

УДК 512.64

О НОРМЕ ШАТТЕНА ЛИНЕЙНОГО ОТОБРАЖЕНИЯ ЕВКЛИДОВЫХ ПРОСТРАНСТВ

Д. Н. Курочкина

Белорусский государственный университет, Минск, Беларусь
e-mail: fpm.kurochkiDN@bsu.by

Поступила: 22.01.2025

Исправлена: 23.03.2025

Принята: 23.05.2025

Ключевые слова: линейное отображение, евклидово пространство, сингулярные числа, ортогональная матрица, неравенство Йенсена, средние степени, неравенство Адамара.

Аннотация. Установлена связь нормы Шаттена линейного отображения $\mathcal{A} : V \rightarrow W$ евклидовых пространств размерностей n и m соответственно со средними степенными, порожденными длинами образов векторов ортонормированного базиса V .

ON THE SCHATTEN NORM OF A LINEAR MAP ON EUCLIDEAN SPACES

D. N. Kurochkina

Belarusian State University, Minsk, Belarus
e-mail: fpm.kurochkiDN@bsu.by

Received: 22.01.2025

Revised: 23.03.2025

Accepted: 23.05.2025

Keywords: linear map, Euclidean space, singular values, orthogonal matrix, Jensen's inequality, power means, Hadamard's inequality.

Abstract. We find a connection between the Schatten norm of the linear mapping $\mathcal{A} : V \rightarrow W$ of Euclidean spaces of dimensions n and m , respectively, and the power means generated by the lengths of the images of vectors of the orthonormal basis V .

1. Введение

Пусть V, W – евклидовы пространства размерностей n и m соответственно, тогда $\mathcal{A} : V \rightarrow W$ есть линейное отображение ранга $k \leq \min\{n, m\}$, $\mathcal{A}^* : W \rightarrow V$ – сопряженное к $\mathcal{A}$ отображение. Известно, что $\mathcal{A}^* \circ \mathcal{A} : V \rightarrow V$ есть самосопряженный оператор, матрицей которого в некотором ортонормированном базисе $\mathcal{E}_V = \langle \varepsilon_1, \dots, \varepsilon_n \rangle$ пространства V является диагональная матрица $\text{diag}[\rho_1^2, \dots, \rho_n^2]$, где $\rho_1 \geq \dots \geq \rho_k > 0 = \rho_{k+1} = \dots = \rho_n$ есть так называемые сингулярные числа оператора $\mathcal{A}$ (все используемые в статье сведения о матрицах и линейных отображениях можно найти в [1] и [2, Гл. 4, § 16]). Теорема о сингулярном разложении утверждает, что существует такой ортонормированный базис $\mathcal{F} = \langle f_1, \dots, f_m \rangle$ пространства W , что

$$\mathcal{A}(\varepsilon_i) = \rho_i f_i \text{ при } i \leq n \text{ и } \mathcal{A}^*(f_j) = \rho_j \varepsilon_j \text{ при } j \leq m.$$

Обозначим через $\mathcal{B} = (\mathcal{A}^* \circ \mathcal{A})^{1/2} : V \rightarrow V$ корень из самосопряженного оператора $\mathcal{A}^* \circ \mathcal{A}$, а через $\mathcal{B}^p : V \rightarrow V$ – p -ю степень оператора $\mathcal{B}$. В базисе $\mathcal{E}_V = \langle \varepsilon_1, \dots, \varepsilon_n \rangle$ матрицей оператора $\mathcal{B}^p : V \rightarrow V$ является диагональная матрица $\text{diag}[\rho_1^p, \dots, \rho_n^p]$.

Введем в рассмотрение число

$$|\mathcal{A}|_p = \text{tr}(\mathcal{B}^p)^{\frac{1}{p}} = (\rho_1^p + \dots + \rho_n^p)^{\frac{1}{p}}, \text{ где } p \geq 1.$$

Ясно, что $|\mathcal{A}|_2 = \sqrt{\rho_1^2 + \dots + \rho_n^2}$, а $|\mathcal{A}|_1 = \rho_1 + \dots + \rho_n$. Известно [1, р. 464–465], что $|\mathcal{A}|_p$ является нормой в векторном пространстве $\text{Lin}(V, W)$ всех линейных отображений, которая называется p -нормой Шаттена отображения $\mathcal{A}$. Норма Шаттена является матричной, т. е. $|\mathcal{A}_2 \circ \mathcal{A}_1|_p \leq |\mathcal{A}_2|_p \cdot |\mathcal{A}_1|_p$ для композиции $V \xrightarrow{\mathcal{A}_1} W \xrightarrow{\mathcal{A}_2} U$ линейных отображений.

