Ordering by vantage points and sign patterns

Talk by Noga Alon

Notes by Sanjana Das

May 8, 2025

This is joint work with Colin Defant, Noah Kravitz, and Daniel Zhu.

§1 Introduction

Maybe one of the main reasons Noga wants to talk about this topic is to show us some simple techniques that he thinks are very powerful. It involves a variant of a method used quite extensively in combinatorics and TCS. He'll first briefly show one of the early applications of this method, and then tell us about this new variant, which he thinks should also be powerful.

§1.1 The problem and motivation

We'll first think about ordering points by a sum of distances. So we're given a finite set of points $C \subseteq \mathbb{R}^d$ (you can think of d as 2 if you want — that case is interesting enough) with $|C| = n$. And we have a generic set $V \subseteq \mathbb{R}^d$ with $|V| = k$. ('Generic' essentially means you don't want some coincidental equalities.) We use V — the 'vantage points' — to define an ordering on C, according to the sum of the Euclidean distances from the points of V to each point in C. More precisely, we look at

$$\sum_{v \in V} \|v - c\|_2 \,,$$

and we order the points $c \in C$ from small to large; this defines an ordering of C.

Question 1.1. How many orderings like this are possible?

More precisely, we want to understand the following quantity:

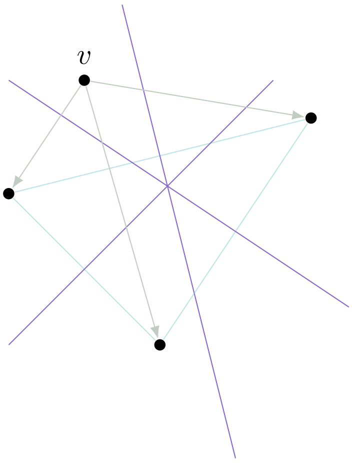
Definition 1.2. We define $F(k, d, C)$ as the set of all orderings of C as V varies, and

$$f(k, d, n) = \max\{|F(k, d, C)| \mid C \subseteq \mathbb{R}^d,\ |C| = n\}.$$

One interpretation of this problem, suggested in a paper that first considered it, has to do with candidate ranking. You can view C as a set of 'candidates,' and their locations in $\mathbb{R}^d$ can represent their 'opinion' on d different issues. And V represents a set of 'voters.' It's natural that every voter also has his own opinions, so is also a location in Euclidean space. And it makes sense that you prefer candidates that are closer to you; so each voter gives a better score to the candidates close to him. And then the average distance from the voters to each candidate defines a ranking on the candidates; and we want to know how many there are.

§1.2 Some results

Good and Tideman (1977) and Zaslavsky (2002) considered the case where $k = 1$, so your ordering is determined according to the distances to just one point. Here it's simple to describe what happens: For every pair of points in C, you can draw the hyperplane that bisects the line segment between them. Then every point to the left of this plane is closer to the first point, and every point to the right is closer to the second.



So we just want to know how many cells we have in this hyperplane arrangement, which consists of $\binom{n}{2}$ bisecting hyperplanes. And there are precise formulas for this (these hyperplanes aren't in general position — they're obtained as the bisectors of lines — but we still an exact formula). We could write this formula down (it's some summation of binomial coefficients), but we'll be interested in just asymptotics. In particular, the formula gives

$$f(1, d, n) = \Theta(n^{2d}).$$

In the plane (when $d = 2$), this is n^4.

But the problem already gets more complicated in dimension 2 with 2 vantage points.

Question 1.3 (CCGKOS 2021). Estimate $f(2, 2, n)$. In particular, is it exponential in n?

So we have n points in the plane $\mathbb{R}^2$, and we have two vantage points; we rank the n points according to their sum of distances to these two points. And the question is, how many orderings are possible? The total number of possible linear orders is $n!$; these authors proved some exponential upper bound (which is a bit better than $n!$), and they asked whether that's the right answer.

