Algorithms for the Separation of Orbit Closures of Matrices (arXiv:1801.02043)

Harm Derksen (University of Michigan)
joint work with Visu Makam (IAS)

SIAM conference on
Applied Algebraic Geometry
July 12, 2019

Invariant Theory

K algebraically closed base field
G reductive algebraic group over K
e.g., GL_n, SL_n, O_n, finite, or products of these

Invariant Theory

K algebraically closed base field
G reductive algebraic group over K
e.g., GL_n, SL_n, O_n, finite, or products of these

V n-dimensional representation of G
$K[V]$ ring of polynomial functions on V
G acts by polynomial automorphisms on $K[V]$

Invariant Theory

K algebraically closed base field
G reductive algebraic group over K
e.g., GL_n, SL_n, O_n, finite, or products of these

V n-dimensional representation of G
$K[V]$ ring of polynomial functions on V
G acts by polynomial automorphisms on $K[V]$

Definition

invariant ring $K[V]^G = \{f \in K[V] \mid \forall g \in G\, g \cdot f = f\}$
$= \{f \in K[V] \mid f \text{ constant on } G\text{-orbits}\}$.

Invariant Theory

K algebraically closed base field
G reductive algebraic group over K
e.g., GL_n, SL_n, O_n, finite, or products of these

V n-dimensional representation of G
$K[V]$ ring of polynomial functions on V
G acts by polynomial automorphisms on $K[V]$

Definition

invariant ring $K[V]^G = \{f \in K[V] \mid \forall g \in G\; g \cdot f = f\}$
$= \{f \in K[V] \mid f \text{ constant on } G\text{-orbits}\}$.

Theorem (Hilbert, Nagata, Haboush)

$K[V]^G$ is a finitely generated K-algebra

Geometry of Orbits

Definition

an invariant $f \in K[V]^G$ *separates* $v, w \in V$ if $f(v) \neq f(w)$

Geometry of Orbits

Definition

an invariant $f \in K[V]^G$ *separates* $v, w \in V$ if $f(v) \neq f(w)$

$\overline{G \cdot v}$ is Zariski closure of orbit $G \cdot v$.

Proposition

$\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset \Leftrightarrow f(v) \neq f(w)$ for some $f \in K[V]^G$

$\Leftarrow$ is easy to see

Geometry of Orbits

Definition

an invariant $f \in K[V]^G$ *separates* $v, w \in V$ if $f(v) \neq f(w)$

$\overline{G \cdot v}$ is Zariski closure of orbit $G \cdot v$.

Proposition

$\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset \Leftrightarrow f(v) \neq f(w)$ for some $f \in K[V]^G$

$\Leftarrow$ is easy to see

Orbit Closure Problem

given $v, w \in W$ determine whether $\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset$
if so, find explicit $f \in K[V]^G$ with $f(v) \neq f(w)$

Geometry of Orbits

Definition

an invariant $f \in K[V]^G$ *separates* $v, w \in V$ if $f(v) \neq f(w)$

$\overline{G \cdot v}$ is Zariski closure of orbit $G \cdot v$.

Proposition

$\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset \Leftrightarrow f(v) \neq f(w)$ for some $f \in K[V]^G$

$\Leftarrow$ is easy to see

Orbit Closure Problem

given $v, w \in W$ determine whether $\overline{G \cdot v} \cap \overline{G \cdot w} = \emptyset$
if so, find explicit $f \in K[V]^G$ with $f(v) \neq f(w)$

$\mathcal{N} := \{v \in V \mid 0 \in \overline{G \cdot v}\}$ Null cone
$v \in \mathcal{N} \Leftrightarrow \overline{G \cdot v} \cap \overline{G \cdot 0} \neq \emptyset \Leftrightarrow \forall f \in K[V]^G, f(v) = f(0)$

Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}$ $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by conjugation: $g \cdot A = gAg^{-1}$

Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}$ $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by conjugation: $g \cdot A = gAg^{-1}$

characteristic polynomial of $A \in \mathrm{Mat}_{n,n}$:
$\chi_A(t) := \det(tI - A) = t^n + f_1(A)t^{n-1} + \cdots + f_n(A)$
$K[V]^G = K[f_1, f_2, \ldots, f_n]$

Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}$ $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by conjugation: $g \cdot A = gAg^{-1}$

characteristic polynomial of $A \in \mathrm{Mat}_{n,n}$:
$\chi_A(t) := \det(tI - A) = t^n + f_1(A)t^{n-1} + \cdots + f_n(A)$
$K[V]^G = K[f_1, f_2, \ldots, f_n]$

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \chi_A(t) = \chi_B(t)$

Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}$ $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by conjugation: $g \cdot A = gAg^{-1}$

characteristic polynomial of $A \in \mathrm{Mat}_{n,n}$:
$\chi_A(t) := \det(tI - A) = t^n + f_1(A)t^{n-1} + \cdots + f_n(A)$
$K[V]^G = K[f_1, f_2, \ldots, f_n]$

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \chi_A(t) = \chi_B(t)$

$A \in \mathcal{N} \Leftrightarrow f_1(A) = \cdots = f_n(A) = 0 \Leftrightarrow \chi_A(t) = t^n \Leftrightarrow A$ is nilpotent

Simultaneous Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by simultaneous conjugation:
$g \cdot (A_1, \dots, A_m) = (gA_1g^{-1}, \dots, gA_mg^{-1})$

Simultaneous Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by simultaneous conjugation:
$g \cdot (A_1, \dots, A_m) = (gA_1 g^{-1}, \dots, gA_m g^{-1})$

for a word $w = w_1 w_2 \cdots w_r$ with $w_1, \dots, w_r \in \{1, 2, \dots, m\}$ define
$A_w = A_{w_1} A_{w_2} \cdots A_{w_r}$
the length $\ell(w)$ of w is r

Simultaneous Matrix Conjugation

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by simultaneous conjugation:
$g \cdot (A_1, \dots, A_m) = (gA_1g^{-1}, \dots, gA_mg^{-1})$

for a word $w = w_1w_2\cdots w_r$ with $w_1, \dots, w_r \in \{1, 2, \dots, m\}$ define
$A_w = A_{w_1}A_{w_2}\cdots A_{w_r}$
the length $\ell(w)$ of w is r

Theorem (Procesi, Razmyslov, char$(K) = 0$)

*$K[V]^G$ generated by all $A = (A_1, \dots, A_m) \mapsto \mathrm{Trace}(A_w)$
for all w of length $\leq n^2$*

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$G = \mathrm{GL}_n$ acts on V by simultaneous conjugation:
$g \cdot (A_1, \dots, A_m) = (gA_1g^{-1}, \dots, gA_mg^{-1})$

for a word $w = w_1w_2 \cdots w_r$ with $w_1, \dots, w_r \in \{1, 2, \dots, m\}$ define
$A_w = A_{w_1}A_{w_2} \cdots A_{w_r}$
the length $\ell(w)$ of w is r

Theorem (Procesi, Razmyslov, $\mathrm{char}(K) = 0$)

$K[V]^G$ generated by all $A = (A_1, \dots, A_m) \mapsto \mathrm{Trace}(A_w)$ for all w of length $\leq n^2$

Theorem (Donkin, $\mathrm{char}(K)$ arbitrary)

$K[V]^G$ generated by all coefficients of $\chi_{A_w}(t)$ for all w

D.-Makam: only need w with $\ell(w) \leq (m+1)n^4$

Simultaneous Matrix Conjugation

Algorithm

Forbes and Shpilka (2013) gave a (parallel) polynomial time algorithm for the orbit closure problem if $\mathrm{char}(K) = 0$ but algorithm does not explicitly construct a separating invariant if orbit closures are disjoint

