Zeyuan Allen-Zhu
Ankit Garg
Yuanzhi Li
Rafael Oliveira
Avi Wigderson

Geodesically Convex Optimization & Applications to Operator Scaling and Invariant Theory

Contents

- 2nd order methods for Matrix Scaling
- Geodesic Convexity
- Operator Scaling – Setup & Algorithm
- Application: Orbit Closure Intersection

Recap - Non-Negative Matrices & Scaling

$\boldsymbol{X} \in \boldsymbol{M}_n(\mathbb{R}_{\geq 0})$ is **doubly stochastic (DS)** if row/column sums of $\boldsymbol{X}$ are equal to **1**.

Y is **scaling** of X if $\exists$ positive $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ s.t. $y_{ij} = \alpha_i x_{ij} \beta_j$.

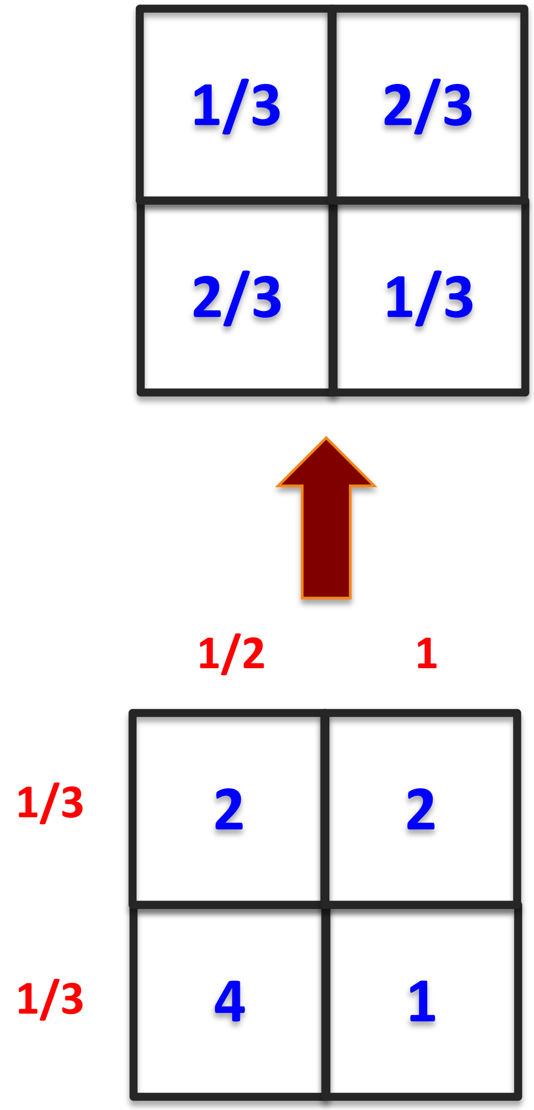
$\boldsymbol{X}$ has DS scaling if $\exists$ scaling $\mathbf{Y}$ of $\boldsymbol{X}$ s.t. all row/column sums of $\boldsymbol{Y}$ equal **1**.

$$\boldsymbol{ds}(\boldsymbol{A}) = \sum_{\boldsymbol{i}} (\boldsymbol{r_i} - \mathbf{1})^2 + \sum_{\boldsymbol{j}} (\boldsymbol{c_j} - \mathbf{1})^2$$

$\boldsymbol{A}$ has approx. DS scaling if $\forall \boldsymbol{\epsilon} > \mathbf{0}$ there is scaling $\boldsymbol{B_\epsilon}$ of $\boldsymbol{A}$ s.t. $\mathbf{ds}(\mathbf{B}_\epsilon) < \boldsymbol{\epsilon}$.

1. When does $\boldsymbol{X}$ have approx. DS scaling?
2. Can we find it efficiently?

Has **convex** formulation!



