Записки научных
семинаров ПОМИ
Том 466, 2017 г.

М. С. Ермаков

МИНИМАКСНОЕ НЕПАРАМЕТРИЧЕСКОЕ ОЦЕНИВАНИЕ НА НАИБОЛЬШИХ МНОЖЕСТВАХ

§1. Введение

Для задачи непараметрического оценивания сигнала в гауссовском шуме оптимальный порядок скорости сходимости оценок изучен для широкого класса функциональных пространств и совершенно различных постановок задач (см. [4, 7, 9, 12, 16, 21] и ссылки в них). В то же время асимптотически минимаксные оценки с точной асимптотикой скорости сходимости известны, только если дана априорная информация, что сигнал принадлежит эллипсоиду в L_2 [9, 14, 17, 18, 21], шарам в L_∞ [1, 3, 13, 15] или телам в пространствах Бесова [4, 9]. Цель статьи – обратить внимание, что такие оценки могут быть получены и для других множеств функций. Для тригонометрической ортогональной системы функций эти множества могут рассматриваться как шары в пространстве Бесова $B^\alpha_{2\infty}$ для некоторой нормы. Мы обозначим эти множества $B(\alpha, P_0)$, где $\alpha > 0$ и $P_0 > 0$.

Шары $B(\alpha, P_0)$ обладают замечательными свойствами в задачах непараметрического оценивания. Они несут разумную информацию о гладкости сигнала.

На этих множествах наиболее распространенные линейные непараметрические оценки имеют заданную скорость сходимости [11, 12]. Для широкого класса линейных оценок они являются наибольшими множествами с заданной скоростью сходимости [11, 19].

Ранее задачи непараметрического оценивания решений линейных некорректных обратных задач в гауссовском шуме на множествах $B(\alpha, P_0)$ рассматривались в задачах эконометрики [10].

Возникающие асимптотически минимаксные оценки являются оценками максимума правдоподобия с квадратичной штрафной функцией

Ключевые слова: алгоритм регуляризации Тихонова, оценка максимума правдоподобия, асимптотически минимаксное оценивание, непараметрическое оценивание.

Работа поддержана грантом РФФИ 17-01-00828.

[5, 9, 22]. Таким образом, оказывается, что оценки максимума правдоподобия с квадратичной штрафной функцией являются оптимальными не только как оценки байесовского отношения правдоподобия [9, 21, 22], но и в минимаксном смысле. Полученные результаты можно рассматривать и как решение обратной задачи: найти для байесовских оценок и оценок максимума правдоподобия со штрафной функцией наибольшие множества, на которых эти оценки являются асимптотически минимаксными.

В работе исследуется также неасимптотическая постановка задачи. Для нее мы показываем, что рассматриваемые оценки минимаксны в классе линейных оценок.

Результаты легко переносятся на случай минимаксного оценивания решений линейных некорректно поставленных задач. Для этой постановки асимптотически минимаксные оценки являются оценками алгоритма регуляризации Тихонова [20].

Мы показываем, что оценки Пинскера [9, 18, 21] имеют на множествах $B(\alpha, P_0)$ порядок сходимости хуже, чем асимптотически минимаксные оценки на $B(\alpha, P_0)$, и для них наибольшими множествами являются шары в пространстве Соболева.

Результаты будут даны в терминах коэффициентов Фурье оцениваемого сигнала. В этой постановке задача допускает следующую интерпретацию. Мы наблюдаем случайную последовательность $y = \{y_j\}_{j=1}^{\infty}$,

$$y_j = x_j + \varepsilon\, \sigma_j\, \xi_j, \quad \varepsilon > 0, \quad 1 \leqslant j < \infty,$$

где $\sigma_j > 0$ – известные постоянные и ξ_j, $1 \leqslant j < \infty$, – независимые гауссовские случайные величины, $\mathbf{E}\, \xi_j = 0$ и $\mathbf{E}\, \xi_j^2 = 1$.

Нужно оценить параметр $x = \{x_j\}_{j=1}^{\infty}$. Обозначим $\sigma = \{\sigma_j\}_{j=1}^{\infty}$ и $\xi = \{\xi_j\}_{j=1}^{\infty}$. В задаче оценивания с фиксированным $\varepsilon > 0$ предположим, что имеется следующая априорная информация

$$x \in B = B(a, P_0) = \Big\{ x = \{x_i\}_{i=1}^{\infty} : \sup_k a_k^{-1} \sum_{j=k}^{\infty} x_j^2 \leqslant P_0 \Big\}, \tag{1.1}$$

где $a = \{a_k\}_{k=1}^{\infty}$ и $a_k > 0$ – строго убывающая последовательность.