Theorem 1.4 (Alon–Defant–Kravitz–Zhu 2024+)

We have $f(2, 2, n) = \Theta(n^8)$. More generally, for fixed $k \geq 1$ and $d \geq 2$, we have $f(k, n, d) = \Theta_{d,k}(n^{2dk})$.

(They also proved that $f(k, 1, n) = \Theta_k(n^{4\lceil k/2\rceil-2})$ — dimension 1 is a special case.)

In this talk, Noga will describe a proof of the upper bound. The lower bound is not easy and requires some work, but it's more of a special thing, while the upper bound is a technique that's good to know — it's simple and powerful. So that's what he'll describe.

§2 Sign patterns

Now Noga will tell us about sign patterns, which are a different thing but will be useful.

Here, the story is that we have a collection of m functions $F = (f_1, \ldots, f_m)$, where $f_i : \mathbb{R}^N \to \mathbb{R}$. Then for each point $x \in \mathbb{R}^N$ where no f_i vanishes, you can look at the sign of each of these functions.

Definition 2.1. If no f_i vanishes at x, we define the sign pattern of F at x as

$$s(F, x) = (\operatorname{sgn}(f_1(x)), \ldots, \operatorname{sgn}(f_m(x))) \in \{\pm 1\}^m.$$

We define $s(F)$ as the number of sign patterns of F (over all x).

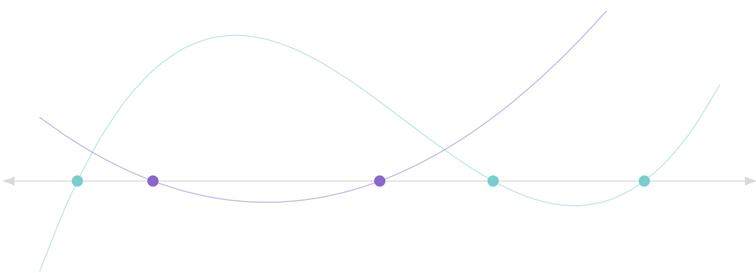
There's an old result of Warren regarding these, which is a sharpening of a more general result of Milner and Thom from real algebraic geometry.

Theorem 2.2 (Warren 1968)

Suppose that F is a collection of m real polynomials in N variables. If the degree of each f_i is at most Δ and $m \geq N$, then $s(F) \leq (4e\Delta m/N)^N$.

The exact statement of the theorem is in terms of some binomial coefficients, but here we only care about asymptotics.

As a sanity check, you can consider $N = 1$ (where we just have polynomials of one variable); in this bound, we're essentially summing the degrees of the polynomials (there are m polynomials of degree Δ). To prove this for $N = 1$, you'd look at the real line and take all the zeros of each polynomial. The total number of zeros will be (at most) the sum of degrees, and this number tells you something about the number of sign patterns (since signs only change at these zeroes).



But this holds in general. In fact, the way it's proved is by bounding the number of connected components:

Theorem 2.3 (Warren 1968)

In the same setup, the number of connected components of the semi-variety

$$\{x \in \mathbb{R}^d \mid f_i(x) \neq 0 \text{ for all } i\}$$

is at most $(4e\Delta m/N)^N$.

It's clear that this implies Theorem 2.2, since two points in the same connected component have to have the same sign patterns (if we have two different sign patterns, then we can move from one point to the other; somewhere in the middle, one of the f_i has to vanish in order to change signs).

Remark 2.4. Milnor's result was for the same setting, but he looked at all Beatty numbers; here we only care about the 0th, which is the number of connected components.

Sign patterns have lots of applications in combinatorics and TCS. We'll describe 1.5 simple applications (but these are still something we don't know how to solve without this tool), and then we'll discuss a variant or extension and how it's relevant to vantage points.

§3 Applications to signrank

Definition 3.1. Given a matrix A with ± 1 entries, its signrank is the minimum possible rank of a real matrix B with the same signs as A, i.e., such that $A_{ij}B_{ij} > 0$ for all i and j.

