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

Scaling Algorithms for Polynomial Identity Testing

Today

Scaling Problems

Null Cone

Alternating Minimization

Orbit Closure Intersection

Geodesic Convexity

Matrix Balancing

$n \times n$ complex matrix $\boldsymbol{A}$ is **doubly balanced (DB)** if ℓ_2 norm of rows/columns of $\boldsymbol{A}$ are equal.

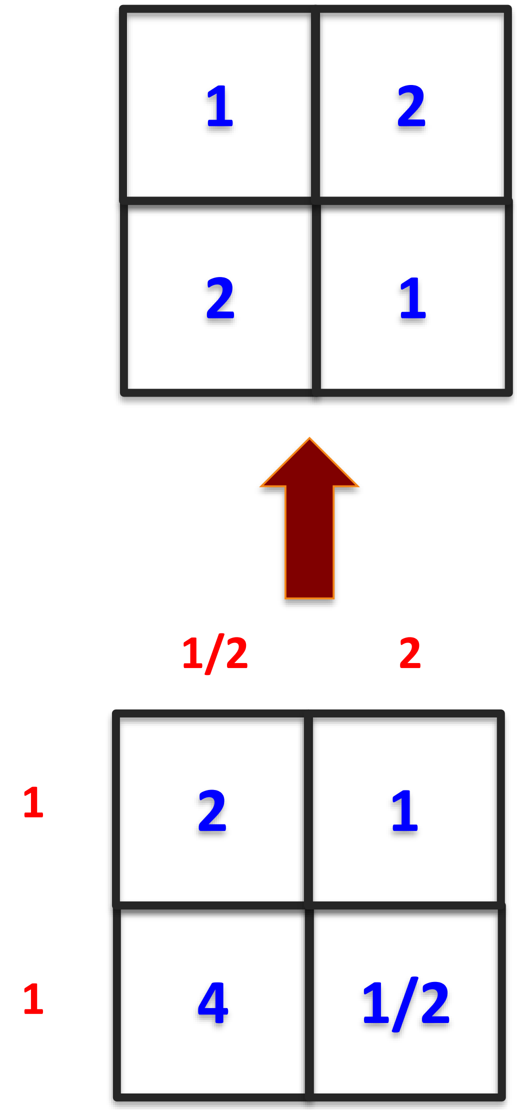
$\boldsymbol{B}$ is **scaling** of $\boldsymbol{A}$ if $\exists$ complex $\boldsymbol{x_1, \ldots, x_n, y_1, \ldots, y_n}$ s.t. $\prod \boldsymbol{x_i} = \prod \boldsymbol{y_j} = \mathbf{1}$ and $\boldsymbol{b_{ij}} = \boldsymbol{x_i a_{ij} y_j}$.

$\boldsymbol{A}$ has DB scaling if $\exists$ scaling $\boldsymbol{B}$ of $\boldsymbol{A}$ s.t. $\boldsymbol{B}$ is DB.

$$\boldsymbol{db}(A) = \sum_i \left(\frac{\boldsymbol{r_i}}{||A||^2} - \frac{\mathbf{1}}{\boldsymbol{n}}\right)^2 + \sum_j \left(\frac{\boldsymbol{c_j}}{||A||^2} - \frac{\mathbf{1}}{\boldsymbol{n}}\right)^2$$

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

1. When does A have approx. DB scaling?
2. Can we find it efficiently?



Matrix Balancing - examples

Column scalings: $\sqrt{2}$, $\left(\sqrt{2}\right)^{-1}$; row scalings: $\sqrt{2}$, $\left(\sqrt{2}\right)^{-1}$

$$\begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix} \Downarrow \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Column scalings: $1/\epsilon$, ϵ; row scalings: ϵ, $1/\epsilon$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \Downarrow \begin{bmatrix} 1 & \epsilon^2 \\ 0 & 1 \end{bmatrix}$$

Matrix Balancing – Algorithm S

Problem: $\boldsymbol{A} \in \boldsymbol{M_n}(\mathbb{C})$, $\boldsymbol{\epsilon > 0}$, is there ϵ-scaling to DB? If yes, find it.

