Signrank vs Margin

Kaave Hosseini

University of Rochester

Based on Joint works with Hamed Hatami, Shachar Lovett, Ben Cheung, Morgan Shirley, Xiang Meng

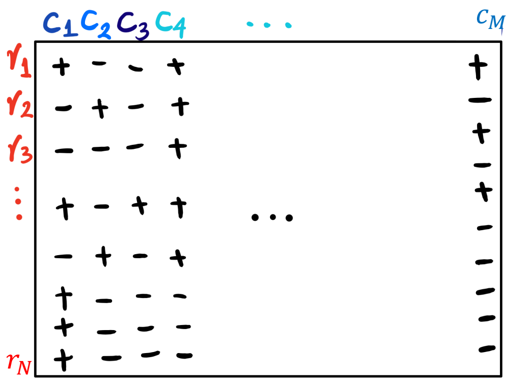
Notions of "rank"

Given $\{+1,-1\}$-matrix $A_{N\times M}$,

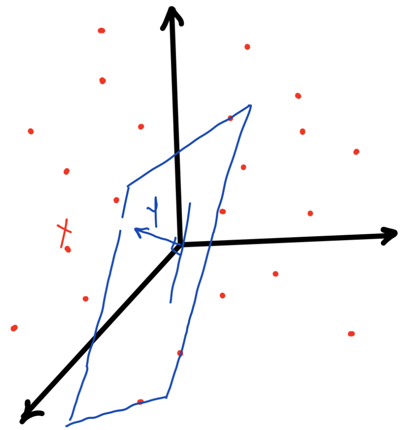
$\mathrm{Rank}(A)$ = smallest d such that $\exists x_1,\dots,x_N,y_1,\dots,y_M \in \mathbb{R}^d$ so that for all i,j $A_{ij} = \langle x_i, y_j\rangle$	$\gamma_2(A)$ = smallest ℓ such that: $\exists x_1,\dots,x_N,y_1,\dots,y_M \in \mathbb{R}^\infty$ so that for all i,j : $\lVert x_i\rVert_2 \cdot \lVert y_j\rVert_2 \leq \ell$ $A_{ij} = \langle x_i, y_j\rangle$
$\widetilde{\mathrm{Rank}}_\alpha(A)$ $1 \leq A_{ij}\cdot\langle x_i, y_j\rangle \leq \alpha$	$\widetilde{\gamma_2^\alpha}(A)$ $1 \leq A_{ij}\cdot\langle x_i, y_j\rangle \leq \alpha$
$\mathrm{Rank}^\pm(A)$ $1 \leq A_{ij}\cdot\langle x_i, y_j\rangle$	$\gamma_2^\infty(A)$ $1 \leq A_{ij}\cdot\langle x_i, y_j\rangle$

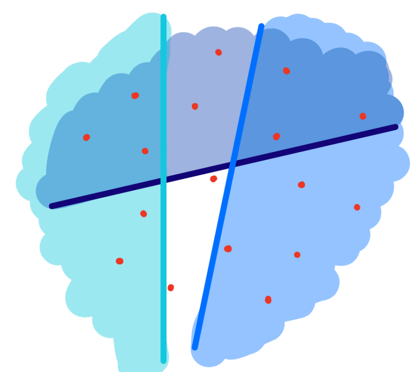
Meta question of this talk:

template: If $X(A)$ is small, how large can $Y(A)$ be?



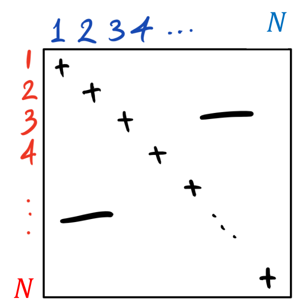
$\langle x, y \rangle > 0$





Example

Identity



$\mathrm{Rank}(A) = \mathrm{N}$	$\gamma_2(A) = 1$
$\widetilde{\mathrm{Rank}}(A) = \Theta(\log(\mathrm{N}))$ (Alon'09)	$\widetilde{\gamma_2}(A) = 1$
$\mathrm{Rank}^{\pm}(A) = 3$	$\gamma_2^{\infty}(A) = 1$

Applications of $\mathrm{Rank}^{\pm}$

- Learning Theory: sign-rank is known as *dimension complexity*
  - Both Upper bounds and Lower bounds
  - Example: fastest known learning algorithm for DNFs (Klivans-Servedio'04)

- Communication complexity:

  (Paturi-Simon '84) $\mathrm{Log}\left(\mathrm{Rank}^{\pm}(A)\right)$ = ***unbounded-error*** communication complexity of A

- circuit complexity lower bounds
  - Lower bounds for Threshold-of-Majority circuits (Razborov-Sherstov'08)

- semi-algebraic graphs

  Open question: Are Semialgebraic graphs of $O(1)$ complexity are exactly those of $\mathrm{Rank}^{\pm} = O(1)$?