Example 3.2

The signrank of the matrix

$$A = \begin{bmatrix} -1 & +1 & +1 & +1 \\ +1 & -1 & +1 & +1 \\ +1 & +1 & -1 & +1 \\ +1 & +1 & +1 & -1 \end{bmatrix}$$

is 3 (this would be true even if the matrix was $n \times n$). As a construction (for $n = 4$), we could take

$$B = \begin{bmatrix} -3 & +1 & +1 & +1 \\ +1 & -3 & +1 & +1 \\ +1 & +1 & -3 & +1 \\ +1 & +1 & +1 & -3 \end{bmatrix}$$

(with -3's on the diagonal and $+1$'s everywhere else); every row has sum 0, so the rank of B is at most $4 - 1 = 3$. (And we can show that this is tight.)

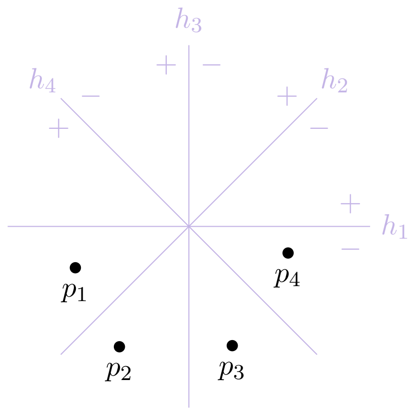
Signrank has a geometric interpretation: It's the minimum dimension d such that there is a hyperplane h_i through the origin in $\mathbb{R}^d$ for each row of A, and a point p_j in $\mathbb{R}^d$ for each column of A, such that the point p_j lies on the positive side of the hyperplane h_i if and only if $A_{ij} > 0$ (and otherwise it's on the negative side). This is because the rank of a matrix being d means that we can write it as a $(n \times d) \times (d \times n)$ matrix product, and we can think of one of these matrices as defining hyperplanes, and the other as defining points. So this gives an equivalent characterization of signrank — if the signrank is small, we can represent A using hyperplanes and points in a small dimension.

Example 3.3

The signrank of the matrix

$$\begin{bmatrix} -1 & -1 & -1 & -1 \\ +1 & -1 & -1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & +1 & +1 & -1 \end{bmatrix}$$

is 2. We can see this by drawing four lines in $\mathbb{R}^2$, and placing points between them appropriately.



One place signrank appears is in communication complexity. Here, the setup is that there's two parties Bob and Alice who both know A; Alice knows a row i and Bob knows a column j. They have to communicate some bits (possibly randomly) so that eventually, one of them finds A_{ij} with probability bigger than $\frac{1}{2}$. The *unbounded error communication complexity* is the number of bits this requires.

Theorem 3.4 (Paturi–Simon 1985)
The signrank of A determines the unbounded error communication complexity of the function given by A — it's roughly $\log_2 \operatorname{signrank}(A)$.

For example, Equality has signrank 3; so this means Alice and Bob just have to communicate 2 bits in order to determine whether their numbers are equal or not (but their success probability will only be a bit bigger than $\frac{1}{2}$).

Remark 3.5. Bob and Alice are the names typically used for two parties communicating, because they begin with B and A; they first appeared in the RSA paper.

Signrank is also related to *linear classifiers* in learning theory, where we try learning concepts by embedding things in an Euclidean space — the picture is that we have hyperplanes with some points on one side and some on the other, and this sort of determines a concept. And the signrank is basically equivalent to the minimum dimension where we can represent a matrix with hyperplanes and points in this way.

Question 3.6 (Paturi–Simon)**.** What is the maximum possible signrank of an $n \times n$ matrix of signs?

In other words, if we're given an $n \times n$ matrix of signs, can we always put in real numbers with those signs to make the rank very small? This is a very natural question. And in the connection to communication complexity, we're asking, are there any functions that really need many bits of communication? The model of unbounded error communication complexity is very powerful; can you solve *any* problem of this form by sending only 5 bits (for example)?

