Math-Net.Ru

Общероссийский математический портал

И. В. Оселедец, Оценки снизу для сепарабельных аппроксимаций ядра Гильберта, *Матем. сб.*, 2007, том 198, номер 3, 137–144

DOI: <https://doi.org/10.4213/sm1530>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.81.55.215

26 апреля 2023 г., 16:36:46

УДК 519.6+517.518.8

И. В. Оселедец

Оценки снизу для сепарабельных аппроксимаций ядра Гильберта

В работе получены асимптотически неулучшаемые оценки снизу на сепарабельные аппроксимации для функции $1/(x+y)$. Метод получения таких оценок основан на использовании обобщения принципа максимального объема для аппроксимации малого ранга.

Библиография: 10 названий.

§ 1. Введение

Сепарабельной аппроксимацией функции от двух переменных $f(x, y)$ называется такая функция $g(x, y)$, что

$$f(x, y) \approx g(x, y), \quad (x, y) \in \Omega$$

(в смысле некоторой нормы), и при этом

$$g(x, y) = \sum_{\alpha=1}^r u_{\alpha}(x)v_{\alpha}(y),$$

где число r будем называть *сепарабельным рангом* функции g .

В качестве нормы мы всегда будем брать C -норму в рассматриваемой области. Основной интерес для нас представляет оценка величин наилучшей сепарабельной аппроксимации

$$\sigma_{r+1} = \inf_{u_{\alpha}, v_{\alpha}} \left\| f(x, y) - \sum_{\alpha=1}^r u_{\alpha}(x)v_{\alpha}(y) \right\|. \quad (1)$$

От функций u_{α} и v_{α} потребуем, чтобы они были непрерывными.

Заметим, что если вместо C -нормы брать L_2 -норму, то σ_r будут сингулярными числами интегрального оператора вида

$$Lu = \int f(x, y)u(y) dy.$$

Асимптотическое поведение сингулярных чисел таких операторов исследовалось в работах [1], [2]. Отметим, что оценки, которые будут получены в настоящей статье, согласуются с уже известными. Отличие состоит в том, что

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты №№ 04-07-90336 и 05-01-00721) и программы фундаментальных исследований Отделения математических наук РАН “Вычислительные и информационные проблемы решения больших задач” по проекту “Матричные методы в интегральных и дифференциальных уравнениях”.

нам удалось получить количественные (а не просто асимптотические) оценки на величины σ_r для каждого r с помощью абсолютно другого подхода.

Сепарабельные аппроксимации играют большую роль в практике, в частности при построении мозаично-скелетонных аппроксимаций больших матриц [3], [4] и тензорных аппроксимаций больших матриц [5].

Для оценки величин σ_r сверху достаточно предъявить пример хорошей аппроксимации. Для этого, конечно, необходимо наложить какие-то условия на функцию f , например потребовать, чтобы она была асимптотически гладкой [3], [4]. В этом случае удается получить оценки вида $\sigma_r \sim e^{-\sqrt{r}}$. Для некоторых конкретных функций f можно получить более тонкие оценки сверху. В частности, для функции

$$f(x, y) = \frac{1}{x + y}$$

(мы будем называть такую функцию *ядром Гильберта*), рассматриваемой в прямоугольнике $[\delta, 1] \times [\delta, 1]$, предложено несколько способов сепарабельной аппроксимации [5], [6]. Естественно спросить: до каких пор можно улучшать оценки сепарабельного ранга функции $f(x, y)$? Ответу на этот вопрос для ядра Гильберта и посвящена эта статья. Основной результат – получение асимптотически неулучшаемой оценки на величины σ_r .

§ 2. Скелетное разложение

В дальнейшем изложении будем считать, что все рассматриваемые функции двух переменных определены на прямоугольнике $\Pi = [\delta, 1] \times [\delta, 1]$, а функции от одной переменной – на отрезке $l = [\delta, 1]$.