Для задачи асимптотически минимаксного оценивания нами будет рассматриваться более узкий класс множеств $B(\alpha, P_0) = B(\widetilde{a}, P_0)$, где $\widetilde{a} = \{k^{-2\alpha}\}$, $\alpha > 0$. Результаты могут быть распространены и на другие последовательности a_k. Однако это требует более аккуратных оценок. Для ортогональной тригонометрической системы функций $\sup\limits_k a_k^{-1} \sum\limits_{j=k}^{\infty} x_j^2$ может рассматриваться в качестве нормы в пространстве Бесова $B^{\alpha}_{2\infty}$. Для тел Бесова в $B^{\alpha}_{2\infty}$, порожденных вейвлетами, асимптотически минимаксные оценки можно найти в книге Джонстоуна [9]. Однако для рассматриваемой нами постановки возникает другая экстремальная задача и получается другое решение.

Есть много публикаций по адаптивному асимптотически минимаксному оцениванию (см. [9, 21] и ссылки в них). Результаты по адаптивному оцениванию для оценок Пинскера [9, 21] легко переносятся на данную постановку задачи для рассматриваемых асимптотически минимаксных оценок.

Напомним определение наибольшего множества [10, 11, 12, 19]. Для оценки $\widehat{x}_\varepsilon$, функции потерь $\|\widehat{x}_\varepsilon - x\|^2$, порядка скорости сходимости ε^γ, $\gamma > 0$, и постоянной $C > 0$, наибольшим множеством является

$$\mathrm{MS}(\widehat{x}_\varepsilon, \gamma)(C) = \{x : \sup_\varepsilon \varepsilon^{-2\gamma}\, \mathbf{E}_x \|\widehat{x}_\varepsilon - x\|^2 < C\}.$$

Здесь $\|x\|$ обозначает норму вектора $x = \{x_j\}_{j=1}^\infty$ в гильбертовом пространстве,

$$\|x\|^2 = \sum_{j=1}^\infty x_j^2.$$

В дальнейшем буквы c, C обозначают положительные постоянные и $a_\varepsilon \asymp b_\varepsilon$ означает, что $c < a_\varepsilon / b_\varepsilon < C$ для всех $\varepsilon > 0$.

§2. Основные результаты

Мы скажем, что линейная оценка $\widehat{x}_\varepsilon = \{\widehat{x}_{\varepsilon j}\}_{j=1}^\infty$ является минимаксной в классе линейных оценок $\widehat{x}_{\varepsilon\lambda} = \{\widehat{x}_{\varepsilon j \lambda_j}\}_{j=1}^\infty$, $\widehat{x}_{\varepsilon j\lambda_j} = \lambda_j\, y_j$, $\lambda_j \in R^1$, $1 \leqslant j < \infty$, $\lambda = \{\lambda_j\}_{j=1}^\infty$, если

$$R_{l\varepsilon} \doteq \sup_{x \in B(a, P_0)} \mathbf{E}_x \|\widehat{x}_\varepsilon - x\|^2 = \inf_\lambda \sup_{x \in B(a, P_0)} \mathbf{E}_x \|\widehat{x}_{\varepsilon\lambda} - x\|^2. \tag{2.1}$$

Мы скажем, что оценка $\widehat{x}_\varepsilon$ является асимптотически минимаксной, если

$$R_\varepsilon \doteq \sup_{x \in B(\alpha, P_0)} \mathbf{E}_x \|\widehat{x}_\varepsilon - x\|^2 = \inf_{\widetilde{x}_\varepsilon \in \Psi} \sup_{x \in B(\alpha, P_0)} \mathbf{E}_x \|\widetilde{x}_\varepsilon - x\|^2 \,(1 + o(1)) \tag{2.2}$$

при $\varepsilon \to 0$. Здесь Ψ обозначает множество всех оценок.

Минимаксная оценка в классе линейных оценок будет найдена при следующих предположениях.

A1. Найдется такое $c > 0$, что $c < \sigma_j^2 < \infty$ для всех j.

A2. Для всех $j > 1$

$$\frac{\sigma_j^2\,(a_{j-1} - a_j)}{\sigma_{j-1}^2\,(a_j - a_{j+1})} > 1. \tag{2.3}$$

Это означает, что последовательность $\sigma_j^2\,(a_{j-1} - a_j)$ – строго возрастающая.

