УДК 519.95

ОБ УСТОЙЧИВЫХ МАТРИЦАХ

А. А. Разборов

Функцией устойчивости (rigidity) $(N\times N)$-матрицы A над некоторым полем k называется функция $R_A:\{1,\ldots,N\}\longrightarrow\{1,\ldots,N^2\}$, где $R_A(r)$ равно наименьшему возможному числу вхождений матрицы A, которые надо изменить с тем, чтобы получить матрицу ранга $\leq r$. В тех случаях, когда надо явно указать основное поле, мы используем запись $R_A^k(r)$.

Это понятие было введено Л. Вэлиентом [15, 16]. Независимо близкое понятие было предложено Д. Ю. Григорьевым [2]. Л. Вэлиент [16] и Д. Ю. Григорьев [3, §15] доказали, что достаточно высокие нижние оценки устойчивости $R_A(r)$ матрицы A для некоторых значений r влекут нелинейные нижние оценки сложности вычисления системы линейных форм AX (X - столбец неизвестных) в различных моделях вычислений. В частности, Л. Вэлиент [16] доказал, что для любой последовательности матриц $\{A_N\}$ (A_N - матрица размером $N\times N$) такой, что $R_{A_N}(N/2)\geq N^{1+\Omega(1)}$ соответствующая последовательность линейных форм $A_N X_N$ не вычисляется арифметическими схемами, имеющими одновременно размер $O(N)$ и глубину $O(\log N)$.

Другой отправной точкой для написания настоящей работы послужило понятие сложности комбинаторных объектов (graph complexity), введенное в [12, 5]. Главная идея, лежащая в основе этого понятия, заключается в предложении изучать сложность вычисления объектов более симметричных по сравнению с булевыми функциями (например, ориентированных, неориентированных или двудольных графов) вместо сложности вычисления самих булевых функций. При этом для каждого сложностного класса в теории булевой

сложности возникает соответствующий ему класс сложности комбинаторных объектов.

Сложность комбинаторных объектов оказалась также тесно связанной с коммуникационной сложностью. В работе [7] были определены классы коммуникационной сложности, аналогичные основным классам Тьюринговой сложности: $P, NP, PH, PSPACE$ и т. д. Оказалось, что эти классы допускают альтернативное определение в терминах сложности комбинаторных объектов (в данном случае - двудольных графов). Например, коммуникационный аналог PH^{CC} класса PH состоит из последовательностей двудольных графов (или, что то же самое, 0-1-матриц), вычислимых схемами в стандартном базисе глубины $O(1)$ и размера $\exp((\log\log N)^{O(1)})$ (N - размер матрицы); $PSPACE^{CC}$ - глубины $(\log\log N)^{O(1)}$ и размера $\exp((\log\log N)^{O(1)})$ и т. д. [7]. Таким образом, ряд задач коммуникационной сложности сводится к аналогичным задачам для сложности двудольных графов.

Поскольку к задачам получения нижних оценок сложности двудольных графов сводятся также задачи получения нижних оценок в булевой сложности, можно ожидать, что первые окажутся труднее последних. Это в самом деле так: в настоящий момент не известны никакие нетривиальные оценки сложности явно заданных графов даже для схем ограниченной глубины, хотя в булевой сложности такие оценки хорошо известны [8, 6, 17, 9, 4, 14, 10].

Первый результат настоящей работы, доказанный в §1, говоря неформально, утверждает, что 0-1 матрицы, достаточно устойчивые в смысле Вэлиента-Григорьева над полем характеристики p, не могут быть вычислены схемами ограниченной глубины и малого размера в базисе $\{\&, \vee, \neg, MOD\text{-}q\}$, если q - степень p. Более точно, если $\mathrm{char}\, k = p$ и $R_A^k(r) \geq \dfrac{N^2}{\exp((\ln r)^{1\setminus d+1})}$, то $C_d^q(A) \geq \exp(\Omega((\log r)^{1\setminus d+1}))$, где $C_d^q(A)$ -

сложность реализации $(N\times N)$-матрицы A в указанном базисе ($r=r(N)$ - произвольно). Из сделанных выше замечаний вытекает также, что из $R_A^k(r)\geq\frac{N^2}{\exp((\log r)^{o(1)})}$ для некоторых $r\geq\exp((\log\log N)^{\omega(1)})$ и некоторого конечного поля k следует $\{A_N\}\notin PH^{CC}$, т.е. достаточно устойчивые матрицы не могут распознаваться полиномиальным коммуникационным алгоритмом с фиксированным числом альтераций.

Вместе с предыдущими работами Вэлиента и Григорьева, этот результат еще раз подчеркивает важность построения явно заданных устойчивых матриц. Вэлиент [15, 16] доказал, что для почти всех A ("почти всех" понимается в смысле мощности, если k конечно и в смысле топологии Зарисского, если k бесконечно) имеет место оценка

$$R_A^k(N/2)\geq\Omega(N^2). \qquad (1)$$

По-видимому, единственная известная до настоящего времени нетривиальная нижняя оценка функции устойчивости явно заданных матриц принадлежит П. Пудлаку, П. Савицкому [13] и автору. Именно, П. Пудлак и П. Савицкий доказали, что если A - матрица Адамара порядка N, то $R_A^{\mathbf{R}}(r)\geq\Omega(N^2/r^4(\log r)^2)$ ($\mathbf{R}$ - поле вещественных чисел). А. Разборов усилил эту оценку до $R_A^{\mathbf{R}}(r)\geq\Omega(N^2/r^3\log r)$.