Пусть функция $g(x, y)$ имеет сепарабельный ранг r :

$$g(x, y) = \sum_{\alpha=1}^r u_{\alpha}(x)v_{\alpha}(y). \quad (2)$$

Возьмем некоторые точки $x_1, \dots, x_r \in l$ и $y_1, \dots, y_r \in l$. Тогда если матрица

$$M_{ij} = g(x_i, y_j), \quad i = 1, \dots, r, \quad j = 1, \dots, r,$$

размера $r \times r$ невырождена, то

$$g(x, y) = g_{\text{skel}}(x, y) = (g(x, y_1), \dots, g(x, y_r)) M^{-1} \begin{pmatrix} g(x_1, y) \\ \vdots \\ g(x_r, y) \end{pmatrix}. \quad (3)$$

Равенство (3) проверяется прямой подстановкой. Заметим, что (3) является прямым обобщением скелетного разложения для матриц. Напомним, что это такое. Пусть $A = (n \times n)$ -матрица ранга r , т.е.

$$A_{ij} = \sum_{\alpha=1}^r u_{i\alpha} v_{j\alpha}. \quad (4)$$

Тогда *скелетным разложением* матрицы A называется равенство вида

$$A = A_{\text{skel}} = C\hat{A}^{-1}R, \quad (5)$$

где C – это $(n \times r)$ -матрица, составленная из некоторых r столбцов матрицы A , R – $(r \times n)$ -матрица, составленная из некоторых r строк матрицы A , а $\hat{A}$ – невырожденная подматрица A , находящаяся на пересечении столбцов из C и строк из R .

Отметим прямую аналогию между (2) и (4), а также между (3) и (5). Отличие состоит в том, что вместо дискретных индексов i, j имеются их непрерывные аналоги x, y . Функции $g(x, y_s)$ являются “столбцами” с номерами y_s , а функции $g(x_s, y)$ – “строками” с номерами x_s .

Если же равенство (2) выполнено не точно, а лишь с некоторой погрешностью, то возникает вопрос о том, как выбрать узлы x_i, y_j так, чтобы получить наименьшую погрешность в аппроксимации функции g ее скелетным разложением (3). Для матриц ответ такой: нужно выбрать такие столбцы и строки, чтобы модуль определителя подматрицы $\hat{A}$ был максимальным [7]. Приведем аналогичный результат и для разложения (3).

ТЕОРЕМА 1. Пусть точки $x_i, y_j, i, j = 1, \dots, r$, выбраны так, что матрица M с элементами

$$M_{ij} = f(x_i, y_j)$$

имеет максимальный по модулю определитель среди всех матриц такого вида. Тогда

$$\sigma_{r+1} \leq \|f(x, y) - f_{\text{skel}}(x, y)\|_C \leq (r+1)^3 \sigma_{r+1}, \quad (6)$$

где σ_{r+1} определено в (1).

ДОКАЗАТЕЛЬСТВО. Оценка снизу следует прямо из определения σ_{r+1} . Для получения оценки сверху нужно всего лишь переписать доказательство из [7]. Для полноты изложения приведем это доказательство. Нам необходимо оценить остаток

$$h(x, y) = f - f_{\text{skel}}$$

для всех x, y . Возьмем произвольные точки x', y' и введем векторы

$$c = [f(x', y_i)], \quad b = [f(x_i, y')]$$

и число $a = f(x', y')$. Рассмотрим $((r+1) \times (r+1))$ -матрицу

$$\widehat{M} = \begin{pmatrix} M & b \\ c^T & a \end{pmatrix}$$

и определим $\gamma = a - c^T M^{-1}b$. Заметим, что

$$\gamma = h(x', y').$$

Если $\gamma = 0$, то $\widehat{M}$ вырождена и значение $h(x', y') = 0$. Поэтому осталось рассмотреть случай, когда $\gamma \neq 0$. Так как

$$|\det \widehat{M}^{-1}| |\det M| = |\gamma^{-1}|,$$

то в силу определения M

$$|\gamma^{-1}| = \|\widehat{M}^{-1}\|_C,$$

где через $\|\cdot\|_C$ обозначается максимальный по модулю элемент матрицы. Легко видеть, что

$$\sigma_{r+1}^{-1}(\widehat{M}) \leq \|\widehat{M}^{-1}\|_C(r+1),$$

где через $\sigma_{r+1}(\widehat{M})$ обозначено $(r+1)$ -е (наименьшее по модулю) сингулярное число матрицы $\widehat{M}$. Его можно оценить следующим образом:

$$\begin{aligned} \sigma_{r+1}(\widehat{M}) &= \min_{u_{i\alpha}, v_{j\alpha}} \left\| M - \sum_{\alpha=1}^r u_{i\alpha} v_{j\alpha} \right\|_F \\ &\leq (r+1)^2 \min_{u_{i\alpha}, v_{j\alpha}} \max_{i,j} \left| M_{ij} - \sum_{\alpha=1}^r u_{i\alpha} v_{j\alpha} \right|. \end{aligned}$$