Simultaneous Matrix Conjugation

Algorithm

Forbes and Shpilka (2013) gave a (parallel) polynomial time algorithm for the orbit closure problem if $\mathrm{char}(K) = 0$ but algorithm does not explicitly construct a separating invariant if orbit closures are disjoint

Algorithm

D. and Makam (2018) gave a polynomial time algorithm for orbit closure problem in arbitary characteristic that also explicitly constructs a separating invariant when orbit closures are disjoint

Orbit Closures for Simultaneous Conjugation

given $A = (A_1, \dots, A_m), B = (B_1, \dots, B_m) \in V = \mathrm{Mat}_{n,n}^m$

define $C_i = \left(\begin{array}{c|c} A_i & 0 \\ \hline 0 & B_i \end{array}\right)$, $i = 1, 2, \dots, m$

Orbit Closures for Simultaneous Conjugation

given $A = (A_1, \dots, A_m), B = (B_1, \dots, B_m) \in V = \mathrm{Mat}_{n,n}^m$

define $C_i = \left(\begin{array}{c|c} A_i & 0 \\ \hline 0 & B_i \end{array}\right)$, $i = 1, 2, \dots, m$

$\mathcal{C} = K\langle C_1, \dots, C_m \rangle = \mathrm{Span}\{C_w \mid w \text{ word}\}$

Orbit Closures for Simultaneous Conjugation

given $A = (A_1, \dots, A_m), B = (B_1, \dots, B_m) \in V = \mathrm{Mat}_{n,n}^m$

define $C_i = \left(\begin{array}{c|c} A_i & 0 \\ \hline 0 & B_i \end{array}\right)$, $i = 1, 2, \dots, m$

$\mathcal{C} = K\langle C_1, \dots, C_m \rangle = \mathrm{Span}\{C_w \mid w \text{ word}\}$

order all words lexicographically

$\emptyset, 1, 2, \dots, m, 11, 12, \dots, 1m, 21, \dots, 2m, \dots, 111, 112, \dots$

Definition

w is called a pivot if $C_w \notin \mathrm{Span}\{C_u \mid u < w\}$

Lemma

$\{C_w \mid w$ *is a pivot*$\}$ *is basis of* $\mathcal{C}$

Orbit Closures for Simultaneous Conjugation

Lemma

every subword of a pivot is also a pivot

so # of pivots is at most $\dim \mathcal{C} \leq 2n^2$
largest pivot has length $< 2n^2$ (actually $O(n \log(n))$ by Shitov)

Lemma

every subword of a pivot is also a pivot

so # of pivots is at most $\dim \mathcal{C} \leq 2n^2$
largest pivot has length $< 2n^2$ (actually $O(n\log(n))$ by Shitov)

suppose we found all pivots of length d
to find pivots of length $d+1$ we only have to check all words wi where w is a pivot of length d and $1 \leq i \leq m$

we can find all pivots in polynomial time

Orbit Closures for Simultaneous Conjugation

Theorem (char$(K) = 0$)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ *for all pivots* w

Theorem (char$(K) = 0$)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ *for all pivots* w

Proof: $\Rightarrow$ clear, $\Leftarrow$:

$$\mathcal{C} \subseteq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right) \,\middle|\, \mathrm{Trace}(A) = \mathrm{Trace}(B) \right\}$$

Orbit Closures for Simultaneous Conjugation

Theorem (char$(K) = 0$)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ *for all pivots* w

Proof: $\Rightarrow$ clear, $\Leftarrow$:

$$\mathcal{C} \subseteq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right) \middle| \, \mathrm{Trace}(A) = \mathrm{Trace}(B) \right\}$$

so $\mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ for all words w

Theorem (char$(K) = 0$)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ *for all pivots* w

Proof: $\Rightarrow$ clear, $\Leftarrow$:

$$\mathcal{C} \subseteq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right) \middle| \mathrm{Trace}(A) = \mathrm{Trace}(B) \right\}$$

so $\mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ for all words w

by Procesi's Theorem $\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset$ $\square$