В §2 настоящей работы мы доказываем новые результаты в этом направлении. Именно, если A - матрица, обратная к матрице Вандермонда над любым полем или матрица обобщенного преобразования Фурье (ОПФ), соответствующего любой абелевой группе над любым полем разложения этой группы, то при любом $r\leq\varepsilon N$ имеет место оценка

$$R_A(r)\geq\Omega(N^2/r). \qquad (2)$$

Заметим, что матрицы (обычного) дискретного преобразования Фурье (ДПФ) являются частным случаем как матриц Вандермонда (и, с точностью до постоянного множителя, совпадают со своими обратными), так и матриц ОПФ. Кроме того, для специального случая матриц

Адамара, а именно, матриц Сильвестра, являющихся также матрицами ОПФ группы $\mathbb{Z}_2^n$, (2) обобщает отмеченный выше результат Пудлака и Савицкого.

При доказательстве оценки (2) используется некоторое сведение к задачам получения нижних оценок сложности вычисления систем билинейных форм арифметическими схемами, в которых всякий раз при выполнении умножения по крайней мере один из входов обязан быть либо константой либо переменной. По аналогии с булевым случаем мы называем такие схемы контактными (switching). Несмотря на кажущуюся простоту таких схем, получение для них нижних оценок, близких к оптимальным, наталкивается на значительные трудности. Мы формулируем одну гипотезу о трудновычислимости семейств булевых форм некоторого вида, утвердительный ответ на которую позволил бы для матриц ОПФ и обратных к матрицам Вандермонда усилить оценку (2) и извлечь отсюда многочисленные следствия, среди которых невычислимость ДПФ схемами размера $O(N)$ и глубины $O(\log N)$, а также $PH^{CC} \neq PSPACE^{CC}$.

Д.Ю.Григорьев и Н.Нисан независимо заметили, что оценка (2) справедлива для произвольных матриц A, у которых все миноры невырождены. Назовем такие матрицы вполне невырожденными. Д.Ю.Григорьев также указал, что матрицы Коши с общим членом $a_{ij}=(x_i-y_j)^{-1}$, где $x_1,\ldots,x_N,y_1,\ldots,y_N$ - произвольные попарно различные элементы основного поля, являются вполне невырожденными. М.А.Фрумкин указал на другой пример таких матриц: именно, так называемые обобщенные матрицы Вандермонда с общим членом $a_{ij}=x_i^{(y_j)}$ для случая, когда $x_1,\ldots,x_N,y_1,\ldots,y_N$ - попарно различные положительные вещественные числа также являются вполне невырожденными [1,с.372].

Отметим, однако, что над любым конечным полем вполне

невырожденные матрицы имеют ограниченный размер (например, в силу теоремы Рамсея). Поэтому развиваемый в §2 настоящей работы метод, по-видимому, является единственным существующим подходом к получению нетривиальных нижних оценок функции устойчивости над конечными полями.

§1. УСТОЙЧИВЫЕ МАТРИЦЫ ТРУДНОВЫЧИСЛИМЫ С ПОМОЩЬЮ СХЕМ ОГРАНИЧЕННОЙ ГЛУБИНЫ

На протяжении всего параграфа понятие функциональной схемы используется в обычном смысле с тем исключением, что входам схем поставлены в соответствие не булевы переменные, а 0-1 матрицы фиксированного размера, множество единиц которых является прямоугольником, т. е. имеет вид $I\times J$. Булевы операции над матрицами производятся поэлеменентно. Размером схемы называется число вершин в ней; глубиной - длина максимального пути, ведущая от входов к выходу. Во всех рассматриваемых в настоящем параграфе схемах степень захода вершин (fan-in) неограничена. Булева функция $MOD\text{-}q$ определяется равенством $MOD\text{-}q(x_1,\ldots,x_n)=1\Leftrightarrow q\mid \sum_{i=1}^{n} x_i$. Через $C_d(A)$ обозначаем сложность вычисления матрицы A схемами только что описанного вида в базисе $\{\&,\vee,\neg\}$; через $C_d^q(A)$ - в базисе $\{\&,\vee,\neg,MOD\text{-}q\}$.

Через $|B|$ обозначается число ненулевых вхождений в матрицу B. Функцией устойчивости (rigidity) $(N\times N)$-матрицы A над некоторым полем k называется функция $R_A^k:\{1,\ldots,N\}\longrightarrow\{1,\ldots,N^2\}$, где $R_A^k(r)=\min\{|B| \mid \mathrm{rank}(A-B)\le r\}$ [15, 16]. Основным результатом настоящего параграфа является следующая теорема.

ТЕОРЕМА 1. Пусть $k=F_p$; $q=p^m$ и d фиксированы; A - 0-1 матрица

размера $N\times N$. Тогда при условии

$$R_A^k(r)\geq\frac{N^2}{\exp((\ln r)^{1\setminus d+1})} \qquad (3)$$

имеет место оценка

$$C_d^q(A)\geq\exp(\Omega(\log r)^{1\setminus d+1}).$$

ДОКАЗАТЕЛЬСТВО. (ср. с доказательством аппроксимационной леммы в [4, 14]). Положим $l=\lfloor 2(\ln r)^{1\setminus d+1}\rfloor$. Заметим, что при $r=O(1)$ утверждение теоремы очевидно, поэтому мы можем считать, что r и l достаточно велики.