Далее,

$$\begin{aligned} \min_{u_{i\alpha}, v_{j\alpha}} \max_{i,j} \left| M_{ij} - \sum_{\alpha=1}^r u_{i\alpha} v_{j\alpha} \right| &= \min_{u_\alpha(x), v_\alpha(y)} \max_{i,j} \left| \frac{1}{x_i + y_j} - \sum_{\alpha=1}^r u_\alpha(x_i) v_\alpha(y_j) \right| \\ &\leq \inf_{u_\alpha(x), v_\alpha(y)} \sup_{x,y} \left| \frac{1}{x+y} - \sum_{\alpha=1}^r u_\alpha(x) v_\alpha(y) \right| = \sigma_{r+1}, \\ \sigma_{r+1}(\widehat{M}) &\leq (r+1)^2 \sigma_{r+1}. \end{aligned}$$

Поэтому

$$|h(x', y')| = |\gamma| = \|\widehat{M}^{-1}\|_C^{-1} \leq (r+1)^3 \sigma_{r+1}.$$

Теорема 1 говорит о том, что хорошая сепарабельная аппроксимация функции f может быть получена с помощью скелетного разложения с подходящим выбором узлов.

Ядро Гильберта симметрично, поэтому для него естественно выбирать узлы x_i, y_i равными: $x_i = y_i, i = 1, \dots, r$. Можно ли тогда получить оценку, аналогичную (6)? Оказывается, что да. Известно, что скелетное разложение для матриц эквивалентно r шагам исключения Гаусса (а точнее, Гаусса–Жордана) с выбором ведущих элементов в позициях $(i_1, j_1), (i_2, j_2), \dots, (i_r, j_r)$, где i_s, j_s – номера строк и столбцов, выбранных в C и R из (5). Если матрица симметрична и положительно определена, то выбор главных элементов на диагонали матрицы ($i_s = j_s$) приводит к разложению Холецкого. Разложение Холецкого имеет несколько замечательных свойств, из которых нам понадобится лишь одно, а именно: матрица, получающаяся после нескольких шагов исключения Холецкого, остается положительно определенной, если исходная матрица была положительно определенной. А для любой положительно определенной симметричной матрицы $A = [a_{ij}]$ имеем

$$\det \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} = a_{ii}a_{jj} - a_{ij}^2 > 0,$$

поэтому

$$|a_{ij}| \leq \max_i |a_{ii}|,$$

т.е. максимальный по модулю элемент в положительно определенной симметричной матрице A всегда находится на ее диагонали. Оказывается, если мы будем выбирать для ядра Гильберта $x_i = y_i$, то для остатка

$$h(x, y) = f - f_{\text{skel}}$$

будет выполняться аналогичное свойство:

$$|h(x, y)| \leq \sup_{x \in I} |h(x, x)|, \quad (x, y) \in \Pi. \quad (7)$$

Это неравенство будет доказано в следующем параграфе, а пока, основываясь на нем, сформулируем следующую теорему.

ТЕОРЕМА 2. Пусть точки $x_i, i = 1, \dots, r$, выбраны так, что матрица

$$M_{ij} = f(x_i, x_j) = \frac{1}{x_i + x_j}$$

имеет максимальный по модулю определитель среди всех матриц такого вида. Тогда

$$\sigma_{r+1} \leq \|f(x, y) - f_{\text{skel}}(x, y)\|_C \leq (r+1)^3 \sigma_{r+1}, \quad (8)$$

где σ_{r+1} определено в (1).

ДОКАЗАТЕЛЬСТВО. Доказательство совпадает с доказательством теоремы 1 с той лишь разницей, что нам необходимо оценить остаток $h(x, y) = f - f_{\text{skel}}$ лишь при $x = y$, а потом воспользоваться неравенством (7).

Неравенства (8) будут использованы в дальнейшем для получения двусторонних оценок для σ_r .

§ 3. Скелетное разложение для ядра Гильберта

Оказывается, что для скелетного разложения (3) ядра Гильберта можно написать явные формулы.

ТЕОРЕМА 3. Пусть даны узлы $x_i, y_j, i, j = 1, \dots, r$. Тогда для скелетного разложения (3) функции $f(x, y) = 1/(x + y)$ выполняется равенство

$$f - f_{\text{skel}} = \frac{1}{x + y} \prod_{i=1}^r \frac{(x - x_i)(y - y_i)}{(x + y_i)(y + x_i)}. \quad (9)$$

ДОКАЗАТЕЛЬСТВО. Достаточно проверить, что

- (а) f_{skel} из (9) совпадает с f при $x = x_i, i = 1, \dots, r$, и при $y = y_i, i = 1, \dots, r$;
- (б) f_{skel} имеет сепарабельный ранг не выше r .

