Математические заметки

том 63 выпуск 4 апрель 1998

УДК 519.142+517.984.4

# НОВЫЕ НИЖНИЕ ОЦЕНКИ УСТОЙЧИВОСТИ МАТРИЦ АДАМАРА

**Б. С. Кашин, А. А. Разборов**

Пусть соотношение $f = \Omega(g)$ означает, что $f(x) \geqslant cg(x)$ для некоторой положительной постоянной $c$ и для всех $x$ из области определения функций $f$ и $g$. Показано, что в произвольной (обобщенной) матрице Адамара необходимо изменить по крайней мере $\Omega(n^2/r)$ элементов, для того чтобы ранг полученной матрицы не превосходил $r$. Это улучшает ранее известную оценку $\Omega(n^2/r^2)$. Если дополнительно потребовать, чтобы измененные элементы были ограничены сверху по модулю некоторой величиной $\theta \geqslant n/r$, то справедлива другая оценка $\Omega(n^3/(r\theta^2))$, усиливающая предыдущую $\Omega(n^2/\theta^2)$.

Библиография: 20 названий.

**1. Введение.** Для поля $k$ и матрицы $A$ над этим полем обозначим через $R_A^k(r)$ *функцию устойчивости* $A$, определяемую как минимальное количество изменений элементов матрицы $A$, необходимое для получения матрицы ранга не выше $r$. Это обозначение ввел Валиант [1], [2]; независимо сходное понятие предложил Григорьев [3].

Основной причиной введения и изучения этого понятия является то обстоятельство, что устойчивые матрицы являются сложными по отношению ко многим важным вычислительным моделям, получение нижних оценок для которых – трудная задача, не решенная к настоящему времени. Такими моделями являются, например, линейные алгебраические схемы [2], [4] и вычисления с ограниченным числом альтераций в теории коммуникационной сложности [5], [6]. В частности, для любого *конкретного* примера достаточно устойчивой матрицы можно было бы построить соответствующую вычислительную задачу, являющуюся сложной с точки зрения указанных моделей. Это привело бы к существенным продвижениям в этой области.

Несмотря на многочисленные усилия, наилучшие известные нижние оценки величины $R_A^k(r)$, полученные для конкретных $(n \times n)$-матриц $A$, имеют вид $\Omega\big(n^2/(r\log(n/r))\big)$ для конечного поля $k$ (Фридман [7]) и $\Omega(n^2/r)$ в случае бесконечного поля $k$ ($r \leqslant n/2$). Простейший способ получения последней оценки был независимо отмечен Григорьевым и Нисаном: для любой *вполне невырожденной* матрицы $A$ (т.е. матрицы, все миноры которой невырожденны) выполнено $R_A^k(r) \geqslant \Omega(n^2/r)$ ($k$ произвольно, $r \leqslant n/2$). Другие методы получения подобной оценки $\Omega(n^2/r)$, реализованные для различных классов

Работа первого автора выполнена при поддержке Российского фонда фундаментальных исследований, гранты № 96-01-00094 и № 96-15-96102, и фонда INTAS, грант № 93-1376. Работа второго автора выполнена при поддержке Российского фонда фундаментальных исследований, гранты № 96-01-01222 и № 96-15-96090.

матриц, были предложены в [5], [8]–[10]; некоторые из них применимы и в случае конечных полей.

Одним из классов матриц, для которого естественно предполагать наличие свойства сильной устойчивости над $\mathbb{R}$, является класс матриц Адамара. Эти матрицы хорошо приспособлены к изучению спектральными методами, которые зарекомендовали себя как чрезвычайно удобный инструмент для решения подобных задач в теории сложности (см. типичные примеры в [11], [12]). Пудлак, Разборов и Савицкий [13] доказали оценку $R_H^{\mathbb{R}}(r) \geqslant \Omega(n^2/(r^3 \log r))$, где $H$ – произвольная матрица Адамара, а Алон [14] усилил ее: $R_H^{\mathbb{R}}(r) \geqslant \Omega(n^2/r^2)$. Исследуя упомянутые выше приложения в теории коммуникационной сложности, Локам [6] рассмотрел модифицированную функцию устойчивости $R_A^{\mathbb{C}}(r,\theta)$, соответствующую случаю, когда дополнительно требуется, чтобы *абсолютные величины* измененных элементов не превосходили $\theta$, и показал, в частности, что $R_H^{\mathbb{C}}(n/2,\theta) \geqslant \Omega(n^2/\theta^2)$ для любой (обобщенной) матрицы Адамара $H$.