Теорема 1. *Пусть выполнены условия* A1, A2. *Тогда линейная оценка* $\widehat{x}_\lambda$ *с координатными коэффициентами*

$$\lambda_j = \frac{P_0\,(a_j - a_{j+1})}{P_0\,(a_j - a_{j+1}) + \varepsilon^2\,\sigma_j^2} \tag{2.4}$$

является минимаксной в классе линейных оценок.

Минимаксный риск равен

$$R_{l\varepsilon} = \varepsilon^2 \sum_{j=1}^{\infty} \frac{P_0\,\sigma_j^2\,(a_j - a_{j+1})}{P_0\,(a_j - a_{j+1}) + \varepsilon^2\,\sigma_j^2}. \tag{2.5}$$

Замечание 1. Оценка $\widehat{x}_\lambda$ является оценкой максимума правдоподобия [5, 22] с квадратичной штрафной функцией

$$P_0^{-1} \sum_{j=1}^{\infty} (a_j - a_{j+1})^{-1} \sigma_j^2\, x_j^2.$$

и байесовской оценкой с априорной мерой, при которой координаты x_j, $1 \leqslant j < \infty$, – независимые нормальные случайные величины, $\mathbf{E} x_j = 0$ и $\mathbf{E}\, x_j^2 = P_0\,(a_j - a_{j+1})$.

Замечание 2. Теорема 1 справедлива также для конечного числа наблюдений $y_1, \ldots, y_k$, где $k < \infty$.

В теореме 2 мы заменим условие A2 более простым условием

B1. Для всех $j > j_0$

$$\frac{\sigma_j^2\, j^{2\alpha+1}}{\sigma_{j-1}^2\,(j-1)^{2\alpha+1}} > 1. \tag{2.6}$$

Это означает, что последовательность $\sigma_j^2\, j^{2\alpha+1}$ – строго возрастающая.

Теорема 2. *Пусть выполнены условия* A1, B1. *Тогда линейная оценка* $\widehat{x}_\lambda$ *с покоординатными коэффициентами*

$$\lambda_j = \frac{2\,\alpha\,P_0\,j^{-2\alpha-1}}{2\,\alpha\,P_0\,j^{-2\alpha-1}+\varepsilon^2\,\sigma_j^2} \tag{2.7}$$

является асимптотически минимаксной в классе всех оценок.

Асимптотически минимаксный риск равен

$$R_\varepsilon = \varepsilon^2 \sum_{j=1}^{\infty} \frac{2\,\alpha\,P_0\,j^{-2\alpha-1}\,\sigma_j^2}{2\,\alpha\,P_0\,j^{-2\alpha-1}+\varepsilon^2\,\sigma_j^2}\,(1+o(1)). \tag{2.8}$$

Замечание 3. Оценка $\widehat{x}_\lambda$ является оценкой максимума правдоподобия [5, 22] с квадратичной штрафной функцией

$$(2\,\alpha\,P_0)^{-1} \sum_{j=1}^{\infty} j^{1+2\alpha}\,\sigma_j^2\,x_j^2$$

и байесовской оценкой с априорной мерой, при которой координаты x_j, $1 \leqslant j < \infty$, – независимые нормальные случайные величины, $\mathbf{E}\,x_j = 0$ и $\mathbf{E}\,x_j^2 = 2\,\alpha\,P_0\,j^{-1-2\alpha}$.

Теоремы 1 и 2 легко распространяются на оценивание решения линейной некорректной обратной задачи. Отметим, что наибольшие множества в этой постановке изучались в работе [10].

Предположим, что мы наблюдаем случайный вектор

$$y = R\,x + \varepsilon\,\xi,$$

где $R : H \to H$ – линейный самосопряженный оператор в сепарабельном гильбертовом пространстве H. Остальные обозначения – те же самые, что и в предыдущей постановке. Предположим, что линейный оператор R имеет дискретный спектр с собственными числами r_j, $1 \leqslant j < \infty$. Тогда мы можем рассмотреть следующую постановку задачи (см. [8, 9, 10, 21]).

Мы наблюдаем случайный вектор

$$z_j = r_j\,x_j + \varepsilon\,\sigma_j\,\xi_j, \quad 1 \leqslant j < \infty,$$

где ξ_j, $1 \leqslant j < \infty$, – независимые гауссовские случайные величины, $\mathbf{E}\,\xi_j = 0$, $\mathbf{E}\,\xi_j^2 = 1$. Стоит та же самая задача оценивания вектора $x = \{x_j\}_{j=1}^{\infty}$. Делением на r_j она сводится к задаче оценивания сигнала.

Ниже приведены две асимптотики минимаксного риска для линейных некорректных обратных задач.

