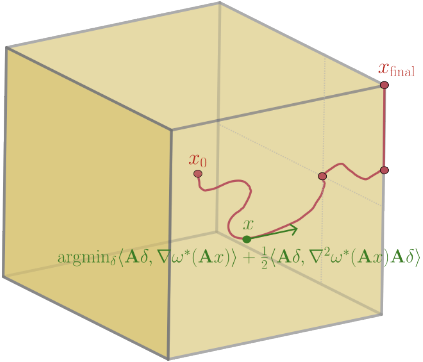
Discrepancy Minimization via Regularization



Lucas Pesenti

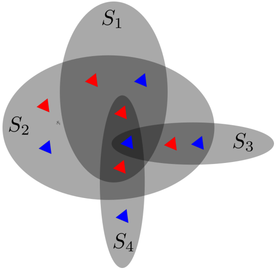
Bocconi University

Adrian Vladu

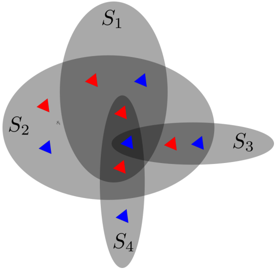
CNRS & Université Paris Cité

Discrepancy Theory

Discrepancy/Vector Balancing problem: Given $u_1, \ldots, u_n \in K \subseteq \mathbb{R}^d$, find $x_1, \ldots, x_n \in \{\pm 1\}$ s.t. the *discrepancy* $\left\| \sum_i x_i u_i \right\|$ is "small"



Discrepancy Theory

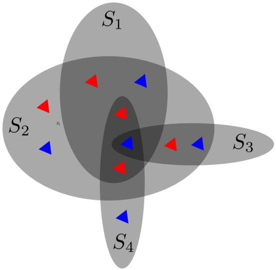


Discrepancy/Vector Balancing problem: Given $u_1, \ldots, u_n \in K \subseteq \mathbb{R}^d$, find $x_1, \ldots, x_n \in \{\pm 1\}$ s.t. the *discrepancy* $\|\sum_i x_i u_i\|$ is "small"

Examples:

- (Spencer's theorem) $K := \{u : \|u\|_\infty \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Komlós conjecture) $K := \{u : \|u\|_2 \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Matrix Spencer conjecture) $K := \{M : \|M\|_{op} \leqslant 1\}$, target $\|\cdot\|_{op}$

Discrepancy Theory



Discrepancy/Vector Balancing problem: Given $u_1, \dots, u_n \in K \subseteq \mathbb{R}^d$, find $x_1, \dots, x_n \in \{\pm 1\}$ s.t. the *discrepancy* $\|\sum_i x_i u_i\|$ is "small"

Examples:

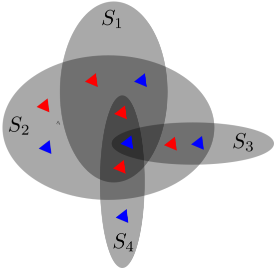
- (Spencer's theorem) $K := \{u : \|u\|_\infty \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Komlós conjecture) $K := \{u : \|u\|_2 \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Matrix Spencer conjecture) $K := \{M : \|M\|_{op} \leqslant 1\}$, target $\|\cdot\|_{op}$

Motivations:

- Prove the existence of rare objects
- Toy problem/building block for "sparsification" tasks

Discrepancy Theory

Discrepancy/Vector Balancing problem: Given $u_1, \ldots, u_n \in K \subseteq \mathbb{R}^d$, find $x_1, \ldots, x_n \in \{\pm 1\}$ s.t. the *discrepancy* $\|\sum_i x_i u_i\|$ is "small"



Examples:

- (Spencer's theorem) $K := \{u : \|u\|_\infty \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Komlós conjecture) $K := \{u : \|u\|_2 \leqslant 1\}$, target $\|\cdot\|_\infty$
- (Matrix Spencer conjecture) $K := \{M : \|M\|_{op} \leqslant 1\}$, target $\|\cdot\|_{op}$

Our contribution:

- A new algorithmic framework for these problems
- A tighter constant in Spencer's theorem
- A proof of Komlós conjecture for "pseudorandom" inputs

Discrepancy and Continuous Methods

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

Discrepancy and Continuous Methods

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

Many different *algorithmic* proofs:

- [Bansal'10, Lovett-Meka'12] random walks, SDP
- [Eldan-Singh'14, Rothvoss'14] LP with random objective
- [Levy-Ramadas-Rothvoss'17] multiplicative weights update
- [Bansal-Laddha-Vempala'22] barrier potential function

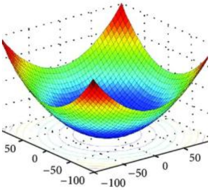
Discrepancy and Continuous Methods

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

Many different *algorithmic* proofs:

- [Bansal'10, Lovett-Meka'12] random walks, SDP
- [Eldan-Singh'14, Rothvoss'14] LP with random objective
- [Levy-Ramadas-Rothvoss'17] multiplicative weights update
- [Bansal-Laddha-Vempala'22] barrier potential function