В настоящей заметке мы устанавливаем усиления некоторых из упомянутых выше результатов. Именно, для произвольной (обобщенной) матрицы Адамара $H$ и любого $r \leqslant n/2$ мы доказываем, что $R_H^{\mathbb{C}}(r) \geqslant \Omega(n^2/r)$, и если $\theta$ является параметром, удовлетворяющим неравенству $\theta \geqslant n/r$, то $R_H^{\mathbb{C}}(r,\theta) \geqslant \Omega(n^3/(r\theta^2))$. Наши оценки слишком слабые, чтобы получить какие-либо содержательные следствия в теории сложности, и они не переходят границу $n^2/r$. Основная цель настоящей заметки – привлечь внимание специалистов к использованию в этой области спектральных методов, так как мы уверены, что их потенциал далеко не исчерпан.

## 2. Устойчивость матриц.

Определение 1 [1], [2], [6]. Для матрицы $A$ над некоторым полем $k$ обозначим через $\mathrm{wt}(A)$ количество ее ненулевых элементов. *Функция устойчивости* $R_A^k(r)$ матрицы $A$ определяется по формуле

$$R_A^k(r) \rightleftharpoons \min_B\{\mathrm{wt}(A-B) \mid \mathrm{rk}(B) \leqslant r\},$$

где $B$ пробегает все матрицы того же размера, что и $A$. Если $k=\mathbb{C}$ и $\theta$ – положительное вещественное число, то положим

$$R_A^{\mathbb{C}}(r,\theta) \rightleftharpoons \min_B\{\mathrm{wt}(A-B) \mid \mathrm{rk}(B) \leqslant r,\ |a_{ij}-b_{ij}| \leqslant \theta\ \ \forall\, i,j\}. \tag{1}$$

Замечание 1. Локам [6] использовал в своем определении функции $R_A^{\mathbb{C}}(r,\theta)$ ограничение $|b_{ij}| \leqslant \theta$ вместо нашего $|a_{ij}-b_{ij}| \leqslant \theta$. Однако, более естественным представляется оценивать $\ell_\infty$-норму *изменений*, а не конечного результата. В частности, наше определение позволяет также устанавливать нетривиальные оценки в случае $\theta \ll 1$ (см. ниже предложение 1).

Определение 2. Комплексная $(n \times n)$-матрица $H$ называется *обобщенной матрицей Адамара*, если $|h_{ij}|=1$ для всех $i,j$ и $HH^*=nI_n$, где $H^*$ – сопряженная к $H$ матрица и $I_n$ – единичная $(n\times n)$-матрица.

Предложение 1. *Для любой обобщенной $(n\times n)$-матрицы Адамара $H$ и любого $r \leqslant n/2$*

а) $R_H^{\mathbb{C}}(r) \geqslant \Omega(n^2/r^2)$ *(см.* [14], [6]);

б) $R_H^{\mathbb{C}}(n/2,\theta) \geqslant \Omega(n^2/\theta^2)$, *где $\theta>0$ – произвольный параметр (см.* [6]);

в) $R_H^{\mathbb{C}}(n/2,\theta) \geqslant \Omega(n^2/\theta)$ *в случае $\theta \leqslant n/r$ (см.* [6]).

Замечание 2. Строго говоря, предложение 1в) сформулировано в [6] для более ограничительного варианта устойчивости $R^+_A(r,\theta)$, получающегося добавлением в (1) дополнительного условия $|b_{ij}| \geqslant 1$. Однако, несложный анализ доказательства Локама показывает, что это условие может быть опущено.