Applications of γ_2^∞

- Machine learning: (γ_2^∞ is known as Margin Complexity)

The sample complexity of Support Vector Machine on a matrix A is $O\left((\gamma_2^\infty)^2\right)$.

- Communication complexity:

Theorem(Linial-Shraibman '07): $\gamma_2^\infty(A) = \Theta(\mathbf{Discrepancy}(A)^{-1})$

(based on Grothendieck inequality and duality)

(Chor-Goldreich'88, Klauck '01)

$\mathbf{log}(\mathbf{Discrepancy}(A)^{-1}) \leq$ Randomized Communication complexity of A

$\mathrm{Rank}^{\pm}$ vs γ_2^{∞}

Question. If $\mathrm{Rank}^{\pm}(A)$ is small, how large can $\gamma_2^{\infty}(A)$ be?

previous work:

[Buhrman-Vereshchagin-de Wolf07, Sherstov08, Sherstov11, Sherstov13, Thaler16, Sherstov19]

Previously known: there is $A_{N\times N}$ such that

$$\mathrm{Rank}^{\pm}(A) = \Theta(\log N) \text{ and } \gamma_2^{\infty}(A) \geq \mathrm{poly}(N)$$

On the other hand, it's well known that for any B with bounded entries

$$\gamma_2(B) \leq \sqrt{\mathrm{rank}(B)} \text{ and } \widetilde{\gamma_2}(B) \leq O\left(\sqrt{\widetilde{\mathrm{rank}}(B)}\right)$$

(Using John's theorem from Convex Geometry.)

Theorem (Hatami-H-Lovett '20): There is $A_{N\times N}$ such that

$$\mathrm{Rank}^{\pm}(A) = 3 \text{ but } \gamma_2^{\infty}(A) \geq \mathrm{poly}(N)$$

Construction: 3-dimensional Inner product over integers

$$x = (x_1, x_2, x_3). \quad x_1, x_2, x_3 \in [-N, N]$$
$$y = (y_1, y_2, y_3). \quad y_1, y_2, y_3 \in [-N, N]$$

$$A(x,y) = \begin{cases} +1 & if\ \langle x, y \rangle \geq 0 \\ -1 & if\ \langle x, y \rangle < 0 \end{cases}$$

$$\text{Rank}^{\pm}(A) = 3$$

Theorem (Hatami-H-Lovett '20): $\gamma_2^{\infty}(A) \geq \sqrt{N}$

$\mathrm{Rank}^{\pm}$ vs γ_2^{∞}

Question. If $\gamma_2^{\infty}(A)$ is small, how large can $\mathrm{Rank}^{\pm}(A)$ be?

Theorem (Linial, Mendelson, Schechtman, and Shraibman ‘07, Arriaga-Vempala ‘06):

$$\mathrm{Rank}^{\pm}(A_{N\times N}) = O\left(\left(\gamma_2^{\infty}(A)\right)^2 \log(N)\right)$$

(Proof based on Johnson-Lindenstrauss lemma.)

Question (Linial, Mendelson, Shechtman, Shraibman ‘07):

Is the $\log(N)$ term necessary?

Theorem (Hatami-H-Meng’23): $\log(N)$ term is necessary for partial matrices.

$\text{Rank}^{\pm}$ vs γ_2^{∞}

Theorem (Newman's lemma): $A_{2^n \times 2^n}$

$$R^{\text{private}}(A) \leq R^{\text{public}}(A) + O(\log n)$$

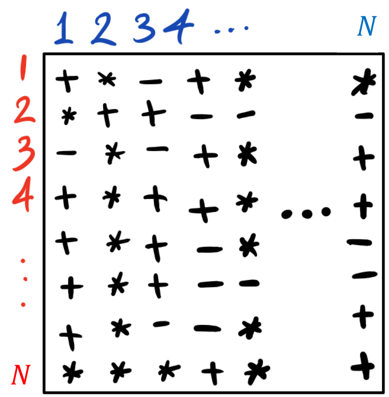
$$R^{\text{private}}_{unbounded}(A) \leq R^{\text{public}}(A) + O(\log n)$$

Question. Is the $O(\log n)$ term necessary above?

