Matrix Rigidity

Mahdi Cheraghchi

mahdi.cheraghchi@epfl.ch

École Polytechnique Fédérale de Lausanne

February 16, 2005

Outline

Linear Programs

- A set of n indeterminates, called *input variables*.
- A set of m *output variables* (typically, $m = n$.).
- A set of *intermediate variables*.
- A sequence of *assignments*, each of the form $x \leftarrow \lambda u + \mu v$.
- No *multiple assignments* or *feedbacks* are allowed, i.e., any variable x occurring on the left-hand side of some assignment cannot occur in any assignment earlier.

Linear Programs

- A set of n indeterminates, called *input variables*.
- A set of m *output variables* (typically, $m = n$.).
- A set of *intermediate variables*.
- A sequence of *assignments*, each of the form $x \leftarrow \lambda u + \mu v$.
- No *multiple assignments* or *feedbacks* are allowed, i.e., any variable x occurring on the left-hand side of some assignment cannot occur in any assignment earlier.

Linear Programs

- Linear programs are also called *linear circuits*.
- We can associate a directed acyclic graph of fan-in 2 with any linear program.
- We put a node $\hat{u}$ for each variable u.
- For any assignment $x \leftarrow \lambda u + \mu v$, we add directed edges from $\hat{u}$ and $\hat{v}$ to $\hat{x}$, and label them with λ and μ, respectively.
- The node corresponding to the left hand side of the assignment acts as an *addition* gate.

Linear Programs

- Linear programs are also called *linear circuits*.
- We can associate a directed acyclic graph of fan-in 2 with any linear program.
- We put a node $\hat{u}$ for each variable u.
- For any assignment $x \leftarrow \lambda u + \mu v$, we add directed edges from $\hat{u}$ and $\hat{v}$ to $\hat{x}$, and label them with λ and μ, respectively.
- The node corresponding to the left hand side of the assignment acts as an *addition* gate.

Linear Programs

- Two complexity measures for linear programs:
  1. *Size*: Sum of the number of nodes and edges in the graph, or simply the number of edges. (Because of the bounded fan-in, they are of the same order.)
  2. *Depth*: Length of the longest path in the graph.
- There is a tradeoff between these two measures.

Linear Programs

- Any expression of the form $\sum_i \lambda_i x_i$ is called a *linear form*, where x_i are indeterminates and λ_i are constants over some field $\mathbb{F}$.
- A set of n linear forms over n indeterminates can be thought of as a matrix by vector multiplication.
- A linear program computes a set of linear forms.
- Thus, any linear program computes a matrix by vector multiplication.

Linear Programs

- Any expression of the form $\sum_i \lambda_i x_i$ is called a *linear form*, where x_i are indeterminates and λ_i are constants over some field $\mathbb{F}$.
- A set of n linear forms over n indeterminates can be thought of as a matrix by vector multiplication.
- A linear program computes a set of linear forms.
- Thus, any linear program computes a matrix by vector multiplication.

Linear Programs

- For computing a set of linear forms on certain fields (e.g., $\mathbb{R}$ and $\mathbb{C}$), linear programs are optimal within a constant factor compared with *straight-line programs*, where unrestricted use of algebraic operations $\{+, -, \times, \div\}$ is allowed.

Question

Find some explicit and natural family of linear forms that any linear program computing it needs a superlinear size.

Linear Programs

- For computing a set of linear forms on certain fields (e.g., $\mathbb{R}$ and $\mathbb{C}$), linear programs are optimal within a constant factor compared with *straight-line programs*, where unrestricted use of algebraic operations $\{+, -, \times, \div\}$ is allowed.

Question

Find some explicit and natural family of linear forms that any linear program computing it needs a superlinear size.

Linear Programs

Remark

This is a major open problem!

However, a few constructions have been suggested, e.g.,

- The family of linear forms defined by $n \times n$ matrices in which the entries are square roots of distinct primes. This needs a size of $\omega(n^2/\log n)$.
- Matrices defined over the rational in which the entries are very large integers.