Algorithm S [Kruithof'37, …, Sinkhorn'64]:

Repeat $\boldsymbol{k}$ times:

1. Normalize rows of $\boldsymbol{A}$ (make norm of rows equal)
2. Normalize columns of $\boldsymbol{A}$ (make norm of cols equal)

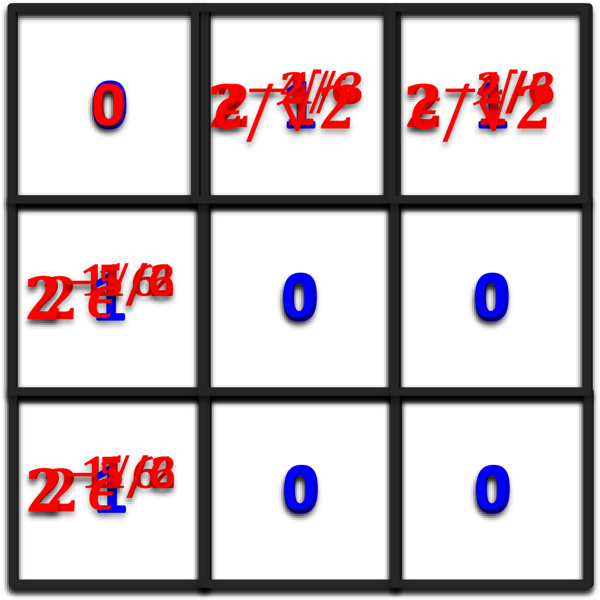
If at any point $\mathbf{db}(\boldsymbol{A}) < \boldsymbol{\epsilon}$, output the scaling.

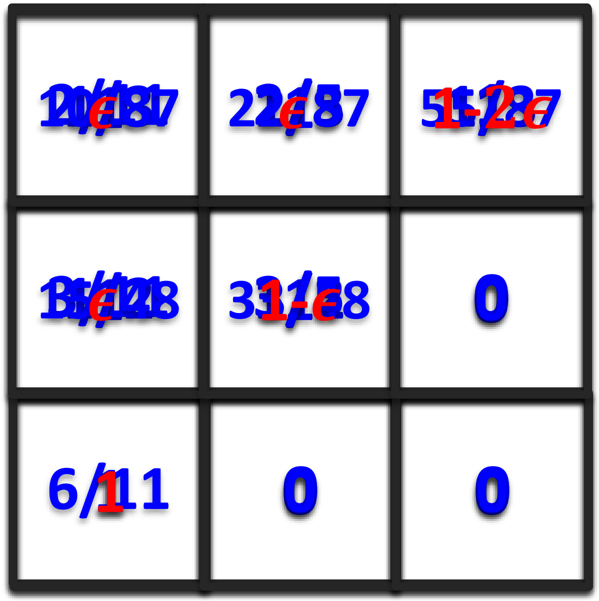
Else, output: **no scaling.**

Questions:

- Are we making progress at all?
- How do we know when to stop? (Which $\boldsymbol{k}$?)
- Is there an $\boldsymbol{\epsilon_0}$ such that if can scale to $\boldsymbol{\epsilon_0}$ then can scale for any $\boldsymbol{\epsilon}$?

Algorithm S – Two Examples





Question: How can we distinguish between these two cases?

Observation: In first example, "huge" block of zeros (Hall blocker). In second, have a matching.

Are these the only cases?

Quantum Operators – Definition

A **quantum operator** is any map $\mathbf{T}: \mathbf{M}_n(\mathbb{C}) \rightarrow \boldsymbol{M}_n(\mathbb{C})$ given by $(\boldsymbol{A}_1, \dots, \boldsymbol{A}_m)$ s.t.

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

Such maps take psd matrices to psd matrices.

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

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

- Analog of scaling?
- Double balanced?

Can scaling solve PIT?

Operator balancing

A quantum operator $\mathbf{T}\colon \mathbf{M}_n(\mathbb{C}) \to M_n(\mathbb{C})$ is **doubly balanced (DB)** if $T(I) = T^*(I) = I$.

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

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

Distance to doubly-balanced:

$$db(T) \stackrel{\text{def}}{=} \left\| \frac{T(I)}{||A||^2} - \frac{1}{n} I \right\|_F^2 + \left\| \frac{T^*(I)}{||A||^2} - \frac{1}{n} I \right\|_F^2$$

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

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

Operator Balancing – Algorithm G

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]:

Repeat $\boldsymbol{k}$ times:

1. Left normalize $\boldsymbol{T}(\boldsymbol{X})$, i.e. make $\boldsymbol{T}(\boldsymbol{I}) \propto \boldsymbol{I}$.
2. Right normalize $\mathbf{T}(\mathbf{X})$, i.e. make $\boldsymbol{T}^*(\boldsymbol{I}) \propto \boldsymbol{I}$.