Theorem 3.7 (Alon–Frankl–Rödl)
There are $n \times n$ matrices A with $\operatorname{signrank}(A) \geq \frac{n}{20}$, and we always have $\operatorname{signrank}(A) \leq (1 + o(1))\frac{n}{2}$.

Note that in the communication problem, Alice can always send $\log_2 n$ bits and just specify i (the number of her row). So *every* function can be computed by sending $\log_2 n$ bits, even with probability 1. And the lower bound here says that even if we only want a *tiny* advantage over $\frac{1}{2}$ (in the success probability), for some functions we need $\log_2 n - 5$ bits.

Proof of lower bound. The lower bound follows by a counting argument using Warren's theorem (Theorem 2.2). We're looking at the signs of a real $n \times n$ matrix of rank d. We can decompose it as a product of $n \times d$ and $d \times n$ matrix; then the entries of the $n \times n$ matrix are given by n^2 quadratic polynomials in the $2dn$ entries of these matrices. So we have n^2 quadratic polynomials in $2dn$ variables, and we're looking at their sign pattern. Then you just substitute — Warren's theorem gives a bound on the total number of sign patterns of n^2 quadratic polynomials in $2dn$ variables, and that should be bigger than the number of possible sign patterns (otherwise there will be a sign pattern that can't be obtained in this way, giving A of signrank greater than d). So we get

$$\left(\frac{8en^2}{2dn}\right)^{2dn} \geq 2^{n^2}.$$

Then you can solve for d, and it ends up being n divided by some constant. □

This proof is simple, but you need this powerful tool (Theorem 2.2) — Noga doesn't know how to prove this without it. (The statement of Theorem 2.2 is basically like Bezout's theorem, but proving it involves some Morse theory, and you have to work a bit.)

Here's another version of this question.

Question 3.8 (Ben-David–Eiron–Simon 2002)**.** What is the maximum possible signrank of an $n \times n$ sign matrix with VC dimension 2?

The VC dimension is another useful measure of how complex a matrix is. The definition is not so essential for this discussion, but here it is:

Definition 3.9. The VC dimension of a matrix of signs is the maximum cardinality of a set of columns I which is shattered, meaning that if you look at all the rows restricted to these columns, you see all the possibilities of $\pm$ signs.

Example 3.10

Consider the matrix

$$\begin{bmatrix} + & - & - & + & + \\ + & - & + & - & + \\ + & + & + & + & + \\ - & + & - & + & - \\ - & + & + & - & - \end{bmatrix}.$$

Its VC dimension is 2, because if we look at the first and third column, each of the four possible rows $(+-)$, $(++)$, $(--)$, and $(-+)$ appears. (And it can't be greater because having VC dimension at least 3 would require at least 8 rows.)

$$\begin{bmatrix} + & - & - & + & + \\ + & - & + & - & + \\ + & + & + & + & + \\ - & + & - & + & - \\ - & + & + & - & - \end{bmatrix} \begin{matrix} \rightsquigarrow & +- \\ \rightsquigarrow & ++ \\ \rightsquigarrow & ++ \\ \rightsquigarrow & -- \\ \rightsquigarrow & -+ \end{matrix}$$

If the VC dimension is small, then the matrix is 'simple' in some sense. So if the VC dimension is 2, since your matrix is simple, you might hope that you could embed it in low-dimensional Euclidean space using hyperplanes and points. But it turns out that this is false — the maximum possible signrank is still pretty big. If you know the VC dimension is 1, then the maximum possible signrank is 3. But for VC dimension 2, the answer is already much bigger:

Theorem 3.11 (Alon–Moran–Yehudayoff 2016)

The maximum possible signrank of a matrix with VC dimension 2 is $\widetilde{\Theta}(n^{1/2})$.

