Rigidity of Vandermonde Matrices

Satyanarayana V. Lokam

Rigidity of matrix A : (Valiant '77)

Least number of entries of A to be changed to reduce its rank to $\leq r$.

$\mathcal{R}_A(r) :=$ min $\{|C| : \text{rank}(A - C) \leq r\}$, where

$|\cdot| :=$ Number of nonzero entries.

Trivial Upper Bound: For any $n \times n$ matrix A, $\mathcal{R}_A(r) \leq (n - r)^2$.

Theorem 1 (Valiant) *For almost all $n \times n$ A,*
$\mathcal{R}_A(r) \geq (n - r)^2$ *over an infinite field,*
$\mathcal{R}_A(r) \geq \Omega((n - r)^2/ \log n)$ *over a finite field.*

Open **Q**: Find EXPLICIT A s.t.
$\mathcal{R}_A(\epsilon n) \geq n^{1+\delta}$, for constants $\epsilon, \delta > 0$.

Motivation:

Theorem 2 (Valiant) $\mathcal{R}_A(\epsilon n) \geq n^{1+\delta} \Rightarrow$ $\mathcal{L}_A : x \longmapsto Ax$ *requires superlinear size log-depth linear circuits.*

Functions similar to rigidity also arise in several lower bound questions, eg. in two-party communication complexity.

Linear Circuits over $\mathbf{F}$

Input: $x \in \mathbf{F}^n$

Output: $y \in \mathbf{F}^m$

Gates: Each gate computes a
linear combination of its inputs.

Circuit computes $\mathcal{L}_A : x \longmapsto Ax$, $A \in \mathbf{F}^{m \times n}$.

Size: Number of edges

Depth: Longest input-output path

In this talk, $\mathbf{F} = \mathbf{C}$.

Candidates for High Rigidity

⋆ A generic (all a_{ij} algebraically independent)
$\Rightarrow$ $\mathcal{R}_A(r)$ is optimal $= (n-r)^2$
But, A is not EXPLICIT!

⋆ Vandermonde $V = \left(x_i^{j-1}\right)_{1 \le i,j \le n}$
(OPEN even with algebraically independent x_i)

⋆ Cauchy $C = ((x_i + y_j)^{-1})_{1 \le i,j \le n}$

⋆ Fourier Transform (Character Table)

⋆ Hadamard $(H \in \{-1,+1\}^n, HH^T = nI)$

⋆ Matrices from good codes

⋆ Incidence Matrix of a Projective Plane

Best known for explicit A: $\mathcal{R}_A(r) = \Omega\left(\frac{n^2}{r}\log\frac{n}{r}\right)$;
only $\Omega(n)$ for $r = \epsilon n$.

Generic Vandermonde Matrices

$V = (x_i^{j-1})_{1 \le i,j \le n}$, where x_i are algebraically independent.

- Expected to have high rigidity.
- We show $\mathcal{R}_V(r) = \Omega(n^2)$ if $r = O(\sqrt{n})$.

Open Q: Prove strong bounds for $r \ge \sqrt{n}$.

NOT *explicit*,
BUT provides

- n-dimensional set of matrices of high rigidity in $\mathbf{C}^{n \times n}$.
- First natural class with rigidity $\Omega(n^2)$ for nonconstant r.

The Shoup-Smolensky Dimension

For $A = (a_{ij}) \in \mathbf{C}^{n\times n}$,
$\mathcal{D}_A(t)$:= $\mathbf{Q}$-dimension of t-wise products of distinct a_{ij}.

Examples:

- A generic $\Rightarrow$ $\mathcal{D}_A(t) = \binom{n^2}{t}$.
- A generic Vandermonde $\Rightarrow$ $\mathcal{D}_A(t) \geq n^t$.

Shoup-Smolensky: $\mathcal{D}_A(\epsilon n) \geq n^{\delta n}$ $\Rightarrow$ $\mathcal{L}_A$ requires superlinear size linear circuits of log-depth.

Valiant: $\mathcal{R}_A(\epsilon n) \geq n^{1+\delta}$ $\Rightarrow$ $\mathcal{L}_A$ requires superlinear size linear circuits of log-depth.

Open Q: Relation between $\mathcal{D}$ and $\mathcal{R}$?

Theorem 3 *V: generic Vandermonde.*

$$\mathcal{R}_V(r) \geq n(n - c \cdot r^2)/2, \ \ c > 0 \textit{ a constant.}$$

In particular, if $r \leq \epsilon\sqrt{n}$, then $\mathcal{R}_V(r) \geq n^2/4$.

Proof: use Shoup-Smolensky dimension $\mathcal{D}$.
C: an optimal matrix of changes for r i.e., $\mathcal{R}_V(r) = |C|$, rank$(V - C) = r$.