A Convex Formulation

$\mathbf{A} \in \boldsymbol{M_n}(\mathbb{R}_{\geq \mathbf{0}})$ input matrix.

$$f(x) = \sum_{1 \leq i \leq n} log\left(\sum_j A_{ij} e^{x_j}\right) - \sum_j x_j$$

Side Note: $f(x)$ is logarithm of **[GY'98]** capacity for matrix scaling

$\boldsymbol{A}$ has DS scaling iff

$$inf\{f(x) : x > \mathbf{0}\} > -\infty$$

How can we solve (really fast) optimization problem above?

- $\nabla^2 f(x)$ **not bounded spectral norm –** bad for 1st order methods
- $f(x)$ **not self-concordant –** cannot apply std 2nd order methods
- But $f(x)$ **"self-robust"** – still hope for some 2nd order methods

Self Concordance & Self Robustness

Self concordance: $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}$ is self concordant if

$$|f'''(x)| \leq 2\left(f''(x)\right)^{3/2}$$

$f : \mathbb{R}^n \to \mathbb{R}$ self concordant if self concordant along each line.
"well-approximated" by quadratic function around every pt.

Unfortunately, log of capacity **NOT** self-concordant.

Self robustness [CMTV'18, ALOW'18]: $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}$ is self robust if

$$|f'''(x)| \leq 2 \cdot f''(x)$$

$f : \mathbb{R}^n \to \mathbb{R}$ self robust if self robust along each line.
"well approximated" by quadratic on small nbhd around each pt.

Log of capacity is self-robust!

Question: Can we efficiently optimize self-robust functions?

Answer: Yes! Perform "box-constrained Newton Method"

Essentially: optimize "quadratic approx" of fncn on small nbhd

Properties of Self Robustness

Self robustness [CMTV'18, ALOW'18]: $\boldsymbol{f}: \mathbb{R} \rightarrow \mathbb{R}$ is self robust if

$$|\boldsymbol{f}'''(\boldsymbol{x})| \leq \mathbf{2} \cdot \boldsymbol{f}''(\boldsymbol{x})$$

$f: \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}$ self robust if self robust along each line.

"well approximated" by quadratic on <u>small nbhd</u> around each pt.

More formally: $\boldsymbol{f}: \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}$ self robust, $\boldsymbol{x}, \boldsymbol{\delta} \in \mathbb{R}^{\boldsymbol{n}}$ s.t. $||\boldsymbol{\delta}||_\infty \leq \mathbf{1}$

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{\delta}) \leq \boldsymbol{f}(\boldsymbol{x}) + \langle \nabla \mathbf{f}(\mathbf{x}), \boldsymbol{\delta} \rangle + \boldsymbol{\delta}^{T} \nabla^2 \mathbf{f}(\mathbf{x}) \boldsymbol{\delta}$$

$$\boldsymbol{f}(\boldsymbol{x}) + \langle \nabla \mathbf{f}(\mathbf{x}), \boldsymbol{\delta} \rangle + \frac{\mathbf{1}}{\mathbf{6}} \boldsymbol{\delta}^{T} \nabla^2 \mathbf{f}(\mathbf{x}) \boldsymbol{\delta} \leq \boldsymbol{f}(\boldsymbol{x}+\boldsymbol{\delta})$$

Idea: iteratively solve minimization problem

$$\boldsymbol{min}_{||\delta||_\infty \leq 1} \langle \nabla \boldsymbol{f}(\boldsymbol{x_t}), \boldsymbol{\delta} \rangle + \boldsymbol{\delta}^{T} \nabla^2 \boldsymbol{f}(\boldsymbol{x_t}) \boldsymbol{\delta}$$

Then update $\boldsymbol{x_{t+1}} \leftarrow \boldsymbol{x_t} + \boldsymbol{\delta}$.

