$$P_A = \{\text{diag}(B) : B \text{ similar to } A\} \subseteq \boldsymbol{R}^n$$
- Horn:
$$P = \{(\lambda_A, \lambda_B, \lambda_C) : A + B = C\} \subseteq \boldsymbol{R}^{3n}$$
- Edmonds: M, M' matroids on $[n]$.
$$P_{M,M'} = \text{conv}\{1_S : S \text{ basis for } M, M'\} \subseteq \boldsymbol{R}^n$$