An Easier Question

Is it possible to relate this computational problem to a combinatorial problem that is potentially more easily tractable?

Linear Programs

Remark

This is a major open problem!

However, a few constructions have been suggested, e.g.,

- The family of linear forms defined by $n \times n$ matrices in which the entries are square roots of distinct primes. This needs a size of $\omega(n^2/\log n)$.
- Matrices defined over the rational in which the entries are very large integers.

An Easier Question

Is it possible to relate this computational problem to a combinatorial problem that is potentially more easily tractable?

Linear Programs

Answer

Yes, the Matrix Rigidity problem!

Definition

The rigidity function of an $n \times n$ matrix A, $\mathcal{R}_A(r)$ is the minimum number of entries of A that must be changed to reduce the rank of A to r or less. [Valiant, 1977]

Another Definition

For any function $f(r)$, a directed acyclic graph $\mathcal{G}$ is called an $f(r)$-grate iff there exists two disjoint subsets A and B of the nodes in $\mathcal{G}$ such that if any set of r nodes of $\mathcal{G}$ are removed, at least $f(r)$ of the pairs in $A \times B$ remain connected.

Linear Programs

Answer

Yes, the Matrix Rigidity problem!

Definition

The rigidity function of an $n \times n$ matrix A, $\mathcal{R}_A(r)$ is the minimum number of entries of A that must be changed to reduce the rank of A to r or less. [Valiant, 1977]

Another Definition

For any function $f(r)$, a directed acyclic graph $\mathcal{G}$ is called an $f(r)$-grate iff there exists two disjoint subsets A and B of the nodes in $\mathcal{G}$ such that if any set of r nodes of $\mathcal{G}$ are removed, at least $f(r)$ of the pairs in $A \times B$ remain connected.

Linear Programs

Answer

Yes, the Matrix Rigidity problem!

Definition

The rigidity function of an $n \times n$ matrix A, $\mathcal{R}_A(r)$ is the minimum number of entries of A that must be changed to reduce the rank of A to r or less. [Valiant, 1977]

Another Definition

For any function $f(r)$, a directed acyclic graph $\mathcal{G}$ is called an $f(r)$-grate iff there exists two disjoint subsets A and B of the nodes in $\mathcal{G}$ such that if any set of r nodes of $\mathcal{G}$ are removed, at least $f(r)$ of the pairs in $A \times B$ remain connected.

Linear Programs

Theorem

The graph of any linear program for computing a set of linear forms Ax is an $\mathcal{R}_A(r)$-grate.

Another Theorem

For sufficiently large n (the number of input nodes), any $f(r)$-grate of depth $k \log_2 n$ with $f(n) > cn^{1+\epsilon}$ has superlinear size.

A Very Interesting Corollary!

Let $\{A_n\}$ be an infinite family of $n \times n$ matrices, and assume $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$, for some positive constants $\epsilon < 1$ and δ. Then no family of linear programs computing the corresponding set of linear forms can achieve linear size and logarithmic depth simultaneously.

Remark Most natural algorithms for computing interesting linear forms achieve a logarithmic depth. (e.g., FFT)

Linear Programs

Theorem

The graph of any linear program for computing a set of linear forms Ax is an $\mathcal{R}_A(r)$-grate.

Another Theorem

For sufficiently large n (the number of input nodes), any $f(r)$-grate of depth $k \log_2 n$ with $f(n) > cn^{1+\epsilon}$ has superlinear size.

A Very Interesting Corollary!

Let $\{A_n\}$ be an infinite family of $n \times n$ matrices, and assume $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$, for some positive constants $\epsilon < 1$ and δ. Then no family of linear programs computing the corresponding set of linear forms can achieve linear size and logarithmic depth simultaneously.

Remark Most natural algorithms for computing interesting linear forms achieve a logarithmic depth. (e.g., FFT)

Linear Programs

Theorem

The graph of any linear program for computing a set of linear forms Ax is an $\mathcal{R}_A(r)$-grate.