Пример 1. Пусть $|r_j| = C\, j^{-\gamma}\,(1 + o(1))$ и $\sigma_j = 1$, $1 \leqslant j < \infty$. Тогда для $\alpha > 0$, $\gamma > 0$ имеет место

$$R_\varepsilon = \varepsilon^{\frac{4\alpha}{1+2\alpha+2\gamma}}\, \frac{\pi}{2\,\alpha\,\sin\bigl(\frac{\pi\,(2\gamma+1)}{2\alpha}\bigr)}\,(2\,\alpha\,P_0)^{\frac{2\gamma+1}{2\gamma+2\alpha+1}}\,C^{-\frac{2\alpha}{2\gamma+2\alpha+1}}\,(1+o(1)). \tag{2.9}$$

Пример 2. Пусть $|r_j| = C\, j^{\kappa}\,\exp\{-B\, j^{\gamma}\}$ и $\sigma_j = 1$, $1 \leqslant j < \infty$. Тогда для $\alpha > 0$, $\gamma > 0$, $B > 0$, $\kappa \in R^1$ имеет место

$$R_\varepsilon = P_0\, B^{2\alpha/\gamma}\,|\log\varepsilon|^{-2\alpha/\gamma}\,(1+o(1)). \tag{2.10}$$

Отметим, что эти асимптотики совпадают с асимптотиками рисков в соответствующей байесовской постановке.

Джонстоун (теорема 3.10, гл. 3, [9]) сравнил точные асимптотики минимаксного риска оценок, основанных на тригонометрических сплайнах, и оценок Пинскера, если неизвестный сигнал принадлежит шару в пространстве Соболева. Оценки, основанные на сплайнах, совпадают с асимптотически минимаксными оценками на наибольших множествах $B(\alpha, P_0)$. Таким образом, мы можем рассматривать этот результат как сравнение асимптотик рисков асимптотически минимаксных оценок на наибольших множествах $B(\alpha, P_0)$ и оценок Пинскера. Ниже мы проведем аналогичное сравнение, если имеется априорная информация, что сигнал принадлежит наибольшему множеству $B(\alpha, P_0)$.

Оценка Пинскера $\widetilde{\theta}_{\varepsilon\mu} = \{\widetilde{\theta}_{\varepsilon j}\}_{j=1}^{\infty}$ является линейной оценкой

$$\widetilde{\theta}_{\varepsilon j} = \lambda_{\varepsilon j}\, y_j,$$

где

$$\lambda_{\varepsilon j} = (1 - \mu\, b_j)_+$$

и $b_j = j^{\beta}$, $\beta > 0$. Параметр μ определяется уравнением

$$\varepsilon^2 \sum_{j=1}^{\infty} b_j^2\,((\mu\, b_j)^{-1} - 1)_+ = P.$$

Оценка Пинскера является асимптотически минимаксной на эллипсоидах (см. [9, 17, 18, 21])

$$S(\beta, P) = \Bigl\{x : \sum_{j=1}^{\infty} b_j^2\, x_j^2 \leqslant P, \quad x = \{x_j\}_1^{\infty}\Bigr\},$$

где $P > 0$.

Обозначим

$$R_\varepsilon(\alpha,\beta) = \inf_\mu \sup_{\theta\in B(\alpha,P_0)} \mathbf{E}_\theta \, \|\widetilde{\theta}_{\varepsilon\mu} - \theta\|^2$$

и

$$C = \frac{2\alpha^2}{(1+\alpha)(1+2\alpha)}.$$

Теорема 3. *Пусть* $0<\alpha<\beta$. *Тогда*

$$R_\varepsilon(\alpha,\beta) = C^{\frac{2\alpha}{1+2\alpha}}\, C_1^{\frac{1}{1+2\alpha}} \left((2\alpha)^{-\frac{2\alpha}{1+2\alpha}} + (2\alpha)^{\frac{1}{1+2\alpha}}\right) \varepsilon^{\frac{4\alpha}{1+2\alpha}}, \tag{2.11}$$

где $C_1 = \frac{\beta}{\beta-\alpha} P_0$.

