Potential Projects

Guy Moshkovitz

Contents

1 Topics 1

2 Reading 2

3 Example project: Complexity of Tensor Ranks 3

One area of research I am particularly interested in is notions of tensor & polynomial rank, and their applications in extremal combinatorics, number theory, theoretical computer science, and more. For some background, see the following documents:

- What is... a tensor?
- Tensor Ranks - References

See also [4, 3, 2, 1] for related past REU projects.

1 Topics

My REU projects span multiple areas within the fields of extremal combinatorics, algebra, and theoretical computer science. Projects in previous years included topics such as structure-vs-randomness of tensors, finite-field Nullstellensatz, and sunflower conjectures.

Here is a general list of topics that I am interested in, and which could be relevant for potential projects with an interested student.

- **Additive combinatorics**: special cases of the Polynomial Gowers Inverse conjecture and of the Polynomial Freiman-Ruzsa conjecture.
(Areas: multivariate polynomials, fields and finite fields.)
- **Algebraic geometry**: rational points on structured varieties, variants of Hilbert's Nullstellensatz.
(Areas: multivariate polynomials, fields and finite fields, commutative algebra.)
- **Circuit complexity**: depth-3 boolean and arithmetic circuits, annihilating polynomials.
(Areas: multivariate polynomials, fields and finite fields, combinatorics.)
- **Extremal set theory**: new variants of the Sauer-Shelah Lemma and applications in machine learning, sunflower conjecture(s).
(Areas: combinatorics.)

- **Information theory**: counting independent sets in graphs using Shearer's Lemma and related inequalities, variants of the Shannon capacity of a graph.
  (Areas: combinatorics, graph theory.)
- **Ramsey theory**: Multi-color Ramsey numbers, applications of Ramsey theory in additive combinatorics.
  (Areas: combinatorics, graph theory.)
- **Szemerédi's regularity lemma**: variants of the graph and hypergraph regularity lemmas, applications to the removal lemma and other applications.
  (Areas: combinatorics, graph theory.)
- **Tensor ranks**: structure-vs-randomness of tensors over small fields, applications to computer science.
  (Areas: multivariate polynomials, fields and finite fields, computational complexity.)
  Below you can find an example of a project involving a computer science application of tensor ranks.

2 Reading

Here is a list of reading materials covering many of the areas mentioned above.

Additive combinatorics:

- Additive Combinatorics and Theoretical Computer Science - L. Trevisan.
- Finite field models in arithmetic combinatorics – ten years on - J. Wolf.
- Selected Results in Additive Combinatorics: An Exposition - E. Viola.

Algebra and algebraic geometry:

- Algebraic Geometry: A Problem Solving Approach (mostly Chapter 4).
- Undergraduate Algebraic Geometry - M. Reid (mostly Chapters 3,4,6).
- Ideals, Varieties, and Algorithms - D. Cox, J. Little, and D. O'Shea.

Circuit complexity:

- A selection of lower bounds for arithmetic circuits - N. Kayal and R. Saptharishi.
- Arithmetic Circuits: a survey of recent results and open questions - A. Shpilka and A. Yehudayof.
- Recent progress on lower bounds for arithmetic circuits - S. Saraf.

Extremal combinatorics:

- Combinatorics, lecture notes - J. Fox.
- Extremal Combinatorics, With Applications in Computer Science - S. Jukna.
- Extremal Set Theory and the Linear Algebra Method - T. Szabo.

Finite fields:

- Finite Fields - T. Murphy.
- Introduction to Finite Fields - D. Forney.
- Topics in Finite Fields (intro) - S. Kopparty.

Information Theory:

- Entropy, Independent Sets and Antichains - J. Kahn.
- Combinatorial Reasoning in Information Theory - N. Alon (mostly Section 2).
- On the number of high-dimensional partitions - C. Pohoata and D. Zakharov.

Ramsey Theory:

- Lower bounds for multicolor Ramsey numbers - D. Conlon and A. Ferber.
- Multicolor Ramsey numbers - Y. Wigderson.
- Recent developments in graph Ramsey theory - D. Conlon, J. Fox, and B. Sudakov.

Regularity lemma:

- Applications of the Szemeredi Regularity Lemma - S. Das.
- Graphs & Algorithms Advanced Topics: Szemeredi's Regularity Lemma - U. Wagner.
- Graph removal lemmas - D. Conlon and J. Fox.