Another Theorem

For sufficiently large n (the number of input nodes), any $f(r)$-grate of depth $k \log_2 n$ with $f(n) > cn^{1+\epsilon}$ has superlinear size.

A Very Interesting Corollary!

Let $\{A_n\}$ be an infinite family of $n \times n$ matrices, and assume $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$, for some positive constants $\epsilon < 1$ and δ. Then no family of linear programs computing the corresponding set of linear forms can achieve linear size and logarithmic depth simultaneously.

Remark Most natural algorithms for computing interesting linear forms achieve a logarithmic depth. (e.g., FFT)

Rigid Matrices

Question

Do rigid matrices really exist? (i.e., $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$.)

- It's easy to see that for any $n \times n$ matrix A, $\mathcal{R}_A(r) \leqslant (n-r)^2$.
- But we need a lower bound!

Valiant's Theorem

Over any arbitrary field, most matrices are much more rigid than needed! In fact, over an infinite field, most matrices meet the upper bound; and over a finite field, most matrices get very close to it (up to a logarithmic factor).

Rigid Matrices

Question

Do rigid matrices really exist? (i.e., $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$.)

- It's easy to see that for any $n \times n$ matrix A, $\mathcal{R}_A(r) \leqslant (n-r)^2$.
- But we need a lower bound!

Valiant's Theorem

Over any arbitrary field, most matrices are much more rigid than needed! In fact, over an infinite field, most matrices meet the upper bound; and over a finite field, most matrices get very close to it (up to a logarithmic factor).

Rigid Matrices

Question

Do rigid matrices really exist? (i.e., $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$.)

- It's easy to see that for any $n \times n$ matrix A, $\mathcal{R}_A(r) \leqslant (n-r)^2$.
- But we need a lower bound!

Valiant's Theorem

Over any arbitrary field, most matrices are much more rigid than needed! In fact, over an infinite field, most matrices meet the upper bound; and over a finite field, most matrices get very close to it (up to a logarithmic factor).

Rigid Matrices

Question

Do rigid matrices really exist? (i.e., $\mathcal{R}_A(\epsilon n) = \Omega(n^{1+\delta})$.)

- It's easy to see that for any $n \times n$ matrix A, $\mathcal{R}_A(r) \leqslant (n-r)^2$.
- But we need a lower bound!

Valiant's Theorem

Over any arbitrary field, most matrices are much more rigid than needed! In fact, over an infinite field, most matrices meet the upper bound; and over a finite field, most matrices get very close to it (up to a logarithmic factor).

Rigid Matrices

A (Not-So-Complicated) Theorem

Let $\mathcal{S}$ be any set of $n \times n$ matrices over some field $\mathbb{F}_q$ such that $|\mathcal{S}| \geqslant q^{pn^2}$, for some $0 < p \leqslant 1$. For large enough n, there exists $A \in \mathcal{S}_n$ such that $\mathcal{R}_A(\epsilon n) = \omega(n^{1+\delta})$, for any $0 \leqslant \delta < 1$ and any $0 < \epsilon < \frac{p}{2}$. i.e., $\mathcal{S}$ contains matrices with rigidity arbitrary close to the upper bound.

Application 1. Matrices constructed by a large number of indeterminates are rigid.

Rigid Matrices

Application 2. *Evaluation* matrices of low degree bivariate polynomials.

Definition

Let $\{\alpha_1, \ldots, \alpha_q\}$ be all the elements of the field $\mathbb{F}_q$, and $P \in \mathbb{F}_q[x, y]$ be of degree $(d-1, d-1)$, for some $d \leqslant q$. P can be specified by its $d \times d$ *coefficient matrix*, $\mathcal{C}_P$. The *evaluation matrix* E is also defined as a $q \times q$ matrix that contains all possible evaluations of P, i.e., $e_{ij} = P(\alpha_i, \alpha_j)$.

Corollary

Let $\mathcal{S}$ be the set of the evaluation matrices of all bivariate polynomials of degree $(d-1, d-1)$. Then $|\mathcal{S}| = q^{d^2}$ and thus, for $d = pq$ $(0 < p \leqslant 1)$, $\mathcal{S}$ contains highly rigid matrices.