Условие (а), очевидно, верно. Для проверки условия (б) несколько преобразуем (9):

$$\begin{aligned} f_{\text{skel}} &= \frac{1}{x + y} \left(1 + \prod_{i=1}^r \frac{(x - x_i)(y - y_i)}{(x + y_i)(y + x_i)} \right) \\ &= \frac{1}{\prod_{i=1}^r (x + x_i)(y + y_i)} \left(\frac{1}{x + y} \left(\prod_{i=1}^r (x + y_i)(y + x_i) + \prod_{i=1}^r (x - x_i)(y - y_i) \right) \right). \end{aligned}$$

Заметим теперь, что выражение в больших скобках на самом деле является полиномом от x, y , так как

$$\prod_{i=1}^r (x + y_i)(y + x_i) + \prod_{i=1}^r (x - x_i)(y - y_i)$$

обращается в нуль при $x = -y$. Этот полином можно представить в виде

$$P(x, y) = \sum_{i=1}^r \sum_{j=1}^r p_{ij} x^i y^j.$$

Легко показать, что сепарабельный ранг $P(x, y)$ не выше r . Действительно,

$$P(x, y) = \sum_{i=1}^r x^i \sum_{j=1}^r p_{ij} y^j.$$

Тем самым мы доказали условие (6).

Теперь мы можем доказать неравенство (7). Действительно, при $x_i = y_i$ имеем

$$h(x, y) = f - f_{\text{skel}} = \frac{1}{x + y} \prod_{i=1}^r \frac{(x - x_i)(y - x_i)}{(x + x_i)(y + x_i)},$$

поэтому

$$h(x, x)h(y, y) = h(x, y)^2 \frac{(x + y)^2}{4xy} \geq h(x, y)^2,$$

откуда сразу следует (7).

§ 4. Оптимальный выбор узлов и окончательная оценка для σ_r

Пусть $x_i, i = 1, \dots, r$, – некоторые узлы и f_{skel} – скелетное разложение ядра Гильберта по этим узлам. Тогда для остатка

$$h(x, y) = f - f_{\text{skel}} = \frac{1}{x + y} \prod_{i=1}^r \frac{(x - x_i)(y - x_i)}{(x + x_i)(y + x_i)}$$

можно поставить задачу *об оптимальном выборе узлов*, а именно нужно найти такие узлы x_i , чтобы функция $h(x, y)$ наименее уклонялась от нуля на Π . Так как мы уже знаем, что максимум $h(x, y)$ достигается при $x = y$, то получаем следующую задачу: среди всех функций вида

$$g(x) = \sqrt{h(x, x)} = \frac{1}{\sqrt{2x}} \prod_{i=1}^r \frac{x - x_i}{x + x_i} \quad (10)$$

найти ту, которая наименее уклоняется от нуля на отрезке $[\delta, 1]$. Несколько ослабим задачу, окончательно поставив ее следующим образом: среди всех функций вида

$$g(x) = \prod_{i=1}^r \frac{x - x_i}{x + x_i} \quad (11)$$

найти ту, которая наименее уклоняется от нуля на отрезке $[\delta, 1]$. Наименьшие уклонения E' и E для задач (10) и (11) связаны следующим очевидным неравенством:

$$\frac{E}{\sqrt{2}} \leq E' \leq \frac{E}{\sqrt{2\delta}}. \quad (12)$$

Решением экстремальной задачи (11) является функция Золотарёва с узлами x_i , определяемыми как

$$x_i = \operatorname{dn}\left(\frac{2(r-i)+1}{2r}K(\delta'), \delta'\right), \quad (13)$$

где $\operatorname{dn}(u, k)$ – эллиптическая функция Якоби, $K(k)$ – эллиптический интеграл и $\delta' = \sqrt{1-\delta^2}$. Золотарёвым было доказано, что функция вида (11) с узлами (13) имеет $r+1$ точку альтернанса на отрезке $[\delta, 1]$ и является наименее уклоняющейся от нуля среди всех функций вида (11) на этом отрезке [8], [9].