Пусть дана некоторая функциональная схема над базисом $\{\&,\vee,\neg,MOD\text{-}q\}$ глубины $\leq d$ и размера s, вычисляющая матрицу A. Нам надо доказать, что

$$s\geq\exp(\Omega(\log r)^{1\setminus d+1}). \qquad (4)$$

Разместим вершины данной схемы по d уровням так, что в каждую вершину на i-м уровне ведут ребра лишь из вершин меньших уровней. Пусть $A(v)$ - матрица, вычисляемая в вершине v. Индукцией по i припишем каждой вершине v i-го уровня некоторую 0-1 матрицу $\overline{A}(v)$ так, что выполнены следующие два условия:

а) $\operatorname{rank}_k\overline{A}(v)\leq(s+1)^{O(l^i)}$,

б) если v - вход, то $\overline{A}(v)=A(v)$. В противном случае, если $A(v)=A(v_1)*\ldots*A(v_h)$ - операция, производимая в вершине v, то $|\overline{A}(v)-\overline{A}(v_1)*\ldots*\overline{A}(v_h)|\leq N^2p^{-1}$.

<u>База</u> <u>$i=0$</u> очевидна: все вершины 0-го уровня являются входами и мы просто полагаем $\overline{A}(v)=A(v)$; заметим, что $\operatorname{rank}_k A(v)\leq 1$.

<u>Шаг индукции</u>. Пусть v - вершина i-го уровня; $A(v)=A(v_1)*\ldots*A(v_h)$ и $\overline{A}(v_1),\ldots,\overline{A}(v_h)$ уже определены. Разберем четыре возможных случая, соответствующих операции *.

Случай $*=\neg$ очевиден - полагаем $\overline{A}(v)=\neg\overline{A}(v_1)$; при этом $\mathrm{rank}_k\overline{A}(v)\le\mathrm{rank}_k\overline{A}(v_1)+1$, и оценка ранга а) вытекает из $1=\omega(1)$.

Случай $*=MOD\text{-}q$. Здесь мы используем известную AC^0-сводимость $MOD\text{-}q$ к $MOD\text{-}p$, а именно,

$$MOD\text{-}q(x_1,\ldots,x_n)=\bigwedge_{j=0}^{m-1} MOD\text{-}p(\{x_I\mid I\subseteq\{1,\ldots,n\},\ |I|=p^j\}), \qquad (5)$$

где $x_I=\&_{i\in I}x_i$. Формула (5) следует из того наблюдения, что a делится на p^m тогда и только тогда, когда все биномиальные коэффициенты $\binom{a}{1}$, $\binom{a}{p}$, ..., $\binom{a}{p^{m-1}}$ делятся на p.

Полагаем $\overline{A}(v)=MOD\text{-}q(\overline{A}(v_1),\ldots,\overline{A}(v_h))$. Теперь для получения верхней оценки на $\mathrm{rank}_k\overline{A}(v)$ достаточно воспользоваться оценками

$$\mathrm{rank}_k(B_1\&\ldots\&B_n)\le\mathrm{rank}_k(B_1)\cdot\ldots\cdot\mathrm{rank}_k(B_n), \qquad (6)$$

$$\mathrm{rank}_k(MOD\text{-}p(B_1,\ldots,B_n))\le(\mathrm{rank}_k(B_1)+\ldots+\mathrm{rank}_k(B_n))^{p-1}, \qquad (7)$$

первая из которых вытекает из того факта, что операция &, определенная на 0-1 матрицах, совпадает с операцией $\circ$ покомпонентного умножения матриц, а последняя билинейна на пространстве строк. Оценка (7) следует из формулы $MOD\text{-}p(B_1,\ldots,B_n)=(B_1+\ldots+B_n)^*$, где $C^*=C\circ\ldots\circ C$ ($(p-1)$ раз). Применяя (6), (7) и индуктивное предположение к правой части (5), получим $\mathrm{rank}_k(MOD\text{-}q(\overline{A}(v_1),\ldots,\overline{A}(v_h))\le (h\cdot\max(\mathrm{rank}\overline{A}(v_1),\ldots,\mathrm{rank}\overline{A}(v_h)))^{O(1)}\le \left(h\cdot(s+1)^{O(l^{i-1})}\right)^{O(1)}\le(s+1)^{O(l^i)}$, т.к. $h\le s$.

Случай $*=\vee$. Пусть L - линейное пространство над $\mathbf{F}_p$, порожденное матрицами $\overline{A}(v_1),\ldots,\overline{A}(v_h)$. Тогда

$$\forall B\in L\ \mathrm{rank}B\le h\cdot(s+1)^{O(l^{i-1})}. \qquad (8)$$

Если $\mathbf{B}$ - случайный элемент из L, то всякий раз, когда некоторое вхождение $\overline{a}_{ij}$ хотя бы одной из матриц $\overline{A}(v_1),\ldots,\overline{A}(v_h)$ обращается в единицу, линейный функционал на L, заданный этим вхождением,

невырожден и мы имеем $\mathbf{P}[b_{ij}=0]=1/p$. Переходя к математическим ожиданиям, получаем, что можно выбрать $B_1,\ldots,B_l\in L$ так, что импликация

$$\bar{A}_{ij}(v_1)\ \ldots\ \bar{A}_{ij}(v_h)=1 \Rightarrow \exists 1\le\nu\le l\,(B_\nu)_{ij}\neq 0$$

выполнена для всех, кроме $\le N^2p^{-1}$ вхождений (i,j). Полагаем $\bar{A}(v)=B_1^*\vee\ldots\vee B_l^*$. Тогда б) вытекает из только что сделанного замечания. Что касается оценки а) для ранга, то в силу (8), (6), $h\le s$ и законов де Моргана, получаем $\mathrm{rank}\bar{A}(v)\le 1+\left(1+\left(h\cdot(s+1)^{O(l^{i-1})}\right)^{p-1}\right)^l\le (s+1)^{O(l^i)}$.