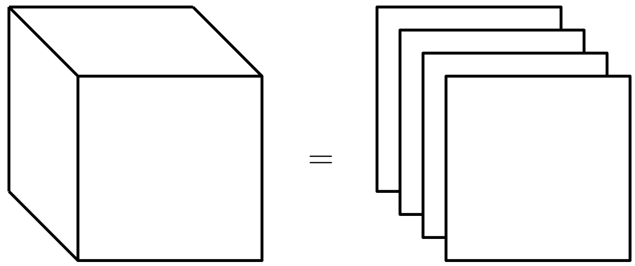
Sunflowers:

- Sunflowers: from soil to oil - A. Rao.
- On Sunflowers and Matrix Multiplication - N. Alon, A. Shpilka, and C. Umans.
- Extremal problems on Δ-systems - A. Kostochka.

3 Example project: Complexity of Tensor Ranks

A (3-) *tensor* is a 3-dimensional array $T \in \mathbb{F}^{n\times n\times n}$ with entries in some field $\mathbb{F}$. For matrices (2-dimensional arrays) there is only one notion of *rank* with many equivalent definitions: column rank, row rank, decomposition rank, and determinantal rank, to name a few. For tensors, however, there is a myriad of different notions of rank, coming from different fields of math: combinatorics, analysis, algebraic geometry, computer science, and more. Many questions are open about the interplay between these notions, and equally importantly, about their computational complexity. Linear algebra tells us that the rank of an $n \times n$ matrix can be computed quite efficiently, for example by Gaussian elimination, in time $O(n^3)$ (and possibly in linear time $O(n^2)$!). Finding an efficient algorithm for any of the various notions of tensors rank would be extremely interesting. On the flip side, showing NP-hardness for many of these notions is open as well.

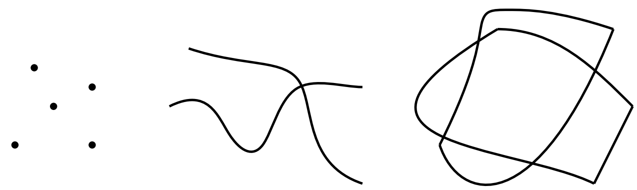
One notion of tensor rank whose computational complexity is unknown is *geometric rank* [8]. It is defined using the concept of a *variety* from algebraic geometry. A variety is the solution set of polynomial equations, $\{\mathbf{x} \in \overline{\mathbb{F}}^n \mid f_1(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0\}$, where $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n]$ are polynomials with coefficients in $\mathbb{F}$, and $\overline{\mathbb{F}}$ denotes the algebraic closure of $\mathbb{F}$. To define the geometric rank of a tensor $T = (t_{i,j,k})_{i,j,k\in[n]} \in \mathbb{F}^{n\times n\times n}$, first note that its 2-dimensional slices $T_1, \dots, T_n \in \mathbb{F}^{n\times n}$ along any of the three axes, say $T_k = (t_{i,j,k})_{i,j\in[n]}$, determine a *bilinear map* $\widetilde{T}\colon \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$ given by $\widetilde{T}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^t T_1 \mathbf{y}, \dots, \mathbf{x}^t T_n \mathbf{y})$.[1]



The *kernel* of $\widetilde{T}$ is the variety $\ker \widetilde{T} = \{(\mathbf{x}, \mathbf{y}) \in \overline{\mathbb{F}}^n \times \overline{\mathbb{F}}^n \mid \widetilde{T}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$, and the geometric rank of T is the codimension of the kernel,

$$\mathrm{GR}(T) = \mathrm{codim}(\ker \widetilde{T}) = 2n - \dim(\ker \widetilde{T}).$$

The *dimension* of a variety V is the maximum length of a chain of irreducible[2] varieties contained in V.[3]



Note that $\mathrm{GR}(T)$ is an integer in the range $\{0, \dots, n\}$, where the upper bound follows from the fact that $\{(\mathbf{0}, \mathbf{y}) \mid \mathbf{y} \in \overline{\mathbb{F}}^n\} \subseteq \ker \widetilde{T}$. We note that geometric rank is independent of the choice of axis along which we slice the tensor [8], analogously to how matrix rank is independent of whether we form the linear map by slicing the matrix along the rows or along the columns (a.k.a column rank = row rank). Practically speaking, one can compute the geometric rank of small tensors using computer software like Macaulay2 or Sage that can compute the dimension of varieties. However, for general tensors it is not clear whether there is an efficient algorithm or not.