If at any point $\mathbf{db}(\mathbf{T}) < \boldsymbol{\epsilon}$ output scaling.

Else output **no scaling**.

- Which $\boldsymbol{k}$ should we choose?
- Is there an $\boldsymbol{\epsilon_0}$ such that if can scale to $\boldsymbol{\epsilon_0}$ then can scale for any $\boldsymbol{\epsilon}$?

Analysis – General Approach

Three steps:

1. **[Upper bound]** Potential function $\boldsymbol{\Phi}$ is norm of input.
   - $\boldsymbol{\Phi}$ upper bounded by input size
2. **[Progress/step]** If we are $\boldsymbol{\epsilon}$-far from DB then normalization decreases value of $\boldsymbol{\Phi}$ by $\times\ \boldsymbol{exp(O(\epsilon))}$
3. **[Lower bound]** If there is scaling, *"some property"* tells us that $\boldsymbol{\Phi} \geq \mathbf{exp}(-\boldsymbol{poly(n)})$
   - Bounded away from zero

Approach proves correctness & running time of $\boldsymbol{poly(nb/\epsilon)}$

Analysis – Revisited (matrix scaling)

Three steps:

1. **[Upper bound]** Potential function $\mathbf{\Phi} = ||\mathbf{A}||^2$
   - $\mathbf{\Phi}$ upper bounded by input size
2. **[Progress/step]** If we are $\boldsymbol{\epsilon}$-far from DB then normalization decreases value of $\mathbf{\Phi}$ by $\times\, \boldsymbol{exp}(\boldsymbol{O}(\boldsymbol{\epsilon}))$
3. **[Lower bound]** $\mathbf{A}$ not in null cone, there is "nice" invariant $\boldsymbol{p}(\boldsymbol{Z})$ s.t. $\boldsymbol{p}(\boldsymbol{A}) \neq \mathbf{0}$.
   - $\boldsymbol{p}(\boldsymbol{Z})$ invariant $\Rightarrow$ $\mathbf{p}(\mathbf{A}) = \mathbf{p}(\mathbf{B}),\ \ \forall\, \boldsymbol{B} \in \boldsymbol{G} \cdot \boldsymbol{A}$
   - $\boldsymbol{p}(\boldsymbol{Z})$ integer coeffs. $\Rightarrow |\mathbf{p}(\mathbf{A})|^2 \geq \mathbf{1} \Rightarrow\ ||\boldsymbol{B}|| \geq \boldsymbol{exp}(-\boldsymbol{n})$

Proves correctness & running time of $\boldsymbol{poly}(\boldsymbol{n} \cdot \boldsymbol{log}(\boldsymbol{v})/\boldsymbol{\epsilon})$

Invariants – Matrix Scaling

Matrix Scaling: $ST_n \times ST_n \curvearrowright M_n(\mathbb{C})$

- Matching monomials are invariants:

$$B = XAY \Rightarrow \prod \mathbf{b}_{\mathbf{i}\sigma(i)} = \prod x_i a_{i\sigma(i)} y_{\sigma(i)} = \prod x_i y_i \cdot \prod a_{i\sigma(i)} = \prod a_{i\sigma(i)}$$

- They generate all other invariants
  - If A not in null cone then $p(A) \neq 0$ for some matching
- A integer coeffs. & $\mathbf{p}(\mathbf{A}) \neq \mathbf{0} \Rightarrow |\mathbf{p}(\mathbf{A})|^2 \geq \mathbf{1}$
  - $B \in \overline{G \cdot A} \Rightarrow |p(B)|^2 = |\mathbf{p}(\mathbf{A})|^2 \geq \mathbf{1}$
  - $p(B)$ is a matching monomial $\Rightarrow ||B||^{2n} \geq |p(B)|^2 \geq 1$

Algorithm S – Analysis

Algorithm S: matrix $\boldsymbol{A}$ integer entries bounded by $\boldsymbol{\nu}$, param. $\boldsymbol{\epsilon > 0}$.