$$\boldsymbol{f}(\boldsymbol{x_{t+1}}) - \boldsymbol{f}(\boldsymbol{x}^*) \leq (\mathbf{1} - \mathbf{1}/||\boldsymbol{x_t} - \boldsymbol{x}^*||_\infty)(\boldsymbol{f}(\boldsymbol{x_t}) - \boldsymbol{f}(\boldsymbol{x}^*))$$

(Kind of) Faster Algorithm & Analysis

Algorithm [ALOW'17, CMTV'17]

- Start with $\boldsymbol{x_0 = 1}$, $\ell = \boldsymbol{O}(\boldsymbol{R} \cdot \boldsymbol{log}(\mathbf{1}/\boldsymbol{\epsilon}))$.
- For $\boldsymbol{t} = \mathbf{0}$ to $\ell - \mathbf{1}$
  - ➢ $\boldsymbol{f}^{(t)}(\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x_t} + \boldsymbol{y})$.
    - ➢ $\boldsymbol{q_t}$ *quadratic-approximation* to $\boldsymbol{f}^{(t)}$.
  - ➢ $\boldsymbol{y_t} = \mathrm{argmin}_{||y||_\infty \le 1}\, \boldsymbol{q_t}(\boldsymbol{y})$.
  - ➢ $\boldsymbol{x_{t+1}} = \boldsymbol{x_t} + \boldsymbol{y_t}$.
- Return $\boldsymbol{x_\ell}$.

Analysis:

1. There is approx. minimizer $\boldsymbol{x}^* \in \boldsymbol{B}_\infty(\mathbf{0}, \boldsymbol{R})$ (add regularizer)
2. Each step gets us $\times(\mathbf{1} - \mathbf{1}/\boldsymbol{R})$ closer to OPT
3. After $\boldsymbol{Rlog}(\mathbf{1}/\boldsymbol{\epsilon})$ iterations $\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}^*) \le \boldsymbol{\epsilon}$
4. This $\boldsymbol{x}$ gives us $\boldsymbol{\epsilon}$-approximate scaling

Getting scaling from minimizer

$\mathbf{A} \in M_n(\mathbb{R}_{\geq 0})$ input matrix.

$$f(x) = \sum_{1\leq i\leq n} log\left(\sum_j A_{ij}e^{x_j}\right) - \sum_j x_j$$

Let

$$(A_x)_{ik} = \frac{A_{ik}e^{x_k}}{\sum_j A_{ij}e^{x_j}}$$

Claim: $||\nabla f(z)||_2^2 = ds(A_z)$

If z s.t. $f(z) \leq inf_{x>0} f(x) + \epsilon$ and $||\nabla f(z)||_2^2 \leq \epsilon$ thus

$$ds(A_z) \leq \epsilon$$

Thus ϵ-close to DS.

Quantum Operators – Definition

A **completely positive operator** is any map $\mathbf{T}\colon \mathbf{M}_n(\mathbb{C}) \to \boldsymbol{M}_n(\mathbb{C})$ given by $(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m)$ s.t.

$$\boldsymbol{T}(\boldsymbol{X}) = \sum_{1 \le i \le m} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{A}_i^{\dagger}$$

Such maps take psd matrices to psd matrices.

Dual of $\mathbf{T}(\mathbf{X})$ is map $\mathbf{T}^*\colon \mathbf{M}_n(\mathbb{C}) \to \boldsymbol{M}_n(\mathbb{C})$ given by:

$$\boldsymbol{T}^*(\boldsymbol{X}) = \sum_{1 \le i \le m} \boldsymbol{A}_i^{\dagger} \boldsymbol{X} \boldsymbol{A}_i$$

- Analog of scaling?
- Doubly stochastic?

Operator Scaling

A quantum operator $\mathbf{T}\colon \mathbf{M}_{n}(\mathbb{C}) \to M_n(\mathbb{C})$ is **doubly stochastic (DS)** if $T(I) = T^*(I) = I$.