Пусть $\alpha>\beta>0$. *Тогда*

$$R_\varepsilon(\alpha,\beta) = C^{\frac{2\beta}{1+2\beta}}\, C_1^{\frac{1}{1+2\beta}} \left((2\beta)^{-\frac{2\beta}{1+2\beta}} + (2\beta)^{\frac{1}{1+2\beta}}\right) \varepsilon^{\frac{4\beta}{1+2\beta}}, \tag{2.12}$$

где

$$C_1 = \sum_{j=1}^{\infty} j^{2\beta}\big(j^{-2\alpha} - (j+1)^{-2\alpha}\big).$$

Если $\alpha=\beta$, *то*

$$R_\varepsilon(\alpha,\beta) = C_\alpha\, P_0^{\frac{1}{1+2\alpha}}\, C^{\frac{2\alpha}{1+2\alpha}}\, \varepsilon^{\frac{4\alpha}{1+2\alpha}}\, |2\ln\varepsilon|^{\frac{1}{1+2\alpha}}, \tag{2.13}$$

где

$$C_\alpha = \left((2\alpha^2)^{\frac{1}{1+2\alpha}} + 2^{-\frac{2\alpha}{1+2\alpha}}\alpha^{\frac{1-2\alpha}{1+2\alpha}}\right)(1+2\alpha)^{-\frac{1}{1+2\alpha}}.$$

Наибольший интерес представляет сравнение рисков оценок Пинскера и асимптотически минимаксных оценок на наибольших множествах при $\alpha=\beta$. Для этой постановки задачи мы сравниваем риски оценок на множествах, имеющих почти ту же самую гладкость. Риски оценок Пинскера на наибольших множествах $B(\alpha,P_0)$ имеют дополнительный логарифмический член асимптотики. Оценки Пинскера не принадлежат классу линейных оценок, имеющих наибольшими множествами множества $B(\alpha,P_0)$. Оказывается, что шары в пространстве Соболева являются наибольшими множествами для оценок Пинскера.

Теорема 4. *Найдется такое* $C<\infty$, *что для всех* $\varepsilon>0$

$$R_\varepsilon(\beta,x) = \varepsilon^{-\frac{4\beta}{1+2\beta}} \inf_\mu \mathbf{E}_x \, \|\widetilde{x}_{\varepsilon\mu} - x\|^2 < C, \tag{2.14}$$

если и только если, x *принадлежит пространству Соболева*

$$S^\beta = \Big\{x:\ \sum_{j=1}^{\infty} b_j^2\, x_j^2 < \infty,\ x=\{x_j\}_{j=1}^\infty\Big\}.$$

В теории линейных некорректных обратных задач часто предполагается, что решение x удовлетворяет условию источника [2]

$$x \in \{x : x = Bu,\ \|u\| \leqslant 1,\ u \in H\},$$

где B – линейный самосопряженный компактный оператор. Это означает, что решение x принадлежит эллипсоиду. Теоремы 3 и 4 показывают, что оптимальные линейные решения, найденные на таких множествах, могут иметь худший порядок скорости сходимости на более широких множествах, чем другие линейные оценки.

§3. Доказательство теорем

3.1. Доказательство теоремы 1. Начнем с доказательства границы снизу. Обозначим $\theta_j^2 = P_0(a_j - a_{j+1})$ и $\theta = \{\theta_j\}_{j=1}^{\infty}$.

Имеем

$$\inf_{\lambda} \sup_{x \in B} \mathbf{E}_x \|\widehat{x}_\lambda - x\|^2 \geqslant \inf_{\lambda} \mathbf{E}_\theta \|\widehat{x}_{\varepsilon\lambda} - \theta\|^2 = \varepsilon^2 \sum_{j=1}^{\infty} \frac{\theta_j^2\, \sigma_j^2}{\theta_j^2 + \varepsilon^2\, \sigma_j^2}, \tag{3.1}$$

и инфимум достигается для

$$\lambda_j = \frac{\theta_j^2}{\theta_j^2 + \varepsilon^2\, \sigma_j^2} = \frac{P_0\,(a_j - a_{j+1})}{P_0\,(a_j - a_{j+1}) + \varepsilon^2\, \sigma_j^2}.$$

Доказательство верхней границы основано на следующих рассуждениях. Пусть $x = \{x_j\}_{j=1}^{\infty} \in B$. Для всех k обозначим

$$u_k = a_k^{-1} \sum_{j=k}^{\infty} x_j^2.$$

Тогда $x_k^2 = a_k u_k - a_{k+1} u_{k+1}$.