Случай $*=\&$ сводится к только что разобранному с помощью законов де Моргана.

Итак, матрицы $\bar{A}(v)$, удовлетворяющие свойствам а) и б), построены. Пусть, в частности, $\bar{A}$ - матрица, соответствующая выходу. Если $\mathrm{rank}\bar{A}\ge r$, то из оценки ранга а) мы имеем $r\le(s+1)^{O(l^d)}$, откуда непосредственно вытекает (4). Если же $\mathrm{rank}\bar{A}<r$, то в силу (3), $|A-\bar{A}|\ge R_A^k(r)\ge\dfrac{N^2}{\exp((\ln r)^{1\setminus d+1})}$. С другой стороны, в силу б), $|A-\bar{A}|\le s\cdot N^2\cdot p^{-1}$, так как каждой позиции, в которой A и $\bar{A}$ различаются, соответствует хотя бы одна вершина v такая, что $\bar{A}(v)$ и $\bar{A}(v_1)*\ldots*\bar{A}(v_h)$ различаются в этой позиции. Сравнивая эти две оценки, мы получаем (4).■

Бабаи, Франкл и Симон [7] рассмотрели класс PH^{CC}, состоящий из матриц $\{A_N\}$, распознаваемых полиномиальными (относительно $\log N$) коммуникационными алгоритмами с ограниченным числом альтераций и доказали, что PH^{CC} совпадает с классом матриц $\{A_N\}$ таких, что $C_{O(1)}(A_N)\le\exp((\log\log N)^{O(1)})$. Из теоремы 1 непосредственно вытекает следующее следствие.

СЛЕДСТВИЕ 1. Если $\{A_N\}$ - последовательность 0-1 матриц (A_N - матрица размера $N\times N$); $\exp((\log\log N)^{\omega(1)})\le r=r(N)\le N$ и для некоторого конечного поля k

$$R_A^k(r)\ge\frac{N^2}{\exp((\log r)^{O(1)})}\,, \qquad (9)$$

то $\{A_N\}\notin PH^{CC}$. ■

ЗАМЕЧАНИЕ 1. По аналогии с булевым случаем, класс AC^0 состоит из тех матриц $\{A_N\}$, для которых $C_{O(1)}(A_N)\le(\log N)^{O(1)}$. Очевидно, $AC^0\subseteq PH^{CC}$, поэтому следствие к теореме 1 применимо также и к AC^0. Любопытно отметить, что теорема 1 не улавливает разницы между классами AC^0 и PH^{CC} (если, конечно, она вообще существует!), т.е. никакого утверждения для AC^0, более сильного, чем это следствие, из теоремы 1 вывести не удается.

ЗАМЕЧАНИЕ 2. В [5] было показано, что для целого ряда свойств графов, двудольных графов, турниров (свойства рамсеевского типа, экстремальные свойства, свойства ассиметричности), присущих случайным объектам, существуют объекты, обладающие этими свойствами, матрицы смежности которых принадлежат $(AC^0)^{\oplus}$ $\left(=\left\{\{A_N\}\,|\,C^2_{O(1)}(A_N)\le(\log N)^{O(1)}\right\}\right)$. Из теоремы 1 вытекает, что при $\exp((\log\log N)^{\omega(1)})\le r=r(N)\le N$ над полем $k=\mathbf{F}_2$ имеет место верхняя оценка устойчивости матриц смежности A_N таких объектов $R^k_{A_N}(r)\le\frac{N^2}{\exp((\log r)^{\Omega(1)})}$. Иными словами, матрицы могут быть достаточно неустойчивыми над $\mathbf{F}_2$ и в то же время обладать рядом свойств, характерных для случайных матриц.

ЗАМЕЧАНИЕ 3. Наиболее естественный кандидат на принадлежность классу $PSPACE^{CC}\setminus PH^{CC}$ - это предикат "СКАЛЯРНОЕ ПРОИЗВЕДЕНИЕ ПО mod2" [7]. Из следствия к теореме 1 вытекает, что для доказательства факта

$PSPACE^{CC} \neq PH^{CC}$ достаточно установить оценку (9) при каких-либо $\exp((\log\log N)^{\omega(1)}) \le r = r(N) \le N$ для матриц скалярного произведения в конечномерных евклидовых пространствах над $\mathbf{F}_2$ (N - число элементов пространства!). В следующем параграфе мы докажем, что для любых $r \le \varepsilon N$ (ε - достаточно малая положительная константа) и любого поля k характеристики, отличной от 2, имеет место более слабая оценка $R_A(r) \ge \Omega(N^2/r)$.