Scaling of $T(X)$ consists of $L, R \in GL_n(\mathbb{C})$ s.t.

$$(A_1, \dots, A_m) \to (LA_1R, \dots, LA_mR)$$

Distance to doubly-stochastic:

$$ds(T) \stackrel{\text{def}}{=} \|T(I) - I\|_F^2 + \|T^*(I) - I\|_F^2$$

$T(X)$ has approx. DS scaling if $\forall \epsilon > 0$, $\exists$ scaling L_ϵ, R_ϵ s.t. operator $T_\epsilon(X)$ given by $(\mathrm{L}_\epsilon A_1 R_\epsilon, \dots, L_\epsilon A_m R_\epsilon)$ has $ds(T_\epsilon) \leq \epsilon$.

1. When does $(A_1, \dots, A_m)$ have approx. DS scaling?
2. Can we find it efficiently?

NO convex formulation!

Previous work

Problem: operator $\mathbf{T} = (\boldsymbol{A_1}, \dots, \boldsymbol{A_m})$, $\boldsymbol{\epsilon} > \mathbf{0}$, can $\boldsymbol{T}$ be $\boldsymbol{\epsilon}$-scaled to double stochastic? If yes, find scaling.

Algorithm G [Gurvits' 04, GGOW'15]:

Repeat $\boldsymbol{k} = \boldsymbol{poly}(\boldsymbol{n}, \mathbf{1}/\boldsymbol{\epsilon})$ times:

1. Left normalize $\boldsymbol{T}(\boldsymbol{X})$, i.e.,$(\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \leftarrow (\boldsymbol{LA_1}, \dots, \boldsymbol{LA_m})$ s.t. $\boldsymbol{T}(\boldsymbol{I}) = \boldsymbol{I}$.
2. Right normalize $\mathbf{T}(\mathbf{X})$, i.e., $(\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \leftarrow (\boldsymbol{A_1 R}, \dots, \boldsymbol{A_m R})$ s.t. $\boldsymbol{T}^*(\boldsymbol{I}) = \boldsymbol{I}$.

If at any point $\mathbf{ds}(\mathbf{T}) \leq \boldsymbol{\epsilon}$, output the current scaling.

Else output **no scaling**.

Potential Function (Capacity) [Gur'04]:

$$\boldsymbol{cap}(\boldsymbol{T}) = \boldsymbol{inf}\left\{\frac{\boldsymbol{det}(\boldsymbol{T}(\boldsymbol{X}))}{\boldsymbol{det}(\boldsymbol{X})} : \boldsymbol{X} \succ \mathbf{0}\right\}.$$

For $\epsilon < 1/n^2$, can scale $\boldsymbol{T}$ to $\boldsymbol{\epsilon}$-close to DS iff $\boldsymbol{cap}(\boldsymbol{T}) > \mathbf{0}$.

Previous work – Analysis

Algorithm G:

Repeat $\boldsymbol{k}$ times:

1. Left normalize: $(A_1, \dots, A_m) \leftarrow (RA_1, \dots, RA_m)$ s.t. $T(I) = I$.
2. Right normalize: $(A_1, \dots, A_m) \leftarrow (A_1C, \dots, A_mC)$ s.t. $T^*(I) = I$.

If at any point $T(X)$ is close to DS, output current scaling.

Else output **no scaling**.

Potential Function (Capacity) [Gur'04]:

$$cap(T) = inf\left\{\frac{det\big(T(X)\big)}{det(X)} : X \succ 0\right\}.$$

Analysis [Gur'04, GGOW'15]:

1. $cap(T) > 0 \Rightarrow cap(T) > e^{-poly(n)}$ **(GGOW'15)**
2. $ds(T) \Rightarrow cap(T)$ grows by $(1 + 1/n)$ after normalization
3. $\mathbf{cap(T)} \leq \mathbf{1}$ for normalized operators.