Для последовательности λ_j, определенной в теореме 1, имеем

$$\begin{aligned}
\mathbf{E}_x \sum_{j=1}^{\infty} (\lambda_j\, y_j - x_j)^2 &= \varepsilon^2 \sum_{j=1}^{\infty} \lambda_j^2\, \sigma_j^2 + \sum_{j=1}^{\infty} (1 - \lambda_j)^2\, x_j^2 \\
&= \varepsilon^2 \sum_{j=1}^{\infty} \lambda_j^2\, \sigma_j^2 + \sum_{j=1}^{\infty} (\theta_j^2\, \sigma_j^{-2}\, \varepsilon^{-2} + 1)^{-2}\, (a_j\, u_j - a_{j+1}\, u_{j+1}) \\
&= \varepsilon^2 \sum_{j=1}^{\infty} \lambda_j^2\, \sigma_j^2 + (\theta_1^2\, \sigma_1^{-2}\, \varepsilon^{-2} + 1)^{-2}\, u_1
\end{aligned} \tag{3.2}$$

$$-\sum_{j=2}^{\infty} u_j\, a_j\, \big((\theta_{j-1}^2\, \sigma_{j-1}^{-2}\, \varepsilon^{-2}+1)^{-2}-(\theta_j^2\, \sigma_j^{-2}\, \varepsilon^{-2}+1)^{-2}\big).$$

В силу условия A2, последнее слагаемое в правой части (3.2) отрицательно. Следовательно максимум правой части (3.2) достигается для $u_j = P_0$, $1 \leqslant j < \infty$. Теорема 1 доказана. □

3.2. Доказательство теоремы 2. Граница сверху следует из теоремы 1. Ниже приведено доказательство границы снизу. Оно во многих чертах похоже на доказательство нижней границы в теореме Пинскера [9, 18, 21].

Зафиксируем значения δ_1, $0<\delta_1<1$, и δ, $0<\delta<P_0$. Определим семейство натуральных чисел k_ε, $\varepsilon>0$, такое что $2\,\alpha\, P_0\, \varepsilon^{-2}\, \sigma_{k_\varepsilon}^2\, k_\varepsilon^{-2\alpha-1} = 1+o(1)$ при $\varepsilon\to 0$. Определим последовательность гауссовских независимых случайных величин $\eta=\{\eta_j\}_{j=1}^{\infty}$, где $\eta_j=\eta_{j\delta\delta_1}$, $\mathbf{E}\,[\eta_j]=0$, $\mathbf{E}\,[\eta_j^2] = 2\,\alpha\,(P_0-\delta)\, j^{-2\alpha-1}$, если $\delta\, k_\varepsilon \leqslant j \leqslant \delta^{-1}\, k_\varepsilon$, и $\eta_j=0$, если $j<\delta_1\, k_\varepsilon$ или $j>\delta_1^{-1}\, k_\varepsilon$.

Обозначим μ вероятностную меру η. Обозначим $\widetilde{x}$ байесовскую оценку с априорной мерой μ.

Определим условную вероятностную меру ν_δ случайного вектора η при условии $\eta\in B(\alpha,P_0)$. Обозначим $\bar{x}$ байесовскую оценку вектора x, имеющую априорную меру ν_δ. Обозначим θ случайную величину, имеющую вероятностную меру ν_δ.

Для любой оценки $\widehat{x}$ имеем

$$\begin{aligned}\sup_{x\in B(\alpha,P_0)} \mathbf{E}_x\, \|\widehat{x}-x\|^2 &\geqslant \mathbf{E}_{\nu_\delta}\, \mathbf{E}_\theta\, \|\widehat{x}-\theta\|^2 \geqslant \mathbf{E}_\mu\, \mathbf{E}_\eta\, \|\widetilde{x}-\eta\|^2 \\ &- \mathbf{E}_\mu\, \mathbf{E}_\eta\, (\|\bar{x}-\eta\|^2,\, \eta\notin B(\alpha,P_0))\, \mathbf{P}_\mu^{-1}\, (\eta\in B(\alpha,P_0)).\end{aligned} \tag{3.3}$$

Имеем

$$\mathbf{E}_\mu\, \mathbf{E}_\eta\, \|\widetilde{x}-\eta\|^2 = I(P_0-\delta)(1+o(1)), \tag{3.4}$$

где

$$I(P_0-\delta) = \varepsilon^2 \sum_{j=l_1}^{l_2} \frac{\sigma_j^2}{1+(2\,\alpha\,(P_0-\delta_1))^{-1}\, \varepsilon^2\, \sigma_j^2\, j^{2\alpha+1}}$$

и $l_1=[\delta_1 k_\varepsilon]$, $l_2=[\delta_1^{-1} k_\varepsilon]$. Здесь $[a]$ обозначает целую часть числа $a\in R^1$.