Repeat $\boldsymbol{k}$ times:

1. Normalize rows of $\boldsymbol{A}$
2. Normalize columns of $\boldsymbol{A}$

If at any point $\boldsymbol{db(A) \leq \epsilon}$, output the scaling so far.

Else, output: **no scaling.**

Analysis [∼LSW'00]:

1. $||\boldsymbol{A}||^2 \leq \boldsymbol{\nu}^2 \cdot \boldsymbol{n}^2$ (bound on input)
2. $\mathbf{d}\boldsymbol{b}(\mathbf{A}) \geq \boldsymbol{\epsilon} \Rightarrow ||\boldsymbol{A}||^2$ decreases by $\exp(\boldsymbol{O}(\boldsymbol{\epsilon}))$

   after normalization (AM-GM)
3. $||\boldsymbol{B}|| \geq \mathbf{1}$ for any scaling of $\boldsymbol{A}$

Invariants – Operator Scaling

Operator Scaling: $SL_n \times SL_n \curvearrowright M_n(\mathbb{C})^m$

- Invariants: given $B_i = LA_iR$

$$det(\sum B_i \otimes Y_i) = det(\sum (LA_iR) \otimes Y_i)$$

$$= det(\sum A_i \otimes Y_i) \cdot det(L)^d det(R)^d = det(\sum A_i \otimes Y_i)$$

- They generate all other invariants
  - If (A_i) not in null cone then $p(A) \neq 0$ for some such inv.
- A_i, Y_i integer coeffs. & $\mathrm{p(A)} \neq 0 \Rightarrow |\mathrm{p(A)}|^2 \geq 1$
  - $B \in \overline{G \cdot A} \Rightarrow |p(B)|^2 = |\mathrm{p(A)}|^2 \geq 1$

$$\Rightarrow exp(nd) \cdot ||B||^{2nd} \geq |p(B)|^2 \geq 1$$

$$||B|| \geq exp(-n)$$

Algorithm G – Analysis

Algorithm G: tuple (A_i) integer entries bounded by ν, $\epsilon > 0$.

Repeat k times:

1. Left normalize $(A_i) \rightarrow \sum A_i A_i^\dagger \sim I_n$
2. Right normalize $(A_i) \rightarrow \sum A_i^\dagger A_i \sim I_n$

If at any point $db(T) < \epsilon$, output scaling.

Else, output: **no scaling.**

Solved Null-Cone Problem!

Analysis [GGOW'15]:

1. $\sum ||A_i||^2 \leq \nu^2 \cdot n^2$ (bound on input)
2. $\mathrm{d}b(\mathrm{A}) \geq \epsilon \Rightarrow \sum ||A_i||^2$ decreases by $\exp(O(\epsilon))$

   after normalization (AM-GM)
3. $\sum ||B_i||^2 \geq exp(-n)$ for any scaling of A

(Recap) Hilbert's Foundational Results

Given vector space V and group G acting (linearly) on it

$$\textbf{Null cone } \mathcal{N}_G(V) = \{ v \in V \mid 0 \in \overline{G \cdot v} \}$$

[Hil'93] Given vector space V and group G acting (linearly) on it $\mathcal{N}_G(V)$ is the common zero set of **all** invariant polynomials. I.e.

$$v \in \mathcal{N}_G(V) \Leftrightarrow p(v) = 0 \ \ \forall\, p \text{ invariant}$$

Null-cone Problem: given $v \in V$, is $v \in \mathcal{N}_G(V)$?

Two ways of solving this problem!

- Optimization: $\inf_{g \in G}(||g \cdot v||^2)$
- Algebraic: decide if all invariants vanish (“**PIT**”)

Why are we talking about this? Where is DB?

Kempf-Ness & Non-commutative duality

Null-cone Problem: given $\boldsymbol{v} \in \boldsymbol{V}$, is $\boldsymbol{v} \in \mathcal{N}_{\boldsymbol{G}}(\boldsymbol{V})$ (i.e. $\mathbf{0} \in \overline{\boldsymbol{G} \cdot \boldsymbol{v}}$)?

- Optimization: $\boldsymbol{cap}(\boldsymbol{v}) = \inf_{\boldsymbol{g} \in \boldsymbol{G}} (||\boldsymbol{g} \cdot \boldsymbol{v}||^2)$

How do we know we are "close" to the optimum?