В настоящей заметке мы докажем, что справедлива

Теорема. *Для любой обобщенной $(n\times n)$-матрицы Адамара $H$ и любого $r \leqslant n/2$*

а) $R^{\mathbb{C}}_H(r) \geqslant \Omega(n^2/r)$;

б) $R^{\mathbb{C}}_H(r,\theta) \geqslant \Omega(n^3/(r\theta^2))$ *в случае* $\theta \geqslant n/r$.

Утверждение б) теоремы усиливает предложение 1в) на множитель $n/r$ при $\theta \geqslant n/r$ (его формулировка и доказательство могут быть рассмотрены как "непрерывное продолжение" предложения 1в) при $\theta \geqslant n/r$).

**3. Спектры матриц.** *Норма Фробениуса* $\|A\|_F$ комплексной матрицы $A$ определяется, как

$$\|A\|_F \rightleftharpoons \Bigl(\sum_{i,j} |a_{ij}|^2\Bigr)^{1/2}.$$

Пусть $(n \times n)$-матрица $A$ является вещественной симметричной и $\lambda_1(A) \geqslant \dots \geqslant \lambda_n(A)$ – ее (вещественные) собственные значения. Тогда

$$\operatorname{Tr}(A) = \lambda_1(A) + \dots + \lambda_n(A), \tag{2}$$

$$\|A\|_F^2 = \lambda_1^2(A) + \dots + \lambda_n^2(A). \tag{3}$$

Пусть $r = \operatorname{rk}(A)$, т.е. в точности $r$ собственных значений среди $\lambda_1(A), \dots, \lambda_n(A)$ ненулевые. Из неравенства Коши согласно (2) и (3) следует, что $\operatorname{Tr}(A)^2 \leqslant \|A\|_F^2 \cdot r$. Таким образом, справедлива

Лемма 1. *Для любой вещественной симметричной матрицы $A$*

$$\operatorname{rk}(A) \geqslant \frac{\operatorname{Tr}(A)^2}{\|A\|_F^2}.$$

Напомним также неравенство Хоффмана–Виландта. Через $\sigma_i(A) \rightleftharpoons \sqrt{\lambda_i(AA^*)}$ обозначим *$i$-е сингулярное число* комплексной матрицы $A$.

Предложение 2 [15], [16, разд. 8.3]. *Для любой пары комплексных $(n \times n)$-матриц $A, B$*

$$\sum_{i=1}^n \bigl(\sigma_i(A) - \sigma_i(B)\bigr)^2 \leqslant \|A - B\|_F^2.$$

**4. Доказательство теоремы.** а) Как и в предшествующих доказательствах (см. [13], [14], [6]), мы будем стремиться показать, что подматрицы произвольной обобщенной матрицы Адамара имеют достаточно высокий ранг (по крайней мере в среднем), а это позволит перенести на наш случай наблюдение Григорьева и Нисана для вполне невырожденных матриц.

В действительности, в функциональном анализе имеется ряд результатов (см., в частности, [17]–[20]), посвященных общему вопросу: "насколько ортогональны выбранные случайным образом подматрицы ортогональной матрицы”? Поскольку “достаточно ортогональные” матрицы являются невырожденными, мы можем непосредственно применить результаты из [19], [20] для получения оценки $R_H^{\mathbb{C}}(r) \geqslant \Omega(n^2/(r\log n))$. С другой стороны, для наших целей необходимо нечто более слабое, чем то, что дает достаточно тонкая техника из [19], [20]. Поэтому мы заменим ее следующей простой леммой, позволяющей избавиться от логарифмического множителя в знаменателе.

Всюду далее мы используем жирный шрифт для обозначения случайных объектов.

Лемма 2. *Пусть $q \leqslant n$ и $A$ – комплексная $(q\times n)$-матрица такая, что $|a_{ij}|=1$ для всех $i,j$ и $AA^* = nI_q$. Пусть $\boldsymbol{B}$ – случайно выбранная $(q\times q)$-подматрица матрицы $A$. Тогда для любого $r$*