§2. НИЖНИЕ ОЦЕНКИ ДЛЯ ФУНКЦИИ УСТОЙЧИВОСТИ

Пусть k - произвольное поле. Если k является полем разложения некоторой конечной абелевой группы G, то матрица обобщенного преобразования Фурье, соответствующего группе G, имеет вид $||\chi(g)||$, где g пробегает группу G, а χ - двойственную группу $\hat{G} = \mathrm{Hom}(G, k^*)$ (изоморфную (неканонически!) группе G). Матрицей Вандермонда называется матрица вида $||a_i^j||$, где $a_0, \ldots, a_{N-1}$ - попарно различные элементы из k; $0 \le j \le N-1$. Основным результатом настоящего параграфа являются следующие две теоремы.

ТЕОРЕМА 2. Пусть A - матрица обобщенного преобразования Фурье некоторой абелевой группы порядка N и $r \le \varepsilon N$ (ε - достаточно малая положительная константа). Тогда имеет место оценка $R_A^k(r) \ge \Omega(N^2/r)$.

ТЕОРЕМА 3. Пусть k бесконечно, A - матрица, обратная к некоторой матрице Вандермонда порядка N и $r \le \varepsilon N$ (ε - достаточно малая положительная константа). Тогда имеет место оценка $R_A^k(r) \ge \Omega(N^2/r)$.

Для доказательства этих теорем мы разовьем метод получения нижних оценок сложности вычисления систем билинейных форм в

некоторой ограниченной модели вычислений. По-видимому, эта модель представляет и самостоятельный интерес.

Рассмотрим арифметические схемы с множеством допустимых операций, состоящим из сложения и умножения вычисленного ранее полинома на константу или переменную. При этом умножение на переменную имеет стоимость 1; остальные операции - бесплатные. Отметим, что если заменить сложение на дизъюнкцию, а умножение - на коньюнкцию, то мы придем к понятию контактных упорядоченных схем (directed switching networks) [11]. По этой причине мы называем арифметические схемы рассматриваемого вида контактными и обозначаем соответствующую меру сложности через S.

Если $L_1,\ldots,L_N$ - семейство билинейных форм, то $S(L_1,\ldots,L_N)$ легко характеризуется через представляющий тензор семейства $L_1,\ldots,L_N$. Пусть P,Q,R - конечномерные пространства, причем в P и Q фиксированы базисы E и F соответственно. Назовем контактным рангом $\mathrm{rk}_S(t)$ тензора $t\in P\otimes Q\otimes R$ наименьшее возможное число тензоров вида

$$e\otimes q\otimes r,\quad e\in E, q\in Q, r\in R \tag{10}$$

или

$$p\otimes f\otimes r,\quad p\in P, f\in F, r\in R, \tag{11}$$

сумма которых равна t. Для семейства $\{L_1,\ldots,L_N\}=\left\{\sum\limits_{i,j} a_{ijk}x_iy_j \mid 1\le k\le N\right\}$ билинейных форм через $t(L_1,\ldots,L_N)$ обозначим его представляющий тензор $\sum\limits_{i,j,k} a_{ijk}e_i\otimes f_j\otimes g_k$. Следующая простая теорема доказывается точно так же, как и ее аналог для случая неограниченных арифметических схем (см., например, [3]).

ТЕОРЕМА 4. $S(L_1,\ldots,L_N)=\mathrm{rk}_S(t(L_1,\ldots,L_N))$. ■

Предположим теперь, что в пространстве R также фиксирован некоторый базис G. Для тензора $t=\sum_{e,f,g} a_{efg}\, e\otimes f\otimes g$ [соответственно, тензора $c=\sum_{e,f} a_{ef}\, e\otimes f$] через $\mathrm{supp}(t)$ [соответственно, $\mathrm{supp}(c)$] обозначим множество $\{\langle e,f,g\rangle \mid a_{efg}\neq 0\}$ [соответственно, $\{\langle e,f\rangle \mid a_{ef}\neq 0\}$]. Тензор t [тензор c] называется диагональным, если всякие два различных элемента из $\mathrm{supp}(t)$ [$\mathrm{supp}(c)$ соответственно] отличаются по меньшей мере в двух местах. Для $S\subseteq E\times F\times G$ и $g\in G$ положим $S_g=\{\langle e,f\rangle\in E\times F \mid \langle e,f,g\rangle\in E\times F\times G\}$. S назовем однородным, если $|S_g|$ не зависит от g.

В дальнейшем нам понадобятся нижние оценки несколько более общего вида, нежели просто оценки на величину rk_S. Именно, пусть $\mathrm{rk}_S(t)\leq(r_P,r_Q)$ по определению обозначает, что тензор t может быть представлен в виде суммы не более r_P тензоров вида (10) и r_Q тензоров вида (11). В следующей теореме устанавливаются нижние оценки для произведения $r_P r_Q$ для широкого класса тензоров t.

ТЕОРЕМА 5. Пусть $t\in P\otimes Q\otimes R$ - диагональный тензор; $S\subseteq\mathrm{supp}(t)$ однородно. Допустим, что $\mathrm{rk}_S(t)\leq(r_P,r_Q)$ и $r_P,r_Q\leq\varepsilon|S|$ (ε - достаточно малая положительная константа). Тогда $r_P r_Q\geq\Omega(\dim R\cdot|S|)$.

В частности, $\mathrm{rk}_S(t)\geq\Omega((\dim R\cdot|S|)^{1/2})$.

ДОКАЗАТЕЛЬСТВО. Положим $d_P=\left[\left(\frac{r_P}{r_Q}\cdot\frac{|S|}{\dim R}\right)^{1/2}\right]$; $d_Q=\left[\left(\frac{r_Q}{r_P}\cdot\frac{|S|}{\dim R}\right)^{1/2}\right]$.