We won't talk about the upper bound, but the lower bound is almost just substituting Warren.

Proof of lower bound. We want to show that there are many sign matrices with VC dimension 2, and then use the same counting argument from before. The way to show that there are many such matrices is to take the incidence graph of the points and lines in a projective plane. This gives a matrix of signs that has roughly $n^{3/2}$ entries which are $+$, and all the rest are $-$; and it doesn't have any 2×2 submatrix of $+$'s. (Explicitly, in the matrix, rows correspond to points and columns to lines, with a $+$ for each incidence; there's no such 2×2 submatrix because two lines can't intersect at two different points.)

Now any matrix obtained from this specific matrix by changing some $+$'s to $-$'s will still not have a 2×2 submatrix of $+$'s. This means it cannot have VC dimension at least 3 — this would require having both the rows $(+++)$ and $(++-)$ in some columns, which would require two columns to have $(++)$ twice.

This gives roughly $2^{n^{3/2}}$ matrices with VC dimension 2. And we can use the same bound on the number of matrices with signrank d given by Warren's theorem (as in the proof of Theorem 3.7) to get

$$\left(\frac{8en^2}{2dn}\right)^{2dn} \geq 2^{n^{3/2}},$$

which gives the desired lower bound on d. □

§4 Orderings and sign patterns

What's the connection to our original problem? In this problem, we have a fixed set $C \subseteq \mathbb{R}^d$ with n points, and we let our k vantage points v_i vary and order C by $\sum_{i=1}^k \|v_i - c\|_2$. And we want to show this number of orderings is not too big.

So it makes sense to look at the following: Consider all $\binom{n}{2}$ functions of the form

$$(v_1, \ldots, v_k) \mapsto \sum_{i=1}^k \|v_i - c\|_2 - \sum_{i=1}^k \|v_i - c'\|_2$$

for all distinct $c, c' \in C$. So we're taking two candidate points c and c', and we're taking some $(v_1, \ldots, v_k)$; the sign of this function tells us whether its sum of distances to c or c' is smaller. Then the signs of all these $\binom{n}{2}$ functions will determine the ordering on C; and we want to bound the number of sign patterns of such functions.

If we were ordering by the sum of *squares* of Euclidean distances, everything would be great — these differences would be polynomials of degree 2 (actually they'd have degree 1, because the squares cancel), and we could plug in Warren's theorem and be happy.

But the trouble is that these functions are not polynomials — they're sums of *Square roots* of quadratic polynomials. So we'd maybe want a theorem that bounds the number of sign patterns of linear combinations of square roots (or more general radicals). Unfortunately, for such functions (linear combinations of square roots of quadratics), the number of sign patterns in general can be big:

Example 4.1

For every m, there is a family of m functions of just one variable, where each is a linear combination of square roots of quadratic polynomials, and still the number of sign patterns is 2^m (the maximum possible).

If we plugged in a formula that looked like Warren (Theorem 2.2) in this example, we'd have 1 variable and m functions, and the degrees would be 2; so we'd get something linear in m. So nothing like Warren can be true in general for linear combinations of square roots.

But the thing is that this example might not worry us too much, because each function is a linear combination of *many* square roots.

Question 4.2. What if our functions don't have too many square roots?

It turns out that then we *can* get a bound like Warren.

Theorem 4.3 (Alon–Defant–Kravitz–Zhu 2024+)
Let F be a collection of m real functions f_i of N real variables, each of the form $f_i = \sum_{j=1}^{r} a_{ij}(g_{ij})^{1/s}$ for some positive polynomials g_{ij} of degree at most Δ (in N variables). Then if $m \geq N$, we have

$$s(F) \leq \left(\frac{4e\Delta s^{r-2}m}{N}\right)^N.$$

In other words, each of our functions is a r-term linear combination of sth roots of polynomials. And we still get a bound of the same form as in Theorem 2.2, with an extra s^{r-2}; if r and s are not too big, this is still a good bound.