$$\mathsf{P}[\mathrm{rk}(\boldsymbol{B}) \leqslant r] \leqslant \frac{2r}{q}.$$

Доказательство. Обозначим $\boldsymbol{B}\boldsymbol{B}^*$ через $\boldsymbol{C}$. Тогда $\boldsymbol{C}$ является (положительно определенной) вещественной симметричной $(q\times q)$-матрицей, все элементы которой на главной диагонали равны $q$ (поэтому $\mathrm{Tr}(\boldsymbol{C}) = q^2$) и которая удовлетворяет неравенству $\mathrm{rk}(\boldsymbol{C}) \leqslant \mathrm{rk}(\boldsymbol{B})$. Таким образом, в силу леммы 1 (примененной к $A := \boldsymbol{C}$)

$$\mathrm{rk}(\boldsymbol{B}) \leqslant r \implies \|\boldsymbol{C}\|_F^2 \geqslant \frac{q^4}{r}. \tag{4}$$

С другой стороны, согласно неравенству Чебышева

$$\mathsf{P}\biggl[\|\boldsymbol{C}\|_F^2 \geqslant \frac{q^4}{r}\biggr] \leqslant \frac{r}{q^4}\cdot \mathsf{E}\bigl[\|\boldsymbol{C}\|_F^2\bigr]. \tag{5}$$

Положим

$$\boldsymbol{\xi}_j \rightleftharpoons \begin{cases} 1, & \text{если } j\text{-й столбец представлен в матрице } \boldsymbol{B}, \\ 0 & \text{в противном случае;} \end{cases}$$

случайный вектор $(\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_n)$ равномерно распределен на множестве всех векторов в $\{0,1\}^n$, имеющих ровно $q$ единиц. Пусть $\boldsymbol{c}_{i_1i_2} = \sum_{j=1}^n a_{i_1j}a_{i_2j}\boldsymbol{\xi}_j$. Тогда

$$\begin{aligned}\|\boldsymbol{C}\|_F^2 &= \sum_{i_1,i_2}(\boldsymbol{c}_{i_1i_2}\boldsymbol{c}^*_{i_1i_2}) = \sum_{i_1,i_2}\sum_{j_1,j_2}(a_{i_1j_1}a_{i_2j_1}a^*_{i_1j_2}a^*_{i_2j_2}\boldsymbol{\xi}_{j_1}\boldsymbol{\xi}_{j_2}) \\ &= \sum_{j_1,j_2}\biggl(\sum_{i_1,i_2}a_{i_1j_1}a_{i_2j_1}a^*_{i_1j_2}a^*_{i_2j_2}\biggr)\boldsymbol{\xi}_{j_1}\boldsymbol{\xi}_{j_2}.\end{aligned}$$

Далее

$$\mathsf{E}[\boldsymbol{\xi}_{j_1}\boldsymbol{\xi}_{j_2}] = \begin{cases} \dfrac{q}{n}, & \text{если } j_1 = j_2, \\ \dfrac{q(q-1)}{n(n-1)}, & \text{если } j_1 \neq j_2. \end{cases}$$

Следовательно,

$$\begin{aligned}\mathsf{E}\bigl[\|\boldsymbol{C}\|_F^2\bigr] &= \frac{q(q-1)}{n(n-1)}\sum_{j_1,j_2}\sum_{i_1,i_2}(a_{i_1j_1}a_{i_2j_1}a^*_{i_1j_2}a^*_{i_2j_2}) \\ &\quad + \biggl(\frac{q}{n} - \frac{q(q-1)}{n(n-1)}\biggr)\sum_{j}\sum_{i_1,i_2}(a_{i_1j}a_{i_2j}a^*_{i_1j}a^*_{i_2j}) \\ &= \frac{q(q-1)}{n(n-1)}\|AA^*\|_F^2 + \biggl(\frac{q}{n} - \frac{q(q-1)}{n(n-1)}\biggr)nq^2 = q^2\biggl(q + \frac{q-1}{n-1}(n-q)\biggr) \leqslant 2q^3.\end{aligned} \tag{6}$$

Теперь лемма 2 непосредственно вытекает из (4)–(6).

Следствие. *Пусть* $H$ – *обобщенная* $(n \times n)$-*матрица Адамара и* $\boldsymbol{H}_0$ – *ее случайная* $(q \times q)$-*подматрица. Тогда* $\mathsf{E}[\mathrm{rk}(\boldsymbol{H}_0)] \geqslant q/8$.