Разберем вначале случай $d_P\geq 1$, $d_Q\geq 1$.

Расширяя в случае необходимости k, можно считать, что k содержит элементы x_{ei}, y_{fj} ($e\in E$, $f\in F$, $1\leq i\leq d_P$, $1\leq j\leq d_Q$), трансцендентные над полем определения тензора t. Пусть P_0 и Q_0 - подпространства в P и Q соответственно, порожденные векторами $\sum_e x_{ei}\, e$

$(1\le i\le d_P)$ и $\sum_f y_{fj} f$ $(1\le j\le d_Q)$ соответственно. Превратим P и Q в ассоциативные алгебры с операцией умножения $\circ$, полагая $e_\nu\circ e_\gamma=\delta_{\nu\gamma}e_\nu$ (соответственно, $f_\nu\circ f_\gamma=\delta_{\nu\gamma}f_\nu$). Пусть $U=P_0\otimes Q_0\otimes R^*$. Рассмотрим билинейное отображение $\mu:(P\otimes Q\otimes R)\otimes U\longmapsto P\otimes Q$, являющееся тензорным произведением ограничений на $P\otimes P_0$, $Q\otimes Q_0$ введенного выше умножения $\circ:P\otimes P\longmapsto P$, $\circ:Q\otimes Q\longmapsto Q$ и естественного спаривания $R\otimes R^*\longmapsto k$. Изучим поведение величины $\dim\mu(a\otimes U)$ для различных тензоров a.

Пусть вначале тензор a имеет вид (10), т.е. $a=e\otimes q\otimes r$. Тогда $\mu(a\otimes U)\le(e\circ P_0)\otimes(q\circ Q_0)\le(ke)\otimes(q\circ Q_0)$, откуда $\dim\mu(a\otimes U)\le d_Q$. Аналогично, $\dim\mu(a\otimes U)\le d_P$ для тензоров a вида (11). Так как по условию t представляется в виде суммы $\le r_P$ тензоров вида (10) и $\le r_Q$ тензоров вида (11), то отсюда

$$\dim\mu(t\otimes U)\le d_P r_Q+d_Q r_P\le 2\left(\frac{r_P r_Q|S|}{\dim R}\right)^{1/2}. \qquad (12)$$

Оценим теперь $\dim\mu(t\otimes U)$ снизу. Для этого разложим t по базису G пространства R: $t=\sum_g c_g\otimes g$; $c_g\in P\otimes Q$. Прежде всего, $\mu(t\otimes U)=\mu\left(t\otimes(P_0\otimes Q_0)\otimes R^*\right)$ является подпространством в $P\otimes Q$, порожденном подпространствами $V_g=c_g\circ(P_0\otimes Q_0)$. Так как t диагонален, то $\mathrm{supp}(c_{g_1})\cap\mathrm{supp}(c_{g_2})=\varnothing$ при $g_1\ne g_2$ и, значит, $\mathrm{supp}(V_{g_1})\cap\mathrm{supp}(V_{g_2})=\varnothing$. Поэтому разложение $\mu(t\otimes U)=\sum_{g\in G}V_g$ прямое и, следовательно,

$$\dim\mu(t\otimes U)=\sum_{g\in G}\dim V_g. \qquad (13)$$

Докажем, что для всех g

$$\dim V_g=d_P d_Q. \qquad (14)$$

Для этого вначале отметим, что $d_P d_Q\le\frac{|S|}{\dim R}=|S_g|$, так как S однородно. Выберем $S'_g\subseteq S_g$ так, что $|S'_g|=d_P d_Q$ и проверим, что проекция V_g на S'_g взаимно-однозначно отображает V_g на подпространство всех двумерных

тензоров с носителями в S'_g.

В самом деле, так как t диагонален, то S'_g имеет вид $\left\{<e_\nu, f_\nu> \mid 1\le\nu\le d_P d_Q\right\}$, где все e_ν, f_ν попарно различны. Пусть α_ν - коэффициент тензора c_g при $e_\nu\otimes f_\nu$ $(1\le\nu\le d_P d_Q)$. Тогда в естественных координатах отображение

$$P_0\otimes Q_0 \xrightarrow{\circ c_g} V_g \xrightarrow{\text{проекция на } S'_g} kS'_g \qquad (15)$$

задается квадратной матрицей H, строки которой индексированы парами (i,j) $(1\le i\le d_P,\ 1\le j\le d_Q)$, а столбцы - индексом ν с общим членом $h_{ij\nu}=\alpha_\nu x_{e_\nu i} y_{f_\nu j}$. Напомним, что x_{ei}, y_{fj} трансцендентны над полем определения тензора t, включающем в себя все α_ν.

Зафиксируем произвольную биекцию $S'_g \longmapsto [d_P]\times[d_Q]$; $<e_\nu, f_\nu> \longmapsto <i(\nu), j(\nu)>$. Сделав подстановку $x_{e_\nu i} \longmapsto \delta_{i,i(\nu)}$; $y_{f_\nu j} \longmapsto \delta_{j,j(\nu)}$ в матрицу H, мы получим невырожденную матрицу, а именно, диагональную матрицу с ненулевыми вхождениями α_ν по диагонали. В силу трансцендентности x_{ei}, y_{fj}, матрица H также невырождена.