Всюду далее V и W будут евклидовы пространства размерностей n и m соответственно.

Определение 1.1. Обозначим для линейного отображения $\mathcal{A} : V \rightarrow W$ число

$$\|\mathcal{A}\|_p = \sup_{\mathcal{E}} \left(\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p \right)^{1/p},$$

где супремум берется по всем ортонормированным базисам $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V (через $\|\mathcal{A}(e_i)\|$ обозначена длина вектора $\mathcal{A}(e_i)$).

Впервые такого рода характеристика линейного отображения была рассмотрена в работе [3] для ортогональной проекции $\mathcal{A} = pr : \mathbb{R}^n \rightarrow W$ на подпространство $W \subset \mathbb{R}^n$ в связи с уточнением оценки дисперсии случайной ошибки в теории Гаусса метода наименьших квадратов.

Используя стандартные рассуждения и неравенство Минковского, несложно показать, что $\|\mathcal{A}\|_p$ при $p \geq 1$ является нормой линейного отображения $\mathcal{A} : V \rightarrow W$, где V и W – пространства размерностей m, n соответственно.

Целью данной заметки является сравнение нормы $\|\mathcal{A}\|_p$ и p -нормы Шаттена $|\mathcal{A}|_p$. С одной стороны они похожи по форме и кроме того $\|\mathcal{A}\|_2 = |\mathcal{A}|_2$ (что следует из цепочки равенств

$$\text{tr}(\mathcal{A}^* \circ \mathcal{A}) = \sum_{i \leq n} (\mathcal{A}^* \circ \mathcal{A}(e_i), e_i) = \sum_{i \leq n} (\mathcal{A}(e_i), \mathcal{A}(e_i)) = \sum_{i \leq n} \|\mathcal{A}(e_i)\|^2 \quad (1)$$

для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$).

С другой стороны, проверка неравенства треугольника для p -нормы Шаттена $|\mathcal{A}|_p$ вызывает большие сложности (что не так для нормы $\|\mathcal{A}\|_p$). Теорема 1.2 еще сильнее подчеркивает нетривиальность во взаимоотношении между ними.

Теорема 1.2. Для линейного отображения $\mathcal{A} : V \rightarrow W$ справедливо

$$\|\mathcal{A}\|_p = \begin{cases} |\mathcal{A}|_p, & \text{если } p \geq 2 \\ c_p \cdot |\mathcal{A}|_2, & \text{если } 1 \leq p < 2 \end{cases}$$

(здесь константа $c_p = n^{1/p-1/2}$).

Доказательство теоремы 1.2 основывается на точных оценках для средних k -степенных, порождаемых линейным отображением (см. теорему 1.3). Напомним, что средним k -степенным неотрицательных чисел $\beta_1, \dots, \beta_n$ называется число

$$A_k(\beta_1, \dots, \beta_n) = \begin{cases} \left(\frac{\beta_1^k + \dots + \beta_n^k}{n} \right)^{\frac{1}{k}} & \text{при } k \neq 0 \\ (\beta_1 \cdot \dots \cdot \beta_n)^{\frac{1}{k}} & \text{при } k = 0. \end{cases}$$

Известно, что

$$A_k(\beta_1, \dots, \beta_n) \geq A_2(\beta_1, \dots, \beta_n) \geq A_l(\beta_1, \dots, \beta_n) \text{ при } k \geq 2 \geq l. \quad (2)$$

Пусть $\mathcal{A} : V \rightarrow W$ есть линейное отображение и $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ – ортонормированный базис V . Через $A_k(\mathcal{E})$ обозначим $A_k(\beta_1, \dots, \beta_n)$, где $\beta_i = \|\mathcal{A}(e_i)\| \geq 0$ при $i \leq n$.

Теорема 1.3. Пусть $\rho_1, \dots, \rho_n$ есть сингулярные числа линейного отображения $\mathcal{A} : V \rightarrow W$. Тогда для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V среднее степенное $A_k(\mathcal{E})$ удовлетворяет двойному неравенству

1. $A_k(\rho_1, \dots, \rho_n) \geq A_k(\mathcal{E}) \geq A_2(\rho_1, \dots, \rho_n)$ при $k \geq 2$;
2. $A_k(\rho_1, \dots, \rho_n) \leq A_k(\mathcal{E}) \leq A_2(\rho_1, \dots, \rho_n)$ при $k < 2$.