Нас интересует величина наименьшего отклонения $E = E_r(\delta)$. Для нее известна (см. [10]) следующая формула:

$$E_r(\delta) = 2q^{r/4} \prod_{n=1}^{\infty} (1 + q^{rn})^{(-1)^n},$$

где

$$q = \exp\left(-\pi \frac{K(\mu_1)}{K(\mu)}\right), \quad \mu = \left(\frac{1 - \delta^{1/2}}{1 + \delta^{1/2}}\right)^2, \quad \mu_1 = \sqrt{1 - \mu^2}. \quad (14)$$

Величина E_r оценивается следующим образом:

$$\frac{2q^{r/4}}{1 + q^{r/2}} \leq E_r \leq 2q^{r/4}. \quad (15)$$

Объединяя неравенства (15), (12), (8), получаем финальную оценку

$$\frac{\sqrt{2}q^{r/4}}{(1 + q^{r/2})(r+1)^3} \leq \sigma_{r+1} \leq \frac{\sqrt{2}q^{r/4}}{\delta}. \quad (16)$$

На практике обычно требуется не зависимость $\sigma_r(\delta)$, а обратная зависимость $r(\delta, \varepsilon)$, т.е. число слагаемых, необходимых для приближения функции с заданной (малой) точностью ε . При этом интересен случай δ близких к нулю. Аналогично работе [10] получается следующая асимптотическая формула для $r = r(\delta, \varepsilon)$:

$$r = O(\ln \varepsilon \ln \delta).$$

Опуская доказательство, отметим лишь, что оно основано на известной асимптотике эллиптического интеграла:

$$K(\mu) \approx \frac{\pi}{2}, \quad K(\mu_1) \approx \ln \frac{4}{\mu_1} \quad \text{при } \mu \rightarrow 1.$$

§ 5. Заключение

Полученные в настоящей работе оценки снизу на ошибку сепарабельной аппроксимации для ядра Гильберта являются неулучшаемыми по порядку, поэтому поиск новых способов сепарабельной аппроксимации может лишь улучшить константы, входящие в эти неравенства. Более того, предложен конструктивный способ построения почти оптимальной сепарабельной аппроксимации, основанный на обобщении скелетного разложения для матриц. В дальнейшем будет сделана попытка обобщить полученные результаты на другие функции, в частности на функцию

$$f(x, y) = \log(x + y).$$

В заключение автор выражаю особую благодарность своему научному руководителю Е. Е. Тыртышникову за внимательное отношение и В. И. Лебедеву за помощь в решении экстремальной задачи (11).

Список литературы

- [1] А. А. Лаптев, “Спектральная асимптотика одного класса интегральных операторов”, *Матем. заметки*, **16**:5 (1974), 741–750; англ. пер.: A. A. Laptev, “Spectral asymptotic behavior of a certain class of integral operators”, *Math. Notes*, **16**:5 (1974), 1038–1043.
- [2] H. Widom, “Hankel matrices”, *Trans. Amer. Math. Soc.*, **121**:1 (1966), 1–35.
- [3] Е. Е. Тыртышников, “Mosaic-skeleton approximations”, *Calcolo*, **33**:1–2 (1996), 47–57.
- [4] С. А. Горейнов, “Мозаично-скелетонные аппроксимации матриц, порожденных асимптотически гладкими и осцилляционными ядрами”, *Матричные методы и вычисления*, ИВМ РАН, М., 1999, 42–76.
- [5] W. Hackbusch, B. N. Khoromskii, Е. Е. Тыртышников, “Hierarchical Kronecker tensor-product approximations”, *J. Numer. Math.*, **13**:2 (2005), 119–156.
- [6] D. Braess, W. Hackbusch, “Approximation of $1/x$ by exponential sums in $[1, \infty)$ ”, *IMA J. Numer. Anal.*, **25**:4 (2005), 685–697.
- [7] S. A. Goreinov, Е. Е. Тыртышников, “The maximal-volume concept in approximation by low-rank matrices”, *Structured matrices in mathematics, computer science, and engineering I* (Boulder, CO, 1999), *Contemp. Math.*, **280**, Amer. Math. Soc., Providence, RI, 2001, 47–51.
- [8] Е. И. Золотарёв, *Собрание сочинений*, т. XXX, вып. 5, Санкт-Петербургская академия наук, 1877.
- [9] А. А. Гончар, “О задачах Е. И. Золотарёва, связанных с рациональными функциями”, *Матем. сб.*, **78(120)**:4 (1969), 640–654; англ. пер.: A. A. Gončar, “Zolotarev problems connected with rational functions”, *Math. USSR-Sb.*, **7**:4 (1969), 623–635.
- [10] А. А. Медовиков, В. И. Лебедев, “Optimization of L_ω stable Crank–Nicolson method”, *Russian J. Numer. Anal. Math. Modelling*, **20**:3 (2005), 283–304.

И. В. Оселедец (I. V. Oseledets)
 Институт вычислительной математики РАН,
 г. Москва
E-mail: oseledets1983@yandex.ru

Поступила в редакцию
 16.02.2006