Rigid Matrices

Application 2. *Evaluation* matrices of low degree bivariate polynomials.

Definition

Let $\{\alpha_1, \ldots, \alpha_q\}$ be all the elements of the field $\mathbb{F}_q$, and $P \in \mathbb{F}_q[x, y]$ be of degree $(d-1, d-1)$, for some $d \leqslant q$. P can be specified by its $d \times d$ *coefficient matrix*, $\mathcal{C}_P$. The *evaluation matrix* E is also defined as a $q \times q$ matrix that contains all possible evaluations of P, i.e., $e_{ij} = P(\alpha_i, \alpha_j)$.

Corollary

Let $\mathcal{S}$ be the set of the evaluation matrices of all bivariate polynomials of degree $(d-1, d-1)$. Then $|\mathcal{S}| = q^{d^2}$ and thus, for $d = pq$ $(0 < p \leqslant 1)$, $\mathcal{S}$ contains highly rigid matrices.

Rigid Matrices

In fact, the set $\mathcal{S}$ we just defined is the code space of a degree two Reed-Muller code.

Not Really a Theorem

The evaluation and the coefficient matrices have the same rank, and thus, our set $\mathcal{S}$ of the evaluation matrices will not contain any rigid matrices when d is a sublinear function of q.

Rigid Matrices

- So far, we have just seen existence of rigid matrices, but not an explicit family.
- As it seems from the Valiant's theorem (that *nearly all* matrices are highly rigid), it must be so easy to find an explicit family of rigid matrices.
- It is not!! The problem is still open after nearly three decades.

Rigid Matrices

- So far, we have just seen existence of rigid matrices, but not an explicit family.
- As it seems from the Valiant's theorem (that *nearly all* matrices are highly rigid), it must be so easy to find an explicit family of rigid matrices.
- It is not!! The problem is still open after nearly three decades.

Variations of Rigidity

- *Norm Rigidity:* $\Delta_A^2(r) \stackrel{\text{def}}{=} \min_B \left\{ \sum_{i,j} |b_{ij}|^2 : \text{rank}(A+B) \leqslant r \right\}$.
- *Bounded Rigidity:*

$$\mathcal{R}_A(r,\theta) \stackrel{\text{def}}{=} \min_B \{\text{wt}(B) \mid \text{rank}(A+B) \leqslant r, \forall i,j : |b_{ij}| \leqslant \theta\}.$$

- *Strong Rigidity:* A is (k,t)-rigid if whenever no more than k entries in each row of A are altered, then A's rank remains at least t. [Friedman, 1993]
- Generalization of rigidity to *tensors* (three-dimensional matrices). [Pudlák and Rödl, 1993]
- ....

Highly Regular Matrices

Question

Are highly regular matrices good candidates for high rigidity?

Definition be patient...

A matrix is called *totally regular* iff all its minors have full rank.

Conjecture

Any totally regular matrix is highly rigid!

Fascinating Fact!

The conjecture is false! There exist nonrigid matrices constructed by linear size superconcentrators.

Highly Regular Matrices

Question

Are highly regular matrices good candidates for high rigidity?

Definition be patient. . .

A matrix is called *totally regular* iff all its minors have full rank.

Conjecture

Any totally regular matrix is highly rigid!

Fascinating Fact!

The conjecture is false! There exist nonrigid matrices constructed by linear size superconcentrators.

Highly Regular Matrices

Question

Are highly regular matrices good candidates for high rigidity?

Definition be patient. . .

A matrix is called *totally regular* iff all its minors have full rank.

Conjecture

Any totally regular matrix is highly rigid!

Fascinating Fact!

The conjecture is false! There exist nonrigid matrices constructed by linear size superconcentrators.

Rigidity of Highly Regular Matrices

- Call a matrix A *highly regular* iff for some constant c, the rank of any $r \times r$ minor of A is at least cr.
- Still it seems plausible to look for rigid examples among highly regular matrices.