Замечание 1.4. В [3] для ортогональной проекции $\mathcal{A} = pr : \mathbb{R}^n \rightarrow W$ для $k = 4$ было установлено двойное неравенство:

$$(*) \left(\frac{m}{n} \right)^{\frac{1}{4}} \geq A_4(\mathcal{E}) \geq \left(\frac{m}{n} \right)^{\frac{1}{2}}. \quad (3)$$

которое согласуется с теоремой 1.3, так как сингулярные числа ортогональной проекции равны $\rho_1 = \dots = \rho_m = 1$, а $\rho_{m+1} = \dots = \rho_n = 0$. В работе [4] для ортогональной проекции была установлена неулучшаемость оценки (3) снизу, а в [5] были получены необходимые и достаточные условия для существования ортогонального базиса, проектирующегося в векторы равной длины.

2. Редукция теоремы 1.2 к теореме 1.3 и ортогональным матрицам

Ясно, что для сингулярного базиса $\mathcal{E}_V$ $A_p(\mathcal{E}_V)$ совпадает с $A_p(\rho_1, \dots, \rho_n)$. Этот факт и теорема 1.3 совместно влекут:

1. $\max_{\mathcal{E}} A_p(\mathcal{E}) = A_p(\rho_1, \dots, \rho_n)$ при $p \geq 2$;
2. $\min_{\mathcal{E}} A_p(\mathcal{E}) = A_p(\rho_1, \dots, \rho_n)$ при $p < 2$.

Поэтому $\|\mathcal{A}\|_p = \max_{\mathcal{E}} (\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p)^{1/p} = n^{1/p} \max_{\mathcal{E}} A_p(\mathcal{E})$ при $p \geq 2$ совпадает с

$$n^{1/p} A_p(\rho_1, \dots, \rho_n) = \left(\sum_{i \leq n} \rho_i^p \right)^{1/p} = |\mathcal{A}|_p.$$

В учебнике С. М. Агеева «Курс линейной алгебры» содержится, что $A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ для так называемого *равнобедренного базиса* $\mathcal{E} = \mathcal{E}_{eq}$, т. е. такого ортонормированного базиса, у которого длины β_i всех векторов $\mathcal{A}(e_i)$ равны между собой (и, следовательно, $\beta_i = A_2(\beta_1, \dots, \beta_n) = A_2(\rho_1, \dots, \rho_n)$). Следовательно,

3. $\max_{\mathcal{E}} A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ при $p < 2$;
4. $\min_{\mathcal{E}} A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ при $p \geq 2$.

Поэтому $\|\mathcal{A}\|_p = \max_{\mathcal{E}} (\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p)^{1/p} = n^{1/p} \max_{\mathcal{E}} A_p(\mathcal{E})$ при $1 \leq p < 2$ совпадает с

$$n^{1/p} A_2(\rho_1, \dots, \rho_n) = n^{1/p-1/2} \left(\sum_{i \leq n} \rho_i^2 \right)^{1/2} = c_p |\mathcal{A}|_2.$$

Покажем, что доказательство теоремы 1.3 сводится к аналогичному утверждению для ортогональных матриц.

Предложение 2.1. Пусть $\mathcal{E}_V$ есть сингулярный, а $\mathcal{E}$ – ортонормированный базисы пространства V . Пусть также T есть ортогональная матрица $T(\mathcal{E}_V | \mathcal{E})$ перехода от сингулярного базиса $\mathcal{E}_V$ к базису $\mathcal{E}$, а сингулярные числа отображения $\mathcal{A}$ стоят на диагонали матрицы $R = \text{diag}[\rho_1, \rho_2, \dots, \rho_n]$, где $\rho_1 \geq \rho_2 \geq \dots \geq \rho_n$. Тогда $\|\mathcal{A}(e_i)\| = \beta_i$ для любого $i \leq n$ совпадает с длиной s_i i -го столбца матрицы $T_R = RT$.