Proof sketch. The idea of the proof is a reduction to the polynomial setting, by using the conjugate trick from Galois theory. Our goal is to bound the number of connected components of the semivariety

$$\{x \in \mathbb{R}^d \mid f_i(x) \neq 0 \text{ for all } i\}.$$

So what we do is take each function f_i and multiply it by all its conjugates, which turns it into a polynomial.

Example 4.4
We have $(a-b)(a+b) = a^2 - b^2$, so if a and b are square roots of polynomials, then this will produce a polynomial.

Of course multiplying by something can only increase the number of connected components (everything that was 0 before remains 0; you have to make sure your functions f_i don't become identically 0, but this is true); and then Warren's theorem gives a bound you can plug in. □

In our case of $f(k,d,n)$, the number of variables is $N = dk$ (representing the coordinates of the k vantage points); the number of functions is $m = \binom{n}{2}$; and each is a linear combination of $r = 2k$ square roots (i.e., $s = 2$) of polynomials of degree $\Delta = 2$. So we get a bound of

$$\left(\frac{4e\Delta 2^{r-1}m}{N}\right)^N = O_{d,k}(n^{2dk}).$$

§5 Some open questions

Question 5.1. By Theorem 3.7, we know there are $n \times n$ matrices of signs that have signrank at least $\frac{n}{20}$, while the upper bound is roughly $\frac{n}{2}$. Which is right?

Noga would guess that the upper bound is the truth. This would be interesting — in the application to communication complexity the difference between $\frac{n}{20}$ and $\frac{n}{2}$ is not so big, but it's still a natural question.

Question 5.2. Suppose we order using the ℓ_p norm instead of the Euclidean norm. Is the maximum possible number of orderings still polynomial in n?

This proof still works if p is a rational number with bounded numerator and denominator. But what if $p = \pi$, for instance? It still looks like you probably can't get so many orders, but we don't know any bound whatsoever.

Question 5.3. For which configurations of points $C \subseteq \mathbb{R}^2$ (or more generally, $C \subseteq \mathbb{R}^d$) is it possible to get *all* permutations with some collection of points V, and how large should V be?

So here we're allowing the number of voters to be much bigger than the number of candidates; and we want to get all possible rankings.

It's not too hard to show that C has to be in convex position, but this is not sufficient. For the other direction, there's a classical result of Schoenberg (which isn't as known as it should be, and is really nice):

Theorem 5.4 (Schoenberg 1937)

For any finite number of points in any dimension, the matrix of Euclidean distances between them (i.e., the matrix $a_{ij} = \|x_i - x_j\|_2$, with 0's on the diagonal) is always nonsingular.

Using this, it can be shown that if C is *transitive* — meaning that there's an isometry of the space moving any point to any other — then you would be able to get all permutations. The idea is that we use Schoenberg to show that you can get all permutations if we allow negative voters as well. Then we add voters at every point to cancel out these negative voters and make everything positive. Transitivity means that if we add the same number of voters to each point, then we're adding the same constant to everything, so this doesn't affect the order.

And for how large V should be, we can use the counting argument to get a lower bound, and something about the *condition number* of such matrices to get an upper bound; both are polynomial in n, but we don't know the correct power.

Finally, here's a very innocent-looking question, which may belong to real algebraic geometry:

Question 5.5. Is the number of connected components in our proof of Theorem 4.3 tight? Specifically, if f is a linear combination of r square roots of positive quadratic polynomials in *one* variable, how many roots can it have?

So we're looking at the very special case where $m = 1$, $N = 1$, $s = 2$, and $\Delta = 2$, and we're interested in the dependence on r. We know that there can't be infinitely many roots (e.g., like a cosine function), because we proved a finite upper bound. But the upper bound we got is exponential in r; on the other side, you can give examples where the number of roots is $2r$. It looks like the answer should be linear in r. And this looks like a natural question you'd think people would've looked at, but we don't know the answer.