Так как

$$\|\bar{x}\|^2 \leqslant \sup_{x\in B(\alpha,P_0)} \|x\|^2 \leqslant P_0,$$

то

$$\mathbf{E}_\mu\, \mathbf{E}_\eta\, (\|\bar{x}-\eta\|^2,\, \eta \notin B(\alpha, P_0)) \leqslant 2\, \mathbf{E}_\mu\, \mathbf{E}_\eta\, (\|\bar{x}\|^2 + \|\eta\|^2,\ \eta \notin B(\alpha, P_0))$$

$$\leqslant 2\, P_0\, \mathbf{P}_\mu\, (\eta \notin B(\alpha, P_0)) + \sum_{j=l_1}^{l_2} (\mathbf{E}_\mu\, \eta_j^4)^{1/2}\ \mathbf{P}_\mu^{1/2}(\eta \notin B(\alpha, P_0)). \tag{3.5}$$

Так как $\mathbf{E}_\mu\,[\eta_j^4] \leqslant Cj^{-2\alpha-2}$, то мы имеем

$$\sum_{j=l_1}^{l_2} (\mathbf{E}_\mu\, \eta_j^4)^{1/2} \leqslant C\, \delta_1^{-r}\, k_\varepsilon^{-2r}. \tag{3.6}$$

Остается оценить

$$\mathbf{P}_\mu\, (\eta \notin B(\alpha, P_0)) = \mathbf{P}\, \Big(\max_{l_1 \leqslant i \leqslant l_2}\ i^{2r}\ \sum_{j=i}^{l_2} \eta_j^2 - P_0\, (1-\delta_1/2) > P_0\, \delta_1/2 \Big) \leqslant \sum_{i=l_1}^{l_2} J_i, \tag{3.7}$$

где

$$J_i = \mathbf{P}\, \Big(i^{2\alpha} \sum_{j=i}^{l_2} \eta_j^2 - P_0\, (1-\delta/2) > P_0\, \delta/2 \Big).$$

Для оценки J_i мы применим следующее утверждение [6].

Предложение 1. *Пусть* $\xi = \{\xi_i\}_{i=1}^{l}$ *– гауссовский случайный вектор, состоящий из независимых случайных величин* ξ_i, $\mathbf{E}\xi_i = 0$, $\mathbf{E}\xi_i^2 = 1$. *Пусть* A *– матрица размера* $l \times l$ *и пусть* $\Sigma = A^T A$. *Тогда*

$$\mathbf{P}\, (\|A\xi\|^2 > \mathrm{tr}\,(\Sigma) + 2\sqrt{\mathrm{tr}\,(\Sigma^2)\, t} + 2\, \|\Sigma\|\, t) \leqslant \exp\{-t\}. \tag{3.8}$$

Здесь $\mathrm{tr}\,(\Sigma)$ *обозначает след матрицы* Σ.

Определим матрицу $\Sigma = \{\sigma_{lj}\}_{l,j=i}^{l_2}$, где $\sigma_{jj} = 2\,\alpha\,(P_0 - \delta)\, j^{-2\alpha-1}\, i^{2\alpha}$ и $\sigma_{lj} = 0$, если $l \neq j$. Тогда

$$2\sqrt{(\mathrm{tr}\,\Sigma^2)\, t} + 2\, \|\Sigma\|\, t = \frac{P_0 - \delta}{\alpha\,(4\,\alpha + 1)} \sqrt{i^{-1}\, t}\, (1 + o(1)) + i^{-1}\, t \doteq V_i(t). \tag{3.9}$$

Положим $t = k_\varepsilon^{1/2}$. Тогда $V_i(t) < C\, k_\varepsilon^{-1/2}$, $l_1 \leqslant i \leqslant l_2$ и, применяя (3.8), имеем

$$J_i < \exp\{-k_\varepsilon^{-1/2}\}. \tag{3.10}$$

Следовательно,

$$\sum_{j=l_1}^{l_2} J_i \leqslant \delta_1^{-1}\, k_\varepsilon\, \exp\{-k_\varepsilon^{1/2}\}. \tag{3.11}$$

Для завершения доказательства остается оценить $R_\varepsilon - I(P_0-\delta)$.

Прямыми вычислениями легко проверить, что

$$|I(P_0) - I(P_0-\delta)| < C\,\delta\, I(P_0). \tag{3.12}$$

Имеем

$$\begin{gathered}\varepsilon^2 \sum_{j=1}^{l_1} \frac{\sigma_j^2}{1+(2\,\alpha\,P_0)^{-1}\,\varepsilon^2\,\sigma_j^2\,j^{2\alpha+1}} \asymp \varepsilon^2 \sum_{j=1}^{l_1} \sigma_j^2 \\ < C\,\delta_1\,\varepsilon^2 \sum_{j=l_1}^{k_\varepsilon} \sigma_j^2 \asymp C\,\delta_1\,\varepsilon^2 \sum_{j=l_1}^{k_\varepsilon} \frac{\sigma_j^2}{1+(2\,\alpha\,P_0)^{-1}\,\varepsilon^2\,\sigma_j^2\,j^{2\alpha+1}}\end{gathered} \tag{3.13}$$