A Nice Theorem:

For any highly regular matrix A, $\mathcal{R}_A(r) = \Omega\left(\frac{n^2}{r}\log\frac{n}{r}\right)$. (For any $\log^2 n \leqslant r \leqslant n/2$.) [Shokrollahi et. al., 1997]

Some explicit examples:

- Cauchy matrices over small fields. ($c_{ij} := 1/(x_i + y_j)$)
- Matrices obtained from asymptotically good algebraic geometric codes.
- The family of $p \times p$ discrete Fourier transform matrices, where p is a prime.

Rigidity of Highly Regular Matrices

- Call a matrix A *highly regular* iff for some constant c, the rank of any $r \times r$ minor of A is at least cr.
- Still it seems plausible to look for rigid examples among highly regular matrices.

A Nice Theorem:

For any highly regular matrix A, $\mathcal{R}_A(r) = \Omega\left(\frac{n^2}{r}\log\frac{n}{r}\right)$. (For any $\log^2 n \leqslant r \leqslant n/2$.) [Shokrollahi et. al., 1997]

Some explicit examples:

- Cauchy matrices over small fields. ($c_{ij} := 1/(x_i + y_j)$)
- Matrices obtained from asymptotically good algebraic geometric codes.
- The family of $p \times p$ discrete Fourier transform matrices, where p is a prime.

Rigidity of Highly Regular Matrices

- Call a matrix A *highly regular* iff for some constant c, the rank of any $r \times r$ minor of A is at least cr.
- Still it seems plausible to look for rigid examples among highly regular matrices.

A Nice Theorem:

For any highly regular matrix A, $\mathcal{R}_A(r) = \Omega\left(\frac{n^2}{r}\log\frac{n}{r}\right)$. (For any $\log^2 n \leqslant r \leqslant n/2$.) [Shokrollahi et. al., 1997]

Some explicit examples:

- Cauchy matrices over small fields. ($c_{ij} := 1/(x_i + y_j)$)
- Matrices obtained from asymptotically good algebraic geometric codes.
- The family of $p \times p$ discrete Fourier transform matrices, where p is a prime.

Rigidity of Highly Regular Matrices

(A Trivial) Theorem

Let A_q be a $q \times q$ minor of an $n \times n$ matrix A, picked uniformly at random. Assume that for any $1 \leqslant q \leqslant n$, $E[\text{rank}(A_q)] = \Omega(q)$. Then $\mathcal{R}_A(r) = \Omega(n^2/r)$. (For any $r \leqslant n/\epsilon$ and some constant $\epsilon > 1$.)

- An interesting example: The family of *generalized Hadamard matrices*. [Kashin and Razborov, 1998]
- A $n \times n$ generalized Hadamard matrix H_n is defined over the field $\mathbb{C}$ such that,
  1. $|h_{ij}| = 1$ for all $1 \leqslant i, j \leqslant n$.
  2. $HH^* = nI_n$ where I_n is the $n \times n$ identity matrix. That is, the rows (columns) are pairwise orthogonal.

Rigidity of Highly Regular Matrices

(A Trivial) Theorem

Let A_q be a $q \times q$ minor of an $n \times n$ matrix A, picked uniformly at random. Assume that for any $1 \leqslant q \leqslant n$, $E[\text{rank}(A_q)] = \Omega(q)$. Then $\mathcal{R}_A(r) = \Omega(n^2/r)$. (For any $r \leqslant n/\epsilon$ and some constant $\epsilon > 1$.)

- An interesting example: The family of *generalized Hadamard matrices*. [Kashin and Razborov, 1998]
- A $n \times n$ generalized Hadamard matrix H_n is defined over the field $\mathbb{C}$ such that,
  1. $|h_{ij}| = 1$ for all $1 \leqslant i, j \leqslant n$.
  2. $HH^* = nI_n$ where I_n is the $n \times n$ identity matrix. That is, the rows (columns) are pairwise orthogonal.