More and more inspired by "continuous" optimization



Our algorithm: Newton's method on a regularized objective

Spencer's Theorem via Regularization (1/3)

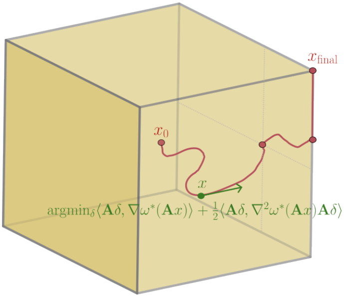
Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Spencer's Theorem via Regularization (1/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$



Algorithm:

Start from $x = (0, \ldots, 0)$

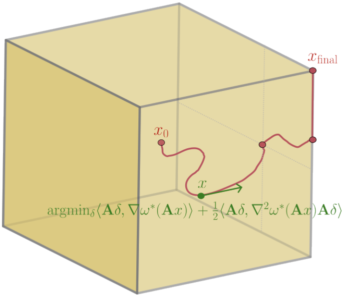
While $x \notin \{\pm 1\}^n$

$\quad F := \{i : x_i \notin \{\pm 1\}\}$

Spencer's Theorem via Regularization (1/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$



Algorithm:

Start from $x = (0, \ldots, 0)$

While $x \notin \{\pm 1\}^n$

$F := \{i : x_i \notin \{\pm 1\}\}$

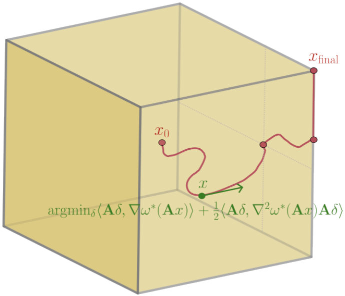
Find δ s.t. $\mathrm{supp}(\delta) \subseteq F, \delta \perp x$ minimizing

$$\langle \mathbf{A}\delta, \nabla\omega^*(\mathbf{A}x)\rangle + \frac{1}{2}\langle \mathbf{A}\delta, \nabla^2\omega^*(\mathbf{A}x)\mathbf{A}\delta\rangle$$

Spencer's Theorem via Regularization (1/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$



Algorithm:

Start from $x = (0, \ldots, 0)$

While $x \notin \{\pm 1\}^n$

$F := \{i : x_i \notin \{\pm 1\}\}$

Find δ s.t. $\text{supp}(\delta) \subseteq F, \delta \perp x$ minimizing

$$\langle \mathbf{A}\delta, \nabla \omega^*(\mathbf{A}x)\rangle + \frac{1}{2}\langle \mathbf{A}\delta, \nabla^2 \omega^*(\mathbf{A}x)\mathbf{A}\delta\rangle$$

x makes a small step in direction δ while staying in $[-1, 1]^n$

Spencer's Theorem via Regularization (2/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Spencer's Theorem via Regularization (2/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Up to tracking $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{A} & \mathbf{0} \end{bmatrix}$, assume WLOG

$$\|\mathbf{A}x\|_\infty = \max_{1 \leqslant i \leqslant n} (\mathbf{A}x)_i = \max_{r \in \Delta_n} \langle \mathbf{A}x, r \rangle \quad \text{where } \Delta_n := \{r \in \mathbb{R}^n_+ : \sum_i r_i = 1\}.$$

Spencer's Theorem via Regularization (2/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Up to tracking $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{A} & \mathbf{0} \end{bmatrix}$, assume WLOG

$$\|\mathbf{A}x\|_\infty = \max_{1 \leqslant i \leqslant n} (\mathbf{A}x)_i = \max_{r \in \Delta_n} \langle \mathbf{A}x, r \rangle \quad \text{where } \Delta_n := \{r \in \mathbb{R}^n_+ : \sum_i r_i = 1\} .$$

Def: *Regularized* maximum

$$\omega^*(y) := \max_{r \in \Delta_n} \langle y, r \rangle + \sum_{i=1}^n r_i^{\frac{1}{2}} .$$

Graph sparsification [Allen-Zhu, Liao, Orecchia'15], bandits [Audibert, Bubeck'09]

Spencer's Theorem via Regularization (2/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Up to tracking $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{A} & \mathbf{0} \end{bmatrix}$, assume WLOG

$$\|\mathbf{A}x\|_\infty = \max_{1 \leqslant i \leqslant n} (\mathbf{A}x)_i = \max_{r \in \Delta_n} \langle \mathbf{A}x, r \rangle \quad \text{where } \Delta_n := \{r \in \mathbb{R}^n_+ : \sum_i r_i = 1\}\,.$$

Def: *Regularized* maximum

$$\omega^*(y) := \max_{r \in \Delta_n} \langle y, r \rangle + \sum_{i=1}^n r_i^{\frac{1}{2}}\,.$$

Graph sparsification [Allen-Zhu, Liao, Orecchia'15], bandits [Audibert, Bubeck'09]