Corollary (Hatami-H-Meng'23): $\mathrm{O}(\log(n))$ is necessary (for partial matrices)

Construction

We give a construction of a partial matrix:



Pick arbitrary $\epsilon > 0$. We give partial matrix $A_{2^n \times 2^n}$ so that

$$\gamma_2^{\infty}(A) = 1 + \epsilon$$

$$\mathrm{Rank}^{\pm}(A) > \Omega\left(\frac{\epsilon \cdot n}{\log(\epsilon^{-1})}\right)$$

Construction

Gap Inner Product(GIP):

$$x, y \in \left\{\frac{-1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right\}^n$$

$$GIP_\epsilon^n(x, y) = \begin{cases} + & \langle x, y\rangle > 1-\epsilon \\ * & -(1-\epsilon) \leq \langle x, y\rangle \leq 1-\epsilon \\ - & \langle x, y\rangle < -(1-\epsilon) \end{cases}$$

Theorem. Let $\epsilon \in (0,1)$.

$$\Omega\left(\frac{\epsilon n}{\log(\epsilon^{-1})}\right) = \mathrm{Rank}^{\pm}(GIP_\epsilon^n) = O(\epsilon n)$$

Main Lemma

Proof idea: first study the continuous version of the problem

$$x, y \in \mathbb{S}^{n-1} \subset \mathbb{R}^n$$

$$\mathbb{H}_\epsilon^n(x, y) = \begin{cases} + & \langle x, y \rangle > 1 - \epsilon \\ * & -(1-\epsilon) \le \langle x, y \rangle \le 1 - \epsilon \\ - & \langle x, y \rangle < -(1-\epsilon) \end{cases}$$

(class of halfspaces with margin $1 - \epsilon$)

Main Lemma. For all $n \in \mathbb{N}$ and $\epsilon \in (0,1)$, $\mathrm{Rank}^{\pm}(\mathbb{H}_\epsilon^n) = n$.

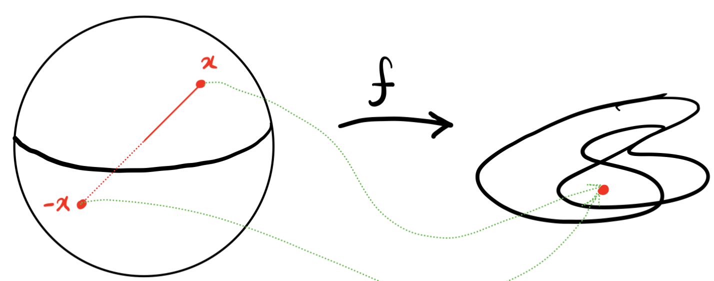
Main Lemma

Main Lemma. For all $n \in \mathbb{N}$ and $\epsilon \in (0,1)$, $\mathrm{Rank}^{\pm}(\mathbb{H}_{\epsilon}^{n}) = n$.

Proof Idea: **Topology**

Borsuk-Ulam theorem: Let $f: \mathbb{S}^{d-1} \to \mathbb{R}^{d-1}$ be an arbitrary **continuous** map.

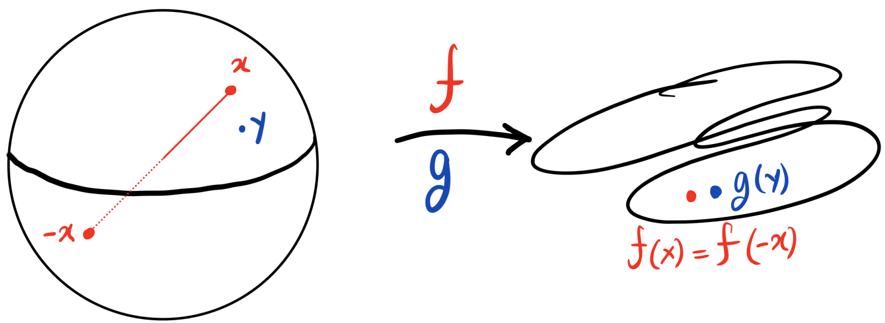
There is a point $x \in \mathbb{S}^{d-1}$ so that $f(x) = f(-x)$



Main Lemma

Main Lemma. For all $n \in \mathbb{N}$ and $\epsilon \in (0,1)$, $\mathrm{Rank}^{\pm}(\mathbb{H}_{\epsilon}^{n}) = n$.