Rigidity of Highly Regular Matrices

(A Trivial) Theorem

Let A_q be a $q \times q$ minor of an $n \times n$ matrix A, picked uniformly at random. Assume that for any $1 \leqslant q \leqslant n$, $E[\text{rank}(A_q)] = \Omega(q)$. Then $\mathcal{R}_A(r) = \Omega(n^2/r)$. (For any $r \leqslant n/\epsilon$ and some constant $\epsilon > 1$.)

- An interesting example: The family of *generalized Hadamard matrices*. [Kashin and Razborov, 1998]
- A $n \times n$ generalized Hadamard matrix H_n is defined over the field $\mathbb{C}$ such that,
  1. $|h_{ij}| = 1$ for all $1 \leqslant i, j \leqslant n$.
  2. $HH^* = nI_n$ where I_n is the $n \times n$ identity matrix. That is, the rows (columns) are pairwise orthogonal.

Generalized Hadamard Matrices

Theorem (almost the last)

Let H be an $n \times n$ generalized Hadamard matrix. Then

1. $\mathcal{R}_H(r) \geqslant \Omega(\frac{n^2}{r})$,
2. *For any* $\theta \geqslant \frac{n}{r}$, $\mathcal{R}_H(r, \theta) \geqslant \Omega(\frac{n^3}{r\theta^2})$.
3. $\Delta_H(r) = n(n-r)$.

[Kashin and Razborov, 1998] [Lokam, 2001]

An interesting special case: The discrete Fourier transform matrix.

Vandermonde Matrices

$$V_n \stackrel{\text{def}}{=} \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix}$$

An interesting special case: Again, the discrete Fourier transform matrix, where $x_i = \zeta_n^{(i-1)}$.

Theorem (the very last one!)

If the x_i are restricted to be algebraically independent over $\mathbb{Q}$, then, for any arbitrary constant $\delta < 1$, there exists an $\epsilon > 0$ such that for every $r \leqslant \epsilon\sqrt{n}$, $\mathcal{R}_{V_n}(r) \geqslant \delta n^2$.
If the x_i are arbitrary but distinct, $\mathcal{R}_{V_n}(r) \geqslant (n-r)^2/(r+1)$.
[Lokam, 2000]

Conclusion

- We wish to prove nontrivial lower bounds on the complexity of linear circuits.
- Valiant proposed the concept of matrix rigidity, which reduces the above problem to a combinatorial property of the matrix defining the linear forms.
- However, nothing much is known: The major question of finding explicit families of rigid matrices remains open.
- Other explicit candidates for high rigidity can also be considered: Toeplitz matrix, circulant matrix, ...

Thank You!

Still Awake?!

References


P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic Complexity Theory,* Grundlehren der mathematischen Wissenschaften, Vol. 315, Springer, Berlin, 1997.


J. Friedman. A note on matrix rigidity. *Combinatorica,* 13(2);235-239, 1993.


B. Kashin and A. A. Razborov. Improved lower bounds on the rigidity of Hadamard matrices. *Mathematical Notes,* 63(4):471-475, 1998.


P. Kimmel and A. Settle. Reducing the rank of lower triangular all-ones matrices. *Technical Report,* CS 92-21, University of Chicago, 1992.


S. V. Lokam. Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity. *Journal of Computer and System Sciences,* 63:449-473, 2001.

References


S. V. Lokam. Note on the rigidity of Vandermonde matrices. *Theoretical Computer Science*, 237(1-2):477-483, 2000.


J. Morgenstern. The linear complexity of computation. *Journal of the ACM*, 22(2):184-194, 1975.


P. Pudlák and V. Rödl. Modified ranks of tensors and the size of circuits. In *Proceedings of the 25th Annual Symposium on Theory of Computing (STOC)*, pages 523-531, 1993.


M. A. Shokrollahi, D. A. Spielman, and V. Stemann. A remark on matrix rigidity. *Information Processing Letters*, 64(6):283-285, 1997.


L. G. Valiant. Graph theoretic arguments in low-level complexity. *Lecture Notes in Computer Science, Springer, Berlin*, 53:162-176, 1977.