Question 1. *Is deciding* $\mathrm{GR}(T) \le r$ *for* $T \in \mathbb{F}^{n\times n\times n}$ NP*-hard?*

Interestingly, for a system of degree-2 homogeneous equations (though not necessarily bilinear!), even deciding if there is a nonzero solution is known to be NP-hard (see Theorem 2.6 in [7] which reduces from graph 3-colorability). However, even if computing geometric rank *exactly* is hard, approximating it—by computing a value guaranteed to be, say, double the right answer—might be easy.

Question 2. *What is the computational complexity of approximating geometric rank up to a multiplicative constant c? (for various values of c)*

[1] Just like a 2-dimensional matrix corresponds to a *linear* map, its n components being linear forms.

[2] A variety is irreducible if it cannot be written as a (proper) union of varieties. Every variety can be written uniquely as a finite union of irreducible varieties; e.g., $\{xy = 0\} = \{x = 0\} \cup \{y = 0\}$.

[3] Equivalently, it is the maximum size of an algebraically-independent set in the quotient ring $\mathbb{F}[x_1, \dots, x_n]/\,\mathrm{I}(V)$. (The *ideal* of V is $\mathrm{I}(V) = \{f \in \mathbb{F}[x_1, \dots, x_n] \mid \forall \mathbf{x} \in V\colon f(\mathbf{x}) = 0\}$.)

Another rank notion for tensors is *slice rank*, introduced by Tao [10] in the context of the recent breakthrough solution of the cap-set problem. To define the slice rank of a tensor $T = (t_{i,j,k})_{i,j,k\in[n]} \in \mathbb{F}^{n\times n\times n}$, first identify it with the trilinear polynomial $\overline{T} = \sum_{i,j,k\in[n]} t_{i,j,k} x_i y_j z_k$. The tensor T is said to have slice rank 1 if $\overline{T}$ is a product of a linear form and a bilinear form (in disjoint variables groups $\mathbf{x}, \mathbf{y}, \mathbf{z}$). The slice rank $\mathrm{SR}(T)$ is the smallest number of tensors of slice rank 1 whose sum is T. Deciding $\mathrm{SR} \leq r$ was shown to be NP-hard [5] via a reduction from a version of the minimum vertex cover problem for 3-uniform hypergraphs. The approximation question, however, remains open.

Question 3. *What is the computational complexity of approximating slice rank up to a multiplicative constant c? (for various values of c)*

One can ask similar questions about other rank notions for tensors, such as *analytic rank*, *subrank*, and *tensor rank*. For the latter notion, which is defined similarly to slice rank except each summand is a product of three linear forms, these questions have been studied in greater depth, starting from Håstad's classical result from 1990 [6] proving NP-hardness, as well as more recent results showing, for example, NP-hardness of approximating tensor rank to within a factor of 1.0005 [9]. These and other questions about ranks of tensors are closely related to algebraic complexity (arithmetic circuit lower bounds, algorithmic matrix multiplication), quantum information theory (quantifying quantum entanglement), and extremal combinatorics (e.g., the cap-set and the sunflower problems).

References

[1] A. Cohen and G. Moshkovitz, Structure vs. randomness for bilinear maps, **Discrete Anal.** 12 (2022). Conference version appeared in 53rd ACM Symposium on Theory of Computing (**STOC 2021**), 800–808.

[2] A. Cohen and G. Moshkovitz, Partition and analytic rank are equivalent over large fields, **Duke Math. J.** 172 (2023), 2433–2470.

[3] G. Moshkovitz and J. Yu, Sharp effective finite-field Nullstellensatz, **Am. Math. Mon.** 130 (2023), 720–727.

[4] G. Moshkovitz and D. Zhu, Quasi-linear relation between partition and analytic rank, arXiv:2211.05780 (2022), submitted.

[5] M. Bläser, C. Ikenmeyer, V. Lysikov, A. Pandey and F. O. Schreyer, Variety membership testing, algebraic natural proofs, and geometric complexity theory, arXiv:1911.02534 (2019).

[6] J. Håstad, Tensor rank is NP-complete, **J. Algorithms** 11, 644–654 (1990).

[7] C. Hillar and L. H. Lim, Most tensor problems are NP-hard, **J. ACM** 60, No. 45 (2013).

[8] S. Kopparty , G. Moshkovitz and J. Zuiddam, Geometric rank of tensors and subrank of matrix multiplication, **Discrete Anal.** 1 (2023). Conference version appeared in 35th Computational Complexity Conference (**CCC 2020**) 35, 1–21.

[9] J. Swernofsky, Tensor rank is hard to approximate, Leibniz International Proceedings in Informatics, LIPIcs 26 (2018).

[10] T. Tao, A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound, `https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound` (2016).