- **[KN'79]** "Gradient is close to zero!"
  - Gradient "along the group action" (Lie Algebra)
  - General notion of convexity (geodesic-convexity)

[KN'79] "Non-commutative duality"

- $\boldsymbol{\mu}(\boldsymbol{w})$ moment map: gradient along group action *(Ankit's talk)*
- Dual program: $\boldsymbol{cap}_{\boldsymbol{\mu}}(\boldsymbol{v}) = \inf_{\boldsymbol{g} \in \boldsymbol{G}} ||\boldsymbol{\mu}(\boldsymbol{g} \cdot \boldsymbol{v})||^2$

Far from DB $\boldsymbol{cap}_{\boldsymbol{\mu}}(\boldsymbol{v}) > \mathbf{0} \Leftrightarrow \boldsymbol{cap}(\boldsymbol{v}) = \mathbf{0}$ **In Null cone**

$\boldsymbol{db}(\boldsymbol{A}), \boldsymbol{db}(\boldsymbol{T})$ norms of moment map for matrix/operator scaling!

Algorithm S – Primal dual approach

Algorithm S: matrix $\boldsymbol{A}$ integer entries bounded by $\boldsymbol{\nu}$, param. $\boldsymbol{\epsilon > 0}$.

Repeat $\boldsymbol{k}$ times:

1. Normalize rows of $\boldsymbol{A}$
2. Normalize columns of $\boldsymbol{A}$

If at any point $\boldsymbol{db(A) \leq \epsilon}$, output the scaling so far.

Else, output: **no scaling.**

Far from dual

Analysis [~LSW'00]:

1. $||\boldsymbol{A}||^2 \leq \boldsymbol{\nu}^2 \cdot \boldsymbol{n}^2$ (bound on input)
2. $\mathbf{d}\boldsymbol{b}(\mathrm{A}) \geq \boldsymbol{\epsilon} \Rightarrow ||\boldsymbol{A}||^2$ decreases by $\exp(\boldsymbol{O}(\boldsymbol{\epsilon}))$

   after normalization (AM-GM)

3. $||\boldsymbol{B}|| \geq \mathbf{1}$ for any scaling of $\boldsymbol{A}$

Progress in primal

Invariant Theory – Orbit Closure Intersection

Invariant Theory:

$G = \mathbb{SL}_n(\mathbb{C})^2$, vector space $\mathbf{V} = \mathbf{M}_n(\mathbb{C})^{\mathrm{m}}$ action by L-R mult:

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

Orbit Closure: given $v = (A_1, \ldots, A_m) \in V$, orbit closure is

$$\overline{\mathcal{O}_v} = \overline{\{(LA_1R, \ldots, LA_mR) \mid (L, R) \in G\}}$$

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

If $v = \mathbf{0}$ problem becomes the *null-cone problem*. From **[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?

What do we need to do?

Why is Operator Balancing not enough?

- Orbit closures can be exponentially close and not intersect
  - Need to have $\boldsymbol{\epsilon} = \mathbf{exp}(-\boldsymbol{poly}(\boldsymbol{n}))$ approximation
  - **Not** the case for null-cone problem
- Operator Balancing runs in time $\boldsymbol{poly}(\boldsymbol{n}/\boldsymbol{\epsilon})$
  - Only good for null cone

We need $\mathbf{log}(\mathbf{1}/\boldsymbol{\epsilon})$ convergence!

How to get it? Different algorithm!

KN'79 – Duality Theory

[KN'79]:

- Elts of min norm in $\overline{\mathcal{O}_{(A_1,\ldots,A_m)}}$, are DB operators
  - ϵ-close to DB implies ϵ-close to min. norm
- $(B_1, \ldots, B_m)$ and $(C_1, \ldots, C_m)$ of minimum norm in $\overline{\mathcal{O}_{(A_1,\ldots,A_m)}}$ then equivalent under unitary