Claim: $\omega^*(\mathbf{A}x) = \|\mathbf{A}x\|_\infty \pm O(\sqrt{n})$

Spencer's Theorem via Regularization (3/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Analysis idea:

Spencer's Theorem via Regularization (3/3)

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there exists $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(\sqrt{n})$

1. Define a "smooth" proxy $\omega^*(\cdot)$ for $\|\cdot\|_\infty$
2. Run "sticky" Newton's Method on $x \mapsto \omega^*(\mathbf{A}x)$

Analysis idea:

1. By studying $\nabla^2 \omega^*(\mathbf{A}x)$, we prove that there always exists a direction $\delta \perp x$, $\mathrm{supp}(\delta) \subseteq F$ s.t.

$$\omega^*(\mathbf{A}(x+\delta)) - \omega^*(\mathbf{A}x) \leqslant \frac{\|\delta\|_2^2}{\sqrt{|F|}} .$$

2. Hence we get charged $\frac{1}{\sqrt{|F|}}$ cost "per unit of $\|x\|_2^2$"
3. Worst case: coordinates get frozen every time $\|x\|_2^2$ increases by 1, total cost

$$1 \times \frac{1}{\sqrt{n}} + 1 \times \frac{1}{\sqrt{n-1}} + \ldots = O(\sqrt{n}) . \quad \square$$

Our Results (1/2)

Improved Constant for Spencer's Theorem:

Thm: [P-V'23] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 3.68\sqrt{n}$.

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 5.32\sqrt{n}$.

Our Results (1/2)

Improved Constant for Spencer's Theorem:

Thm: [P-V'23] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 3.68\sqrt{n}$.

Thm: [Spencer'85] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 5.32\sqrt{n}$.

Proof idea: slightly different regularizer, track constants carefully

Our Results (2/2)

Komlós conjecture for "pseudorandom" inputs:

Thm: [P-V'23] For any $\mathbf{A} = [u_1 \mid \ldots \mid u_n]$ s.t. $\|u_i\|_2 \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t.

$$\|\mathbf{A}x\|_\infty = O(1 + \sqrt{\lambda \log n}),$$

where $\lambda := \max_{\|v\|_2=1, v \perp \mathbf{1}} \|\mathbf{A}^{\odot 2} v\|_2$ ($\odot \equiv$ entrywise product)

Our Results (2/2)

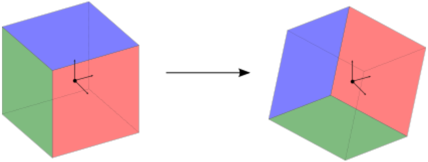
Komlós conjecture for "pseudorandom" inputs:

Thm: [P-V'23] For any $\mathbf{A} = [u_1 \mid \dots \mid u_n]$ s.t. $\|u_i\|_2 \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t.

$$\|\mathbf{A}x\|_\infty = O(1 + \sqrt{\lambda \log n}) \,,$$

where $\lambda := \max_{\|v\|_2=1, v \perp \mathbf{1}} \|\mathbf{A}^{\odot 2} v\|_2$ ($\odot \equiv$ entrywise product)

- Generalizes Banaszczyk's bound and [Potukuchi'18]
- Special case:



"A randomly rotated hypercube has a corner at ∞-distance $O(1)$ from the origin"

Open problem: Komlós conjecture for worst-case rotations?

Conclusion

Our algorithm: Newton's method on a regularized objective

Thm: [P-V'23] For any $\mathbf{A} \in \mathbb{R}^{n\times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 3.68\sqrt{n}$.

Thm: [P-V'23] For any $\mathbf{A} = [u_1 \mid \dots \mid u_n]$ s.t. $\|u_i\|_2 \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(1 + \sqrt{\lambda \log n})$, with

$$\lambda := \max_{\|v\|_2 = 1, v \perp \mathbf{1}} \|\mathbf{A}^{\odot 2} v\|_2 .$$

Conclusion

Our algorithm: Newton's method on a regularized objective

Thm: [P-V'23] For any $\mathbf{A} \in \mathbb{R}^{n \times n}$ s.t. $|\mathbf{A}_{ij}| \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty \leqslant 3.68\sqrt{n}$.

Thm: [P-V'23] For any $\mathbf{A} = [u_1 \mid \ldots \mid u_n]$ s.t. $\|u_i\|_2 \leqslant 1$, there is $x \in \{\pm 1\}^n$ s.t. $\|\mathbf{A}x\|_\infty = O(1 + \sqrt{\lambda \log n})$, with

$$\lambda := \max_{\|v\|_2 = 1, v \perp \mathbf{1}} \|\mathbf{A}^{\odot 2} v\|_2 .$$

Future directions:

- Prove a tight constant in Spencer's theorem
- Application of this framework to matrix discrepancy and graph sparsification
- Bridge this framework and arguments based on Lovász local lemma ($\rightsquigarrow$ Beck-Fiala conjecture)