Доказательство. Подматрица $\boldsymbol{H}_0$ строится в два этапа: сначала мы выбираем случайным образом $q$ строк, затем $q$ столбцов. Поскольку $(q \times n)$-матрица $\boldsymbol{A}$, получающаяся на первом этапе, удовлетворяет требованиям леммы 2, для любого фиксированного выбора строк $1 \leqslant i_1 < \cdots < i_q \leqslant n$ имеет место оценка условной вероятности $\mathsf{P}[\mathrm{rk}(\boldsymbol{H}_0) \leqslant q/4 \mid \boldsymbol{H}_0$ содержит строки $i_1, \ldots, i_q] \leqslant 1/2$. Следовательно, $\mathsf{P}[\mathrm{rk}(\boldsymbol{H}_0) \leqslant q/4] \leqslant 1/2$ и $\mathsf{E}[\mathrm{rk}(\boldsymbol{H}_0)] \geqslant q/4 \cdot \mathsf{P}[\mathrm{rk}(\boldsymbol{H}_0) \geqslant q/4] \geqslant q/8$.

Теперь мы готовы завершить доказательство теоремы. Пусть $H$ – обобщенная $(n \times n)$-матрица Адамара и $r \leqslant n/2$. Поскольку оценка $R_H^{\mathbb{C}}(n/2) \geqslant n/2$ очевидна, не ограничивая общности можем считать, что $r \leqslant n/16$. Пусть матрица $A$ такова, что $\mathrm{rk}(H - A) \leqslant r$. Положим $q = 16r$ и в матрице $H$ случайным образом выберем $(q \times q)$-подматрицу $\boldsymbol{H}_0$. Если $\boldsymbol{A}_0$ – соответствующая подматрица матрицы $A$, то

$$\mathsf{E}[\mathrm{rk}(\boldsymbol{H}_0)] \leqslant \mathsf{E}[\mathrm{rk}(\boldsymbol{H}_0 - \boldsymbol{A}_0)] + \mathsf{E}[\mathrm{rk}(\boldsymbol{A}_0)] \leqslant r + \mathsf{E}[\mathrm{wt}(\boldsymbol{A}_0)] = r + \frac{q^2}{n^2}\,\mathrm{wt}(A).$$

С другой стороны, $\mathsf{E}[\mathrm{rk}(\boldsymbol{H}_0)] \geqslant 2r$ согласно следствию из леммы 2. Объединяя два последних неравенства, заключаем, что $\mathrm{wt}(A) \geqslant rn^2/q^2 = \Omega(n^2/r)$. Это завершает доказательство утверждения а) теоремы.

б) Предложение 1в) было доказано в [6] путем применения неравенства Хоффмана–Виландта к паре матриц $H, \frac{1}{\theta}B$. Мы обобщим этот подход на случай $\theta \geqslant n/r$ варьированием коэффициента при матрице $B$.

Пусть, как и ранее, $H$ – обобщенная $(n \times n)$-матрица Адамара и $r \leqslant n/2$. Пусть матрица $A$ такова, что $\mathrm{rk}(H - A) \leqslant r$ и $|a_{ij}| \leqslant \theta$ для всех $i, j$, где $\theta \geqslant n/r$ – еще один параметр. Применяя предложение 2 к паре матриц $H, \frac{r}{n}(H - A)$, мы, как и в [6], имеем

$$\left\| H - \frac{r}{n}(H - A) \right\|_F^2 \geqslant n(n - r). \tag{7}$$

С другой стороны,

$$H - \frac{r}{n}(H - A) = \left(1 - \frac{r}{n}\right)H + \frac{r}{n}A.$$

Следовательно, эта матрица имеет не более $\mathrm{wt}(A)$ элементов, ограниченных по модулю величиной $1 + r\theta/n$, в то время как остальные элементы по модулю не превосходят $1 - r/n$. Отсюда следует оценка

$$\left\| H - \frac{r}{n}(H - A) \right\|_F^2 \leqslant \mathrm{wt}(A)\left(1 + \frac{r\theta}{n}\right)^2 + \left(1 - \frac{r}{n}\right)^2 n^2. \tag{8}$$

Из неравенств (7) и (8) получаем

$$\mathrm{wt}(A)\left(1 + \frac{r\theta}{n}\right)^2 \geqslant r(n - r).$$

Поскольку $r \leqslant n/2$ и $\theta \geqslant n/r$, отсюда следует требуемая оценка $\mathrm{wt}(A) \geqslant \Omega(n^3/(r\theta^2))$. Это завершает доказательство утверждения б) теоремы.

## СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ


[1] Valiant L. G. Some Conjectures Relating to Superlinear Complexity Bounds. Tech. Report no. 85: Univ. of Leeds, 1976.

[2] Valiant L. G. Graph-Theoretic Arguments in Low-Level Complexity. Tech. Report no. 13-77: Univ. of Edinburgh, Dept. of Comp. Sci., 1977.

[3] Григорьев Д. Ю. Использование понятий отделенности и независимости для получения нижних оценок сложности схем // Записки научн. семин. ЛОМИ. 1976. Т. 60. С. 38–48.

[4] Григорьев Д. Ю. Нижние оценки в алгебраической сложности вычислений // Записки научн. семин. ЛОМИ. 1982. Т. 118. С. 25–82.

[5] Разборов А. А. Об устойчивых матрицах. Препринт. М.: МИАН, 1989.

[6] Lokam S. V. Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity // Proc. of the 36th IEEE Symposium on Foundations of Computer Science. Los Alamitos (C.A.): IEEE, 1995. P. 6–15.

[7] Friedman J. A note on matrix rigidity // Combinatorica. 1993. V. 13. №2. P. 235–239.

[8] Pudlák P., Vavřín Z. Computation of rigidity of order $n^2/r$ for one simple matrix // Comment. Math. Univ. Carolin. 1991. V. 32. №2. P. 213–218.

[9] Kimmel P., Settle A. Reducing the Rank of Lower Triangular All-Ones Matrix. Tech. Report CS 92-21: Univ. of Chicago, 1992.

[10] Pudlák P. Large communication in constant depth circuits // Combinatorica. 1994. V. 14. №2. P. 203–216.

[11] Krause M., Waack S. Variation ranks of communication matrices and lower bounds for depth two circuits having symmetric gates with unbounded fan-in // Proc. of the 32nd IEEE Symposium on Foundations of Computer Science. Los Alamitos (C.A.): IEEE, 1991. P. 777–782.

[12] Nisan N., Wigderson A. On the complexity of bilinear forms // Proceedings of the 27th ACM Symposium on the Theory of Computing. New York: ACM, 1995. P. 723–732.

[13] Pudlák P., Razborov A., Savický P. Observations on Rigidity of Hadamard Matrices. Manuscript, 1988.

[14] Alon N. On the Rigidity of Hadamard Matrices. Manuscript: Tel Aviv Univ., 1990.

[15] Hoffman A. J., Wielandt H. W. The variation of the spectrum of a normal matrix // Duke Math. J. 1953. V. 20. P. 37–39.

[16] Golub G. H., van Loan C. F. Matrix Computations. Baltimore, Maryland: John Hopkins Univ. Press, 1983.

[17] Кашин Б. С. О некоторых свойствах матриц ограниченных операторов из пространства $\ell_2^n$ в $\ell_2^m$ // Изв. АН АрмССР. Матем. 1980. Т. 15. №5. С. 379–394.

[18] Лунин А. А. Об операторных нормах подматриц // Матем. заметки. 1989. Т. 45. №3. С. 94–100.

[19] Kashin B., Tzaffiri L. Some Remarks on the Restriction of Operators to Coordinate Subspaces. Tech. Report no. 12. Jerusalem: The Edmund Landau Center for Research in Math. Anal., Hebrew Univ., 1993/94.

[20] Rudelson M. Almost orthogonal submatrices of an orthogonal matrix // Israel J. Math. (to appear).


Математический институт им. В. А. Стеклова РАН
*E-mail*: `kashin@mi.ras.ru, razborov@mi.ras.ru`

Поступило
01.12.97