[AGLOW'18]: solving orbit closure intersection problem

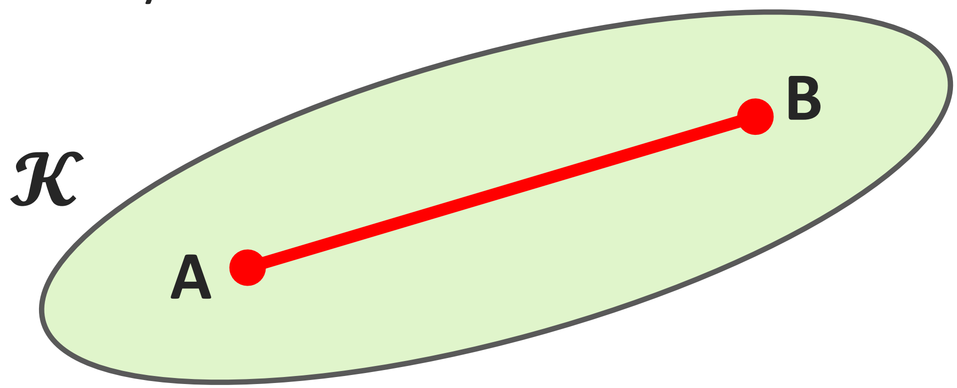
1. g-convex opt finds ***ϵ-approx*** to element of minimum norm (DB)
2. With elements of min norm, test if they are $SU(n)$-equivalent

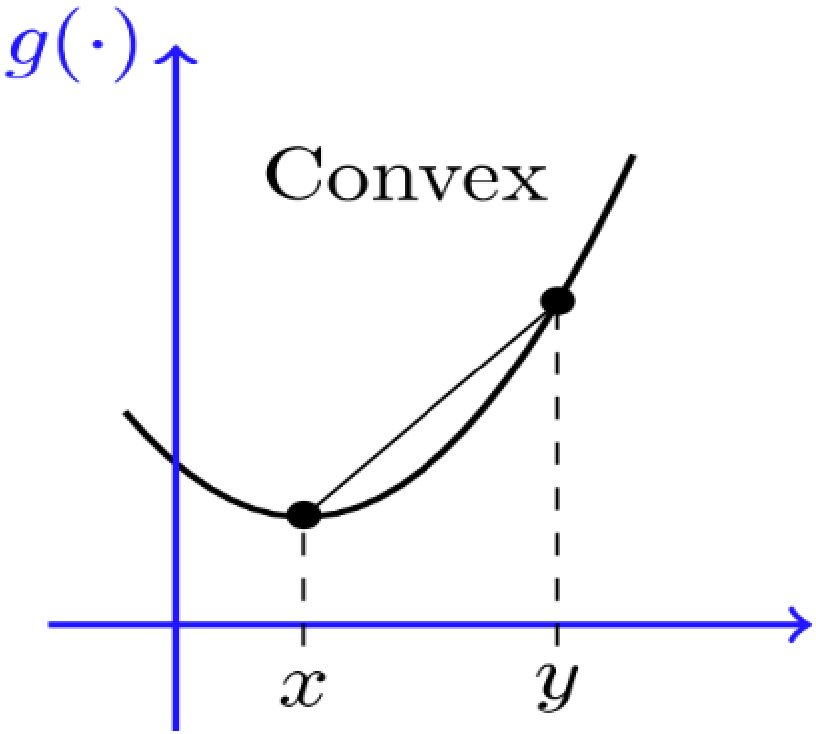
What is this g-convexity?

Convexity

Convexity (Euclidean geometry):

- Shortest path between $\boldsymbol{A}, \boldsymbol{B}$ given by line
- Convex Set $\mathcal{K}$:
- Convex function:





Ellipsoid

Interior Point Methods

2^{nd} order methods

What is Geodesic Convexity?

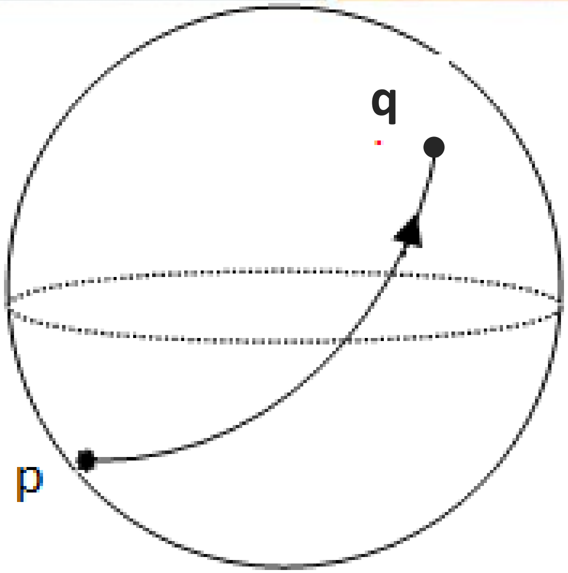
Geodesic Convexity:

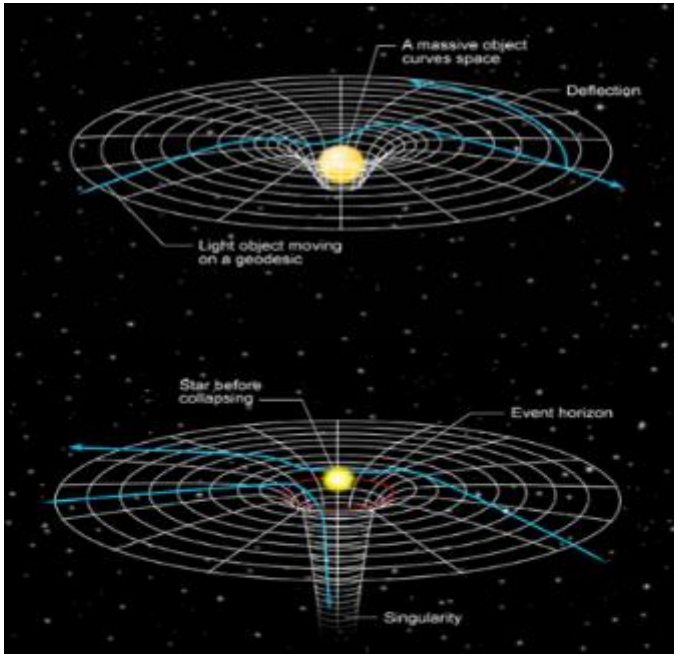
- Shortest path between $\boldsymbol{A}, \boldsymbol{B}$ given by geodesic
- Geodesically Convex Set $\mathcal{K}$:

  $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{K}$ so is its geodesics
- Geodesically Convex function:

  $\boldsymbol{f}$ convex along each geodesic!

Little Known





Geodesic Convexity

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

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

Convexity:

- $\mathbf{K} \subseteq \mathcal{S}_+$ g-convex if $\forall \boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{K}$ geodesic from $\boldsymbol{A}$ to $\boldsymbol{B}$ in $\boldsymbol{K}$
- Function $\boldsymbol{f} : \boldsymbol{K} \to \mathbb{R}$ is g-convex if univariate function $\boldsymbol{f}(\gamma_{A,B}(\boldsymbol{t}))$ is convex in $\boldsymbol{t}$ for any $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{K}$

Geodesically Convex Functions

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

- $\mathbf{log}(||\boldsymbol{g} \cdot \boldsymbol{v}||^2)$

- $\mathbf{log}(\boldsymbol{g} \cdot \boldsymbol{g}^\dagger)$ (geodesically linear)

Log of capacity $\stackrel{\text{def}}{=}$ $\mathbf{log}\left(||\boldsymbol{g} \cdot \boldsymbol{v}||^2\right) - \mathbf{log}(\boldsymbol{g} \cdot \boldsymbol{g}^\dagger)$ g-convex!

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

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 geodesically 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 $\boldsymbol{X}_{\boldsymbol{\epsilon}} \in \boldsymbol{\mathcal{S}}_+$ 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
- Generalizes to other manifolds and metrics

Remark:

- For operator scaling, $\boldsymbol{X}_{\boldsymbol{\epsilon}}$ also gives us scaling $\boldsymbol{\epsilon}$-close to DB

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

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

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

Algorithm

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

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

Remarks & Recap

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)

[AGLOW'18]: solving orbit closure intersection problem

1. g-convex opt finds ***ϵ-approx*** to element of minimum norm (DB)
2. With elements of min norm, test if they are $SU(n)$-equivalent

Advertisement

Amazing workshop at the IAS!
Videos & materials online
https://www.math.ias.edu/ocit2018

Survey on all of this (w/ Ankit) on arxiv & on EATCS complexity column!
(link on my webpage)

Open Questions

- Complexity of null-cone problem? Of OCI?

- Better algorithms for scaling problems?
  - Best algorithms we have are $\boldsymbol{poly(R \cdot log(1/\epsilon))}$

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

Thank you!