Lower Bound: $$\mathcal{D}_{V-C}(n/2) \geq \left(n - \frac{2|C|}{n}\right)^{n/2}.$$

- Pick $n/2$ rows with $\leq 2|C|/n$ changes.
- Products formed by one unchanged entry per each of these rows are $\mathbf{Q}$-independent.

Upper Bound: $$\mathcal{D}_{V-C}(n/2) \leq \binom{nr + n/2}{n/2}^2.$$

Claim: rank$(A) \le r \;\Rightarrow\; \mathcal{D}_A(t) \le \binom{nr+t}{t}^2$.

$$A = \underbrace{P}_{n\times r} \cdot \underbrace{Q}_{r\times n}$$

$$\begin{aligned} \prod a &= \prod(\cdots + pq + \cdots) \\ &= \sum (p_1 \cdots p_t)\cdot(q_1 \cdots q_t) \end{aligned}$$

Hence,

$$\begin{aligned} \mathcal{D}_A(t) &\le \quad (\#\text{ deg-}t\text{ monomials of } p\text{'s}) \cdot \\ &\quad\quad (\#\text{ deg-}t\text{ monomials of } q\text{'s}) \\ &\le \binom{nr+t}{t}^2 \end{aligned}$$

■

Arbitrary Vandermondes

Theorem 4 (Shparlinsky '97; Pudlák '94)

$$\mathcal{R}_V(r) \geq (n-r)^2/(r+1) = \Omega\left(\frac{n^2}{r}\right).$$

Discrete Fourier Transforms

DFT $= (\zeta^{ij})_{0 \leq i,j \leq n-1}$,

where ζ is a primitive n^{th} root of unity.

Theorem 5 (Shokrollahi,Spielman,Stemann '97)
Every submatrix of A is nonsingular *and*
$\log^2 n \leq r \leq n/2$,
$\Rightarrow \ \mathcal{R}_A(r) = \Omega(n^2/r \ \log(n/r))$.

Corollary 6 *For prime n and $\log^2 n \leq r \leq n/2$,*

$$\mathcal{R}_{DFT}(r) = \Omega\left(\frac{n^2}{r} \log \frac{n}{r}\right).$$

– Best known bound for explicit Vandermonde!

Rigidity Bounds for Explicit Matrices

Reference	$\Omega(\cdot)$	Matrices
Razborov '89	$\frac{n^2}{r}$	$(\text{Vandermonde})^{-1}$ Fourier
Friedman '93	$\frac{n^2}{r} \log \frac{n}{r}$	Parity Check
Pudlák '94	$\frac{n^2}{r}$	Vandermonde Full Triangular Shifters
Shokrollahi, Spielman, & Stemann '97	$\frac{n^2}{r} \log \frac{n}{r}$	Cauchy Parity Check Fourier (*via* Coro.5)
Shparlinksy '97	$\frac{n^2}{r}$	Vandermonde
Kashin & Razborov '98	$\frac{n^2}{r}$	Hadamard

Common Proof Strategy

Step 1: most submatrices have a high rank.

Step 2: too few changes $\Rightarrow$ one such submatrix remains untouched.

We note: *This strategy cannot give better than an $\Omega(n^2/r \log n/r)$ bound on rigidity.*

For any $n \times n$ matrix A:

Greedy vertex cover (Lovász '75) $\Rightarrow$

$\exists\, S \subseteq$ entries of A, such that:

- $|S| = O(n^2/r \log n/r)$, and
- S intersects *every* $(r+1) \times (r+1)$ submatrix of A.

Hence,

any proof that relies on Step 2 <u>cannot</u> give superlinear lower bounds on $\mathcal{R}_A(\epsilon n)$.

As for Step 1 ...

Valiant '77: Linear size superconcentrators $\Rightarrow$ $\exists\ n \times n$ matrices A such that

- All submatrices of A are nonsingular, and yet,
- $\mathcal{R}_A(\epsilon n) \leq n^{1+o(1)}$.

Hence,

any proof that relies on Step 1 <u>cannot</u> give a lower bound of $\mathcal{R}_A(\epsilon n) \geq n^{1+\delta}$ for constants $\epsilon, \delta > 0$.

$\therefore$ Techniques quite different from those used in the past seem necessary to prove that $\mathcal{R}_A(\epsilon n) \geq n^{1+\delta}$ for explicit A.

Alternative Techniques

?

Codenotti, Pudlák, and Resta '98: Study graph-theoretic structure of low-rank matrices.

Odd Alternating Cycle Conjecture $\Rightarrow$ superlinear rigidity.