Итак, мы установили, что отображение (15) биективно. Тем самым, (14) доказано. Из (13) и (14) мы теперь получаем $\dim\mu(t\otimes U)=d_P d_Q \dim R\ge\Omega(|S|)$, так как $d_P, d_Q\ge 1$. Отсюда и из (12) вытекает искомая оценка $r_P r_Q\ge\Omega(\dim R\cdot|S|)$.

Осталось разобрать случай, когда одно из натуральных чисел d_P, d_Q (скажем, d_P) равно 0. Докажем, что на самом деле этот случай невозможен. Для этого положим $r'_P=\frac{r_Q \dim R}{|S|}$. Тогда $r_P\le r'_P\le r_Q$ и к паре (r'_P, r_Q) применимы описанные выше рассуждения. Мы докажем таким образом, что $r'_P r_Q\ge\Omega(\dim R\cdot|S|)$, откуда $r_Q\ge\Omega(|S|)$, что противоречит условию, если ε в формулировке теоремы выбрано достаточно малым. ■

СЛЕДСТВИЕ 1. Пусть t_{kG} - тензор умножения в групповой алгебре kG произвольной группы G порядка N (записанный в групповом базисе). Тогда из $\mathrm{rk}_S(t_{kG}) \le (r_P, r_Q)$ при условии $r_P, r_Q \le \varepsilon N^2$ (ε - достаточно малая положительная константа) вытекает оценка $r_P r_Q \ge \Omega(N^3)$. В частности, $\mathrm{rk}_S(t_{kG}) \ge \Omega(N^{3/2})$. ■

СЛЕДСТВИЕ 2. Пусть $a_1, \ldots a_N \in k$; t- тензор умножения в алгебре $k[x]/(x-a_1)\ldots(x-a_N)$, записанный в базисе $(1, x, \ldots, x^{N-1})$. Тогда из $\mathrm{rk}_S(t) \le (r_P, r_Q)$ при условии $r_P, r_Q \le \varepsilon N^2$ (ε - достаточно малая положительная константа) вытекает оценка $r_P r_Q \ge \Omega(N^3)$. В частности, $\mathrm{rk}_S(t) \ge \Omega(N^{3/2})$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим вспомогательный тензор $t' = \sum \left\{ x^i \otimes x^j \otimes x^{i+j} \mid 0 \le i < N/2; \quad 0 \le j < N/2; \quad N/4 \le i+j < N/2 \right\}$ и однородное множество $S = \left\{ \langle x^i, x^j, x^{i+j} \rangle \mid 0 \le i \le N/4; \quad 0 \le j \le N/2; \quad N/4 \le i+j < N/2 \right\}$. Применив теорему 1, докажем следствие 2 для тензора t'. Так как t' получен из t проекцией, сохраняющие выделенные в P и Q базисы, то $\mathrm{rk}_S(t') \le \mathrm{rk}_S(t)$. ■

СЛЕДСТВИЕ 3. $\mathrm{rk}_S(\langle N, N, N \rangle) \ge \Omega(N^{5/2})$ ($\langle N, N, N \rangle$ - тензор матричного умножения). ■

СЛЕДСТВИЕ 4. $\mathrm{rk}_S(t_{pol}) \ge \Omega(N^{3/2})$ (t_{pol}- тензор умножения полиномов N-й степени).

ДОКАЗАТЕЛЬСТВО аналогично доказательству следствия 2. ■

Вернемся теперь к доказательству теорем 2 и 3. Обе они легко

вытекают из уже доказанных результатов и следующей редукции.

ЛЕММА 1. Пусть $\dim P=\dim Q=\dim R=N$ и $t\in P\otimes Q\otimes R$ - тензор, имеющий ранг $\le N$, т.е. допускающий представление $\sum_{i=1}^{N} p_i\otimes q_i\otimes r_i$. Пусть A - матрица, выражающая элементы p_i через базис E, т.е. $p_i=\sum_e a_{ie}e$. Тогда для любого $1\le r\le N$ имеет место $\mathrm{rk}_s(t)\le\left(R_A(r),Nr\right)$.

ДОКАЗАТЕЛЬСТВО. Поставим в соответствие каждой $N\times N$-матрице B тензор $t(B)\in P\otimes Q\otimes R$ по правилу $t(B)=\sum_{i,e} b_{ie}e\otimes q_i\otimes r_i$. Это отображение линейно и $t(A)=t$.

Пусть теперь $A=A_0+A_1$, где $\mathrm{rank}A_0\le r$; $|A_1|\le R_A(r)$. Имеем

$$\mathrm{rk}_s(t)=\mathrm{rk}_s(t(A))\le\mathrm{rk}_s(t(A_0))+\mathrm{rk}_s(t(A_1)). \qquad (16)$$

Если B - матрица ранга 1, т.е. $b_{ie}=x_iy_e$, то $t(B)=\sum_{i,e} x_iy_e e\otimes q_i\otimes r_i=p\otimes a$, где $p\in P$, $a\in Q\otimes R$. Раскладывая a по базису F пространства Q, мы убеждаемся, что $\mathrm{rk}_s(t(B))\le(0,N)$. Следовательно,

$$\mathrm{rk}_s(t(A_0))\le(0,Nr). \qquad (17)$$

Если $|B|=1$, то $t(B)$ имеет вид (10). Значит, $\mathrm{rk}_s(t(B))\le(1,0)$ и

$$\mathrm{rk}_s(t(A_1))\le\left(R_A(r),0\right). \qquad (18)$$

Утверждение леммы вытекает из (16), (17), (18).■

Если G - абелева группа и k - поле разложения для G, то тензор t_{kG} диагонализируем и из следствия 1 к теореме 5 и леммы 1 вытекает теорема 2 (матрица A является матрицей ОПФ, соответствующего G, деленной на $|G|$).