Theorem (char$(K) = 0$)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ *for all pivots* w

Proof: $\Rightarrow$ clear, $\Leftarrow$:

$$\mathcal{C} \subseteq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right) \,\middle|\, \mathrm{Trace}(A) = \mathrm{Trace}(B) \right\}$$

so $\mathrm{Trace}(A_w) = \mathrm{Trace}(B_w)$ for all words w

by Procesi's Theorem $\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset$ $\square$

Using Donkin's theorem one gets (with more effort):

Theorem (char(K) arbitrary)

$\overline{G \cdot A} \cap \overline{G \cdot B} \neq \emptyset \Leftrightarrow \chi_{A_w}(t) = \chi_{B_w}(t)$ *for all pivots* w

Simultaneous Left-Right Action

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$H = \mathrm{SL}_n \times \mathrm{SL}_n$ acts on V by simultaneous left-right action:
$(g, h) \cdot (A_1, \dots, A_m) = (gA_1h^{-1}, \dots, gA_mh^{-1})$

Simultaneous Left-Right Action

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$H = \mathrm{SL}_n \times \mathrm{SL}_n$ acts on V by simultaneous left-right action:
$(g, h) \cdot (A_1, \dots, A_m) = (gA_1 h^{-1}, \dots, gA_m h^{-1})$

Theorem (D. and Makam)

$K[V]^H$ generated by all $A = (A_1, \dots, A_m) \mapsto \det(\sum_{i=1}^m A_i \otimes T_i)$ where $T = (T_1, \dots, T_m) \in \mathrm{Mat}_{d,d}^m$ and $d < mn^3$

Simultaneous Left-Right Action

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$H = \mathrm{SL}_n \times \mathrm{SL}_n$ acts on V by simultaneous left-right action:
$(g,h) \cdot (A_1, \dots, A_m) = (gA_1h^{-1}, \dots, gA_mh^{-1})$

Theorem (D. and Makam)

$K[V]^H$ generated by all $A = (A_1, \dots, A_m) \mapsto \det(\sum_{i=1}^m A_i \otimes T_i)$ where $T = (T_1, \dots, T_m) \in \mathrm{Mat}_{d,d}^m$ and $d < mn^3$

for $T = (T_1, \dots, T_m) \in \mathrm{Mat}_{d,d}^m$, define $f_T \in K[V]^H$ by
$f_T(A) = \det(\sum_{i=1}^m A_i \otimes T_i)$

Simultaneous Left-Right Action

Example: $V = \mathrm{Mat}_{n,n}^m$ m-tuples $n \times n$ matrices
$H = \mathrm{SL}_n \times \mathrm{SL}_n$ acts on V by simultaneous left-right action:
$(g, h) \cdot (A_1, \dots, A_m) = (gA_1h^{-1}, \dots, gA_mh^{-1})$

Theorem (D. and Makam)

$K[V]^H$ generated by all $A = (A_1, \dots, A_m) \mapsto \det(\sum_{i=1}^m A_i \otimes T_i)$ where $T = (T_1, \dots, T_m) \in \mathrm{Mat}_{d,d}^m$ and $d < mn^3$

for $T = (T_1, \dots, T_m) \in \mathrm{Mat}_{d,d}^m$, define $f_T \in K[V]^H$ by
$f_T(A) = \det(\sum_{i=1}^m A_i \otimes T_i)$

Garg-Gurvitz-Oliviera-Wigderson, Ivanyos-Qiao-Subrahmanyan

there is polynomial time algorithm for deciding whether $A = (A_1, \dots, A_m) \in \mathcal{N}$ and algorithm constructs $T \in \mathrm{Mat}_{n,n}^m$ with $f_T(A) \neq 0$ if $A \notin \mathcal{N}$