Previous work – Algorithm G

Potential Function (Capacity) [Gur'04]:

$$\boldsymbol{cap(T) = inf\left\{\frac{det\big(T(X)\big)}{det(X)} :\ X \succ 0\right\}.}$$

For $\epsilon < 1/n^2$, can scale $\boldsymbol{T}$ to $\boldsymbol{\epsilon}$-close to DS iff $\boldsymbol{cap(T) > 0}$.

How can we decide if $\boldsymbol{cap(T) > 0}$? Can we approx. capacity?

[GGOW'15]: natural scaling algorithm decides whether $\boldsymbol{cap(T) > 0}$ in deterministic $\boldsymbol{poly(n)}$ time. Moreover, it finds $\boldsymbol{exp(\epsilon)}$-approx. to capacity in time $\boldsymbol{poly(n, 1/\epsilon)}$.

Can we get convergence in $\boldsymbol{\log\left(\frac{1}{\epsilon}\right)}$?

Need a different algorithm!

Capacity: optimization problem over *Positive Definite* matrices

Is capacity a special function in this manifold?

Geodesic Convexity

Generalizes Euclidean convexity to Riemannian manifolds.

- $\mathbb{R}^n$ becomes a smooth manifold (locally looks like $\mathbb{R}^n$)
- Straight lines become geodesics ("shortest paths")

Example (our setup): complex positive definite matrices $\mathcal{S}_+$ with geodesic from $\boldsymbol{A}$ to $\boldsymbol{B}$ given by:

$$\gamma_{A,B} : [0,1] \to \mathcal{S}_+ \qquad \gamma_{A,B}(t) = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^t A^{1/2}$$

Convexity:

- $K \subseteq \mathcal{S}_+$ g-convex if $\forall A, B \in K$ geodesic from A to B in K
- Function $f : K \to \mathbb{R}$ is g-convex if univariate function

  $f(\gamma_{A,B}(t))$ is convex in t for any $A, B \in K$

Geodesically Convex Functions

Geodesically convex functions over $\mathcal{S}_+$:

- $\mathbf{log}(\mathbf{det}(\boldsymbol{T}(\boldsymbol{X}))$
- $\mathbf{log}(\mathbf{det}(\boldsymbol{X}))$ (geodesically linear)

Thus log of capacity $\stackrel{\text{def}}{=}$ $\mathbf{log}\big(\mathbf{det}\big(\boldsymbol{T}(\boldsymbol{X})\big)\big) - \mathbf{log}(\mathbf{det}(\boldsymbol{X}))$ g-convex!

For $\mathbf{log}(\mathbf{1}/\boldsymbol{\epsilon})$ convergence, need new opt. tools for g-convex fncs.

Known approaches for g-convex functions:

- **[Folklore]** g-self-concordant functions converge in time $\boldsymbol{poly}(\boldsymbol{n} \cdot \boldsymbol{log}(\mathbf{1}/\boldsymbol{\epsilon}))$.

No analog of ellipsoid or interior point method known for this setting.

Self Concordance & Self Robustness

Self concordance: $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}$ is self concordant if

$$|\boldsymbol{f}'''(\boldsymbol{x})| \leq \mathbf{2}\left(\boldsymbol{f}''(\boldsymbol{x})\right)^{\mathbf{3/2}}$$

$f : \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}$ self concordant if self concordant along each line.

$h : \boldsymbol{\mathcal{S}}_{+} \to \mathbb{R}$ g-self concordant if self concordant along each geodesic.

Unfortunately, log of capacity **NOT** self-concordant.

Self robustness: $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}$ is self robust if

$$|\boldsymbol{f}'''(\boldsymbol{x})| \leq \mathbf{2} \cdot \boldsymbol{f}''(\boldsymbol{x})$$

$f : \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}$ self robust if self robust along each line.

$h : \boldsymbol{\mathcal{S}}_{+} \to \mathbb{R}$ g-self robust if self robust along each geodesic.

Log of capacity is self-robust!