Доказательство. Достаточно показать, что T_R есть матрица $T(\mathcal{E}_W | \mathcal{A}(\mathcal{E}))$ перехода от сингулярного базиса $\mathcal{E}_W$ к образу $\mathcal{A}(\mathcal{E})$. Действительно, $e_j = \sum_{i=1}^n t_{ij} \varepsilon_i$ и поэтому

$$\mathcal{A}(e_j) = \sum_{i=1}^n t_{ij} \mathcal{A}(\varepsilon_i) = \sum_{i=1}^n (t_{ij} \rho_i) f_i.$$

Следовательно, $s_j = \sum_{i=1}^n (t_{ij} \rho_i)^2 = \|\mathcal{A}(e_j)\|^2 = \beta_j$. □

Предложение 2.2. Если $\mathcal{E}$ – ортонормированный базис пространства V , T – матрица перехода от сингулярного базиса $\mathcal{E}_V$ к $\mathcal{E}$, то $A_k(\mathcal{E}) = \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{1/k}$, где s_i есть длина i -го столбца матрицы T_R .

Отсюда следует, что при доказательстве теоремы 1.3 (1)–(2) длины векторов $\|\mathcal{A}(e_i)\|$ можно заменить на длины s_i столбцов матрицы T_R .

3. Доказательство теоремы 1.3 (1)

Из неравенства (2) и цепочки равенств (1) следует, что $A_2(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V , а также

$$A_k(\mathcal{E}) \geq A_2(\rho_1, \dots, \rho_n) \text{ при } k \geq 2 \text{ и } A_k(\mathcal{E}) \leq A_2(\rho_1, \dots, \rho_n) \text{ при } k \leq 2.$$

Отсюда и из предложения 2.2 следует, что доказательство неравенства (1) из теоремы 1.3 сводится к неравенству для ортогональных матриц при $k \geq 2$:

$$\left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{1/k} \geq \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{1/k}.$$

Для его установления воспользуемся неравенством Иенсена: если гладкая числовая функция $y = f(x)$ является выпуклой, то выполняется неравенство:

$$f(\alpha_1 x_1 + \dots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \dots + \alpha_n f(x_n),$$

где числа $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ таковы, что их сумма равна единице (если же функция $f(x)$ вогнута, то

$$f(\alpha_1 x_1 + \dots + \alpha_n x_n) \geq \alpha_1 f(x_1) + \dots + \alpha_n f(x_n).$$

Ясно, что (*) функция $f(x) = x^{\frac{k}{2}}$, имеющая 2-ю произвольную $f^{(2)}(x) = \frac{k}{2}(\frac{k}{2} - 1)x^{\frac{k}{2}-2}$, является выпуклой при $k \geq 2$ и $k < 0$ (является вогнутой при $0 < k < 2$).

Пусть $k \geq 2$. Так как $\sum_{i \leq n} t_{ij}^2 = 1$, то для j -го столбца матрицы T_R имеет место

$$(\rho_1^2 t_{1i}^2 + \dots + \rho_n^2 t_{ni}^2)^{\frac{k}{2}} \leq t_{1i}^2 \rho_1^k + \dots + t_{ni}^2 \rho_n^k.$$

Применив неравенство Иенсена для каждого столбца, получим:

$$(\rho_1^2 t_{11}^2 + \dots + \rho_n^2 t_{n1}^2)^{\frac{k}{2}} + \dots + (\rho_1^2 t_{1n}^2 + \dots + \rho_n^2 t_{nn}^2)^{\frac{k}{2}} \leq A,$$

где $A = (t_{11}^2 \rho_1^k + \dots + t_{n1}^2 \rho_n^k) + \dots + (t_{1n}^2 \rho_1^k + \dots + t_{nn}^2 \rho_n^k) = (t_{11}^2 + \dots + t_{1n}^2) \rho_1^k + \dots + (t_{n1}^2 + \dots + t_{nn}^2) \rho_n^k$. Поскольку матрица T ортогональна, то $(t_{11}^2 + \dots + t_{in}^2) \rho_i^k = \rho_i^k$ для всех $i \leq n$, и поэтому $A = \rho_1^k + \dots + \rho_n^k$.

Поскольку функция $f(x) = x^{\frac{k}{2}}$ монотонно возрастает при $k \geq 2$, то

$$\left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \leq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}}.$$

Отсюда и из предложения 2.2 получаем

$$A_k(\mathcal{E}) = \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \leq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}} = A_k(\mathcal{E}_V).$$

4. Доказательство теоремы 1.3 (2)

Доказательство неравенства (2) из теоремы 1.3 сводится к неравенству для ортогональных матриц:

$$\left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \geq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}} \text{ при } k \leq 2, k \neq 0; \quad (4)$$