Orbit Closure Separation Algorithm for Left-Right Action

$(H = \mathrm{SL}_n \times \mathrm{SL}_n,\ G = \mathrm{GL}_n)$
suppose $A = (A_1, \dots, A_m), B = (B_1, \dots, B_m) \in \mathrm{Mat}_{n,n}^m$ are given

Orbit Closure Separation Algorithm for Left-Right Action

$(H = \mathrm{SL}_n \times \mathrm{SL}_n,\ G = \mathrm{GL}_n)$
suppose $A = (A_1, \ldots, A_m), B = (B_1, \ldots, B_m) \in \mathrm{Mat}_{n,n}^m$ are given

if $A, B \in \mathcal{N}$ then $0 \in \overline{H \cdot A} \cap \overline{H \cdot B} \neq \emptyset$

Orbit Closure Separation Algorithm for Left-Right Action

$(H = \mathsf{SL}_n \times \mathsf{SL}_n,\ G = \mathsf{GL}_n)$
suppose $A = (A_1, \ldots, A_m), B = (B_1, \ldots, B_m) \in \mathsf{Mat}_{n,n}^m$ are given

if $A, B \in \mathcal{N}$ then $0 \in \overline{H \cdot A} \cap \overline{H \cdot B} \neq \emptyset$

suppose $A \notin \mathcal{N}$
we find $T \in \mathsf{Mat}_{n,n}$ with $f_T(A) \neq 0$
if $f_T(A) \neq f_T(B)$ then $\overline{G \cdot A} \cap \overline{G \cdot B} = \emptyset$

Orbit Closure Separation Algorithm for Left-Right Action

$(H = \mathrm{SL}_n \times \mathrm{SL}_n,\ G = \mathrm{GL}_n)$
suppose $A = (A_1, \ldots, A_m), B = (B_1, \ldots, B_m) \in \mathrm{Mat}_{n,n}^m$ are given

if $A, B \in \mathcal{N}$ then $0 \in \overline{H \cdot A} \cap \overline{H \cdot B} \neq \emptyset$

suppose $A \notin \mathcal{N}$
we find $T \in \mathrm{Mat}_{n,n}$ with $f_T(A) \neq 0$
if $f_T(A) \neq f_T(B)$ then $\overline{G \cdot A} \cap \overline{G \cdot B} = \emptyset$

suppose $f_T(A) = f_T(B) \neq 0$
(using T) we define a polynomial map $\zeta : \mathrm{Mat}_{n,n}^m \to \mathrm{Mat}_{n,n}^{mn^2}$ of degree n^2 with the property

$$\overline{H \cdot A} \cap \overline{H \cdot B} = \emptyset \Leftrightarrow \overline{G \cdot \zeta(A)} \cap \overline{G \cdot \zeta(B)} = \emptyset$$

Orbit Closure Separation Algorithm for Left-Right Action

$(H = \mathrm{SL}_n \times \mathrm{SL}_n,\ G = \mathrm{GL}_n)$
suppose $A = (A_1, \dots, A_m), B = (B_1, \dots, B_m) \in \mathrm{Mat}_{n,n}^m$ are given

if $A, B \in \mathcal{N}$ then $0 \in \overline{H \cdot A} \cap \overline{H \cdot B} \neq \emptyset$

suppose $A \notin \mathcal{N}$
we find $T \in \mathrm{Mat}_{n,n}$ with $f_T(A) \neq 0$
if $f_T(A) \neq f_T(B)$ then $\overline{G \cdot A} \cap \overline{G \cdot B} = \emptyset$

suppose $f_T(A) = f_T(B) \neq 0$
(using T) we define a polynomial map $\zeta : \mathrm{Mat}_{n,n}^m \to \mathrm{Mat}_{n,n}^{mn^2}$ of degree n^2 with the property

$$\overline{H \cdot A} \cap \overline{H \cdot B} = \emptyset \Leftrightarrow \overline{G \cdot \zeta(A)} \cap \overline{G \cdot \zeta(B)} = \emptyset$$

we reduced the problem to simultaneous conjugation!