Question: Can we efficiently optimize g-self robust functions?

This work – g-convex opt for self-robust fcns

Problem: given $\boldsymbol{f} : \mathcal{S}_+ \to \mathbb{R}$ g-self robust, $\boldsymbol{\epsilon} > \mathbf{0}$, and bound on initial distance $\boldsymbol{R}$ to OPT (diameter) find $\boldsymbol{X_\epsilon} \in \boldsymbol{\mathcal{S}_+}$ such that

$$\boldsymbol{f}(\boldsymbol{X_\epsilon}) \leq \inf_{Y \in \mathcal{S}_+} \boldsymbol{f}(\boldsymbol{Y}) + \boldsymbol{\epsilon}$$

Theorem [AGLOW'18]:

There exists a deterministic $\boldsymbol{poly}(\boldsymbol{n}, \boldsymbol{R}, \boldsymbol{log}(\mathbf{1}/\boldsymbol{\epsilon}))$, algorithm for the problem above.

- Second order method, generalizing recent work of **[ALOW'17, CMTV'17]** for matrix scaling to g-convex setting (Box constrained Newton method)
- Generalizes to other manifolds and metrics

Remark:

- For operator scaling, $\boldsymbol{X_\epsilon}$ also gives us scaling $\boldsymbol{\epsilon}$-close to DS

This paper – g-convex opt for self-robust fcns

Problem: given $\boldsymbol{f} : \mathcal{S}_+ \to \mathbb{R}$ g-self robust, $\boldsymbol{\epsilon} > \mathbf{0}$, and bound on initial distance $\boldsymbol{R}$ to OPT (diameter) find $\boldsymbol{X_\epsilon} \in \boldsymbol{\mathcal{S}_+}$ such that

$$f(X_\epsilon) \leq \inf_{Y \in \mathcal{S}_+} f(Y) + \epsilon$$

Algorithm

- Start with $\boldsymbol{X_0 = I}$, $\boldsymbol{\ell = O(R \cdot log(1/\epsilon))}$.
- For $\boldsymbol{t = 0}$ to $\boldsymbol{\ell - 1}$
  - ➢ $\boldsymbol{f^{(t)}(D) = f(X_t^{1/2} \exp(D) X_t^{1/2})}$.
    - ➢ $\boldsymbol{Q_t}$ *quadratic-approximation* to $\boldsymbol{f^{(t)}}$.
  - ➢ $\boldsymbol{D_t = \mathrm{argmin}_{||D||_F \leq 1} Q_t(D)}$. (*Euclidean convex* opt.)
  - ➢ $\boldsymbol{X_{t+1} = X_t^{1/2} exp(D_t) X_t^{1/2}}$.
- Return $\boldsymbol{X_\ell}$.

- Why would we need this instead of regular scaling?
- What is the bound for $\boldsymbol{R}$ in operator scaling?
  - **[AGLOW'18]** polynomial bound for $\boldsymbol{R}$

Invariant Theory – our setting

Invariant Theory:
$\boldsymbol{G} = \mathbb{SL}_{\boldsymbol{n}}(\mathbb{C})^{\mathbf{2}}$, vector space $\mathbf{V} = \mathbf{M}_{\boldsymbol{n}}(\mathbb{C})^{\mathbf{m}}$ action by L-R mult:

$$(\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \to (\boldsymbol{LA_1R}, \dots, \boldsymbol{LA_mR})$$

Orbit Closure: given $\boldsymbol{v} = (\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \in \boldsymbol{V}$**,** orbit closure is

$$\overline{\mathcal{O}_{\boldsymbol{v}}} = \overline{\{(\boldsymbol{LA_1R}, \dots, \boldsymbol{LA_mR}) \mid (\boldsymbol{L}, \boldsymbol{R}) \in \boldsymbol{G}\}}$$

Orbit Closure Intersection Problem: given two quantum operators $\boldsymbol{u} = (\boldsymbol{A_1}, \dots, \boldsymbol{A_m})$, $\boldsymbol{v} = (\boldsymbol{B_1}, \dots, \boldsymbol{B_m})$**,** is $\overline{\mathcal{O}_{\boldsymbol{u}}} \cap \overline{\mathcal{O}_{\boldsymbol{v}}} \neq \emptyset$?