и

$$\left(\prod_{i=1}^n s_i \right)^{\frac{1}{n}} \geq \left(\prod_{i=1}^n \rho_i \right)^{\frac{1}{n}} \text{ при } k = 0. \quad (5)$$

Сначала докажем неравенство (5), которое можно переписать в виде

$$\left(\prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}} \right)^{\frac{1}{n}} \geq \left(\prod_{i=1}^n \rho_i \right)^{\frac{1}{n}}$$

или в равносильном виде

$$\prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}} \geq \prod_{i=1}^n \rho_i.$$

Для доказательства последнего неравенства воспользуемся неравенством Адамара, примененным к матрице T_R . Учитывая, что $\det(T_R) = \det(T) \prod_{i=1}^n \rho_i$, а матрица T ортогональна, получим требуемое:

$$\prod_{i=1}^n \rho_i = |\det(T)| \prod_{i=1}^n \rho_i \leq \prod_{i=1}^n s_i = \prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}}.$$

Наконец, проводя рассуждения, аналогичные тем, что были приведены в предыдущем параграфе, докажем неравенство (4). Из (*) следует, что $s_i^k = (\rho_1^2 t_{1i}^2 + \dots + \rho_n^2 t_{ni}^2)^{\frac{k}{2}}$ не превосходит $t_{1i}^2 \rho_1^k + \dots + t_{ni}^2 \rho_n^k$ при $k < 0$; и не меньше $t_{1i}^2 \rho_1^k + \dots + t_{ni}^2 \rho_n^k$ при $0 < k < 2$. Отсюда и из того, что

матрица T ортогональна, следует, что

$$\sum_{i=1}^n s_i^k = (\rho_1^2 t_{11}^2 + \dots + \rho_n^2 t_{nn}^2)^{\frac{k}{2}} + \dots + (\rho_1^2 t_{1n}^2 + \dots + \rho_n^2 t_{nn}^2)^{\frac{k}{2}}$$

при $k < 0$ не превосходит

$$(t_{11}^2 + \dots + t_{1n}^2)\rho_1^k + \dots + (t_{n1}^2 + \dots + t_{nn}^2)\rho_n^k = \rho_1^k + \dots + \rho_n^k$$

(и при $0 < k < 2$ не меньше $\rho_1^k + \dots + \rho_n^k$).

Поскольку функция $f(x) = x^{\frac{k}{2}}$ монотонно возрастает при $0 < k < 2$ и монотонно убывает при $k < 0$, то

$$\left(\frac{\sum_{i=1}^n s_i^k}{n}\right)^{\frac{1}{k}} \geq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n}\right)^{\frac{1}{k}} \text{ при } 0 < k < 2 \text{ или } k < 0 - \text{неравенство (4) установлено.}$$

В заключение, автор выражает благодарность профессору С. М. Агееву, под руководством которого выполнена эта работа.

Литература

1. Horn R., Johnson Ch. *Matrix Analysis*. Cambridge: Cambridge University Press, 2014.
2. Мальцев А. И. Основы линейной алгебры. Москва: Лань, 2024.
3. Колмогоров А. Н., Петров А. А., Смирнов Ю. М. Одна формула Гаусса из теории метода наименьших квадратов // Изв. АН СССР. Сер. матем. 1947. Т. 11, № 6. С. 561–566.
4. Мальцев А. И. Замечание к работе А. Н. Колмогорова, А. А. Петрова и Ю. М. Смирнова // Изв. АН СССР. Сер. матем. 1947. Т. 11, № 6. С. 567–568.
5. Нисневич Л. Б., Брызгалов В. И. Об одной задаче n -мерной геометрии // УМН. 1953. Т. 8, № 4(56). С. 169–172.

References

1. Horn R., Johnson Ch. *Matrix Analysis*. Cambridge, Cambridge University Press, 2014.
2. Maltsev A. I. *Fundamentals of Linear Algebra*. Moscow, Lan, 2024 (in Russian).
3. Kolmogorov A. N., Petrov A. A., Smirnov Yu. M. One Gauss Formula from Least Squares Theory. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 1947, vol. 11, no. 6, pp. 561–566 (in Russian).
4. Maltsev A. I. A Remark on the Paper of A. N. Kolmogorov, A. A. Petrov, and Yu. M. Smirnov. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 1947, vol. 11, no. 6, pp. 567–568 (in Russian).
5. Nisnevich L. B., Bryzgalov V. I. On a Problem of N-Dimensional Geometry. *UMN*, 1953, vol. 8, no. 4(56), pp. 169–172 (in Russian).