Теорема 3 вытекает из следствия 2 к теореме 5 и леммы 1, т.к. $k[x]/(x-a_1)\dots(x-a_N)\approx k^N$ и матрица перехода от базиса $(1,x,\dots,x^{N-1})$ к базису в k^N обратна соответствующей матрице Вандермонда.

По-видимому, оценки, предоставляемые теоремой 5, не являются

оптимальными. В частности, выглядит правдоподобной следующая гипотеза:

ГИПОТЕЗА. Пусть $\dim P=\dim Q=\dim R=N$; $t\in P\otimes Q\otimes R$ - диагональный тензор; $S\subseteq \mathrm{supp}(t)$ однородно и $|S|=\Omega(N^2)$. Тогда $\mathrm{rk}_s(t)\geq\Omega(N^2)$ (оценка теоремы 5 дает $\Omega(N^{3/2})$).

Как отмечалось во введении, следствиями этой гипотезы являются $PH^{CC}\neq PSPACE^{CC}$ и невычислимость дискретного преобразования Фурье арифметическими схемами размера $O(n)$ и глубины $O(\log n)$.

Автор благодарен Д. Ю. Григорьеву, Н. Нисану, П. Пудлаку, П. Савицкому и М.А.Фрумкину за полезные обсуждения, касающиеся устойчивых матриц.

Литература

1. Гантмахер Ф.Р. "Теория матриц", М.:Наука, 1988.

2. Григорьев Д.Ю. "Использование понятий отделенности и независимости для получения нижних оценок сложности схем". - Зап. научн. семинаров ЛОМИ АН, 1976, т. 60, с. 38-48.

3. Григорьев Д.Ю. "Нижние оценки в алгебраической сложности вычислений". - Зап. научн. семинаров ЛОМИ АН, 1982, т. 118, с. 25-82.

4. А. А. Разборов, "Нижние оценки размера схем ограниченной глубины в полном базисе, содержащем функцию логического сложения" - Матем. зам., 1987, т. 41, вып. 4, с. 598-607. (Engl. transl. in: *Mathem. Notes of the Academy of Sci. of the USSR* 41:4, 333-338).

5. А. А. Разборов, "Формулы ограниченной глубины в базисе $\{\&,\oplus\}$ и некоторые комбинаторные задачи" - в сб. "Вопросы кибернетики.

Сложность вычислений и прикладная математическая логика", М.:1988, с. 149-166.

6. Ajtai M. "Σ^1_1-formulae on finite structures" - *Annals of Pure and Applied Logic*, 1983, 24, p.1-48.

7. L.Babai, P.Frankl, J.Simon, "Complexity classes in communication complexity theory" - *Proc. 27th IEEE FOCS*, 1986, pp. 337-347.

8. Furst M., Saxe J., Sipser M. "Parity, circuits and the polynomial time hierarchy", *Mathematical Systems Theory*, 1984, 17, p.13-27.

9. Hastad J. "Computational Limitations for Small-Depth Circuits" - ACM Doctoral Dissertation Award 1986, the MIT Press.

10. M.Paterson, "Bounded Depth Circuits over {&,⊕}", Warwick, Britain, preprint, 1986.

11. Pudlak P. "The hierarchy of Boolean circuits" - *Computers and Artificial Intelligence*, 1987, v.6, No 5, p.449-468

12. P.Pudlak, V. Rödl, P.Savicky, "Graph Complexity" - *Acta Informatica*, 25, 1988, p. 515-535.

13. Pudlak P., Savicky P., частное сообщение

14. R.Smolensky, "Algebraic methods in the theory of lower bounds for boolean circuit complexity",*Proc. 19th ACM STOC*, 1987, pp. 77-82.

15. Valiant L.G. "Some conjectures relating to super-linear complexity bounds" - Techn. Rep. No 85, Univ. Leeds, 1976.

16. Valiant L.G. "Graph-theoretic arguments in low-level complexity" - Comput. Sci. Rep. 13-77, Univ. Edinburgh, 1977.

17. Yao A.C. "Separating the polynomial-time hierarchy by oracles", *Proc.* 26th *IEEE FOCS*, 1985, p.1-10.

Математический институт

им. В. А. Стеклова АН СССР

ON RIGID MATRICES

The rigidity of an $N\times N$ matrix A over a field k, as defined by L.Valiant [15,16] (see also [2]), is the function from $\{1,\ldots,N\}$ to $\{1,\ldots,N^2\}$ given by $R_A(r)=\min\{|B| \mid \mathrm{rank}(A-B)\le r\}$ where $|B|$ denotes the number of non-zero entries in B. We prove that sufficiently rigid over $\mathbf{F}_p$ matrices are hard to compute by constant-depth circuits over the basis $\{\&,\vee,\neg,MOD\text{-}q\}$, q is a degree of p. As a corollary, we get that the class PH^{CC} contains no sufficiently rigid matrices.

We present also some lower bounds for the rigidity of:

a) matrices of the generalized Fourier transformations corresponding to abelian groups,

b) matrices which are inverse to the van der Monde matrices. Proofs rely upon a reduction to so-called arithmetic switching circuits.