If $\boldsymbol{v} = \mathbf{0}$ problem becomes the *null-cone problem*.
[GGOW'16]: connections to non-commutative PIT, non-commutative algebra, combinatorics, functional analysis…

How can we solve the orbit intersection problem for L-R action?

Randomized Algorithm

[Mum'65]: alg. structure of orbit closures

- $\overline{\mathcal{O}_{(A_1,\dots,A_m)}} \cap \overline{\mathcal{O}_{(B_1,\dots,B_m)}} = \emptyset$ iff invariant polynomial s.t.

$$\boldsymbol{p}((A_1,\dots,A_m)) \neq \boldsymbol{p}\big((B_1,\dots,B_m)\big)$$

Randomized algorithm:

Given $(A_1,\dots,A_m)$ and $(B_1,\dots,B_m)$, does $\overline{\mathcal{O}_{(A_1,\dots,A_m)}} \cap \overline{\mathcal{O}_{(B_1,\dots,B_m)}} \neq \emptyset$?

1. **[IQS'17, DM'17]:** Invariants of degree $\boldsymbol{n}^6$ suffice
2. Take random invariant polynomial and evaluate it on $(A_1,\dots,A_m)$ and $(B_1,\dots,B_m)$

KN'79 – Duality Theory

[KN'79]:

- Elts of min norm in $\overline{\mathcal{O}_{(A_1,\ldots,A_m)}}$, are DS operators
  - ϵ-close to DS implies ϵ-close to min. norm
- $(B_1, \ldots, B_m)$ and $(C_1, \ldots, C_m)$ elts of min norm in $\overline{\mathcal{O}_{(A_1,\ldots,A_m)}}$ then there exist $\mathrm{U}, \mathrm{V} \in \boldsymbol{SU}(\boldsymbol{n})$ s.t. $C_i = UB_iV$

[AGLOW'18]: solving orbit closure intersection problem. Given $(A_1, \ldots, A_m)$ and $(B_1, \ldots, B_m)$, does $\overline{\mathcal{O}_{(A_1,\ldots,A_m)}} \cap \overline{\mathcal{O}_{(B_1,\ldots,B_m)}} \neq \emptyset$

1. Our g-convex opt finds ***ϵ-approx*** to element of min norm (DS)
2. With elements of min norm, test if they are $SU(n)$-equivalent
   - we give efficient algorithm for testing equivalence

Remarks

Why do we need $\log(1/\epsilon)$ convergence?

- Orbit closures can be exponentially close and not intersect
  - Need to have $\epsilon = \exp(-poly(n))$ approximation
  - **Not** the case for null-cone problem
- $SU(n)$-equivalence algorithm also approximate (and lossy)

Independently, **[DM'18]** solved orbit closure intersection for LR-action in algebraic way.

- Solution also works for fields of positive characteristic
  - Our solution works only over $\mathbb{C}$

Prior to **[AGLOW'18, DM'18]** only *randomized* polynomial time algorithm known for orbit closure intersection (PIT instance).

Open questions

- Efficient algorithms for more classes of g-convex functions?

- Efficient algorithms for null-cone and orbit closure intersection for more general actions?
    - Recent developments for tensor scaling, though still $\boldsymbol{poly(1/\epsilon)}$
    - Upcoming work gets $\boldsymbol{poly(nR \cdot log(1/\epsilon))}$, but still have bad bounds on $\boldsymbol{R}$

- More applications of g-convexity?
    - Recent work **[VY’18]** on Brascamp-Lieb showing it is g-convex

Thank you!