Proof Idea: If the maps f, g are continuous.



$\langle x, y \rangle > \gamma$ hence $\mathbb{H}_{\epsilon}^{n}(x, y) = +1$ and $\langle f(x), g(y) \rangle > 0$

Also $\langle -x, y \rangle < -\gamma$ hence $\mathbb{H}_{\epsilon}^{n}(x, y) = -1$, however, by Borsuk-Ulam:

$$\langle f(-x), g(y) \rangle = \langle f(x), g(y) \rangle > 0$$

If not continuous, find a careful continuation $\tilde{f}, \tilde{g}$ that preserves most of the inner-product signs.

γ_2 vs $\widetilde{\gamma_2}$

Question. If $\widetilde{\gamma_2}$ is small, how large can γ_2 be?

Linial-Shraibman'09:

$$\log\left(\widetilde{\gamma_2}(A)\right) \leq \mathrm{R}^{public}(A) \qquad \leq \widetilde{\gamma_2}(A)$$

Question: **Linial-Shraibman'09 , also Pitassi, Shirley, Shraibman'23**

Can one substitute $\log\left(\widetilde{\gamma_2}(A)\right)$ by $\log\left(\gamma_2(A)\right)$ above?

Theorem (Cheung-Hatami-H-Shirley'23). No.

There is a matrix $A_{N\times N}$ such that $\mathrm{R}^{public}(A) \leq O(\log\log N)$ but $\gamma_2(A) \geq \mathrm{poly}(N)$.

Hence $\widetilde{\gamma_2}(A) = \mathrm{polylog}(n)$ but $\gamma_2(A) \geq \mathrm{poly}(N)$

γ_2 vs $\widetilde{\gamma_2}$

Theorem (Cheung-Hatami-H-Shirley'23).

There is a matrix $A_{N\times N}$ such that $\widetilde{\gamma_2}(A) \leq O(\text{poly}\log N)$ but $\gamma_2(A) \geq \text{poly}\, N$.

$$x = (x_1, x_2, x_3). \quad x_1, x_2, x_3 \in [-N, N]$$
$$\text{y} = (y_1, y_2, y_3). \quad y_1, y_2, y_3 \in [-N, N]$$

$$A(x, y) = \begin{cases} +1 & if\ \langle x, y\rangle = 0 \\ -1 & if\ \langle x, y\rangle \neq 0 \end{cases}$$

Open problems

$\mathrm{Rank}(A)$	$\gamma_2(A)$
$\widetilde{\mathrm{Rank}}(A)$	$\widetilde{\gamma_2}(A)$
$\mathrm{Rank}^{\pm}(A)$	$\gamma_2^{\infty}(A)$

Problem 1. If $\gamma_2^{\infty}(A) = O(1)$, how large can $\gamma_2(A)$ be?

Linial-Shraibman ($\gamma_2(A)$ can not be larger than $\sqrt{N}$)

Problem 2. Construct a *total* matrix that $\gamma_2^{\infty}(A) = O(1)$ but $\mathrm{Rank}^{\pm}(A) = \omega(1)$.

Problem 3. If $\gamma_2(A) = O(1)$, does it imply that $\mathrm{Rank}^{\pm}(A) = O(1)$?

(Hatami-Hatami-Pires-Tao-Zhao'22) It is true for Cayley graphs of abelian groups:

$$\mathrm{Rank}^{\pm}(A) \leq 2^{2^{\gamma_2(A)}}$$

Problem 4. If $\gamma_2^{\infty}(A) = O(1)$ is there a monochromatic rectangle of density $\Omega(1)$?

True for $\mathrm{Rank}^{\pm}(A)$: Alon-Pach-Pinchasi-Radoičić-Sharir'09, Fox-Pach-Suk'16:

A has a monochromatic rectangle of density at least $2^{-\mathrm{Rank}^{\pm}(A)}$

Thank you!