и

$$\begin{gathered}\varepsilon^2 \sum_{j=l_2}^{\infty} \frac{\sigma_j^2}{1+(2\,\alpha\,P_0)^{-1}\varepsilon^2\,\sigma_j^2\,j^{2\alpha+1}} \asymp \varepsilon^2 \sum_{j=l_2}^{\infty} j^{-2\alpha-1} \\ \leqslant \varepsilon^2\delta_1^{2\alpha}\,C \sum_{k_\varepsilon}^{l_2} j^{-2\alpha-1} \asymp \varepsilon^2\delta_1^{2\alpha}\,C \sum_{k_\varepsilon}^{l_2} \frac{\sigma_j^2}{1+(2\,\alpha\,P_0)^{-1}\,\varepsilon^2\,\sigma_j^2\,j^{2\alpha+1}}.\end{gathered} \tag{3.14}$$

Из (3.12)–(3.14) вытекает, что $R_\varepsilon - I(P_0-\delta) \to 0$ для некоторого $\delta = \delta(\varepsilon) \to 0$ и $\delta_1 = \delta_1(\varepsilon) \to 0$ при $\varepsilon \to 0$.

3.3. Доказательство теоремы 3. Рассуждения основаны на следующей лемме.

Лемма 1. *Справедливо равенство*

$$\sup_{x \in B(\alpha,P_0)} \mathbf{E}_x\, \|\widetilde{x}_\varepsilon - x\|^2 = \mathbf{E}_{\theta_\varepsilon}\, \|\widetilde{x}_\varepsilon - \theta_\varepsilon\|^2, \tag{3.15}$$

где $\theta_\varepsilon = \{\theta_{\varepsilon k}\}_{k=1}^\infty$, $\theta_{\varepsilon k} = P_0\,(a_k - a_{k+1})$.

Доказательство леммы 1. Обозначим $u_k = a_k^{-1} \sum\limits_{j=k}^{\infty} x_j^2$. Тогда

$$x_k^2 = a_k\,u_k - a_{k+1}\,u_{k+1}.$$

Обозначим $l = [\mu^{-\frac{1}{\beta}}]$. Имеем

$$\mathbf{E}_x \|\widetilde{x}_\varepsilon - x\|^2 = \mu^2 \sum_{j=1}^{l} b_j^2 x_j^2 + \sum_{j=l+1}^{\infty} x_j^2 + \varepsilon^2 \sum_{j=1}^{l} \lambda_j^2 \doteq J_1 + J_2 + J_3 \quad (3.16)$$

соответственно. Тогда

$$\begin{aligned} J_1 + J_2 &= \mu^2 \sum_{j=1}^{l} b_j^2 (a_j u_j - a_{j+1} u_{j+1}) + a_{l+1} u_{l+1} \\ &= \mu^2 a_1 b_1^2 u_1^2 - \mu^2 a_{l+1} b_l^2 u_{l+1}^2 \\ &+ \mu^2 \sum_{j=2}^{l} a_j u_j (b_j^2 - b_{j-1}^2) + a_{l+1} u_{l+1}. \end{aligned} \quad (3.17)$$

Максимум правой части (3.17) достигается для $u_j = P_0$, $1 \leqslant j < \infty$, где $x_j^2 = P_0 (a_j - a_{j+1})$. Непосредственными вычислениями получаем $J_3 = C \varepsilon^2 l$.

Если $\beta > \alpha$, то

$$J_1 + J_2 = \frac{\beta}{\beta - \alpha} l^{-2\alpha} (1 + o(1)).$$

Если $\alpha > \beta$, то

$$J_1 + J_2 = P_0 l^{-2\beta} C_1 (1 + o(1)).$$

Если $\alpha = \beta$, то

$$J_1 + J_2 = \alpha P_0 l^{-2\alpha} \ln l.$$

Минимизируя $J_1 + J_2 + J_3$ относительно l, получаем теорему 3. □

3.4. Доказательство теоремы 4. Достаточно доказать необходимость. Имеем

$$\begin{aligned} \mathbf{E}_x \|\widetilde{x}_{\varepsilon\mu} - x\|^2 &= \varepsilon^2 \sum_{j=1}^{l} (1 - l^{-\beta} j^\beta) + l^{-2\beta} \sum_{j=1}^{l} j^{2\beta} x_j^2 + \sum_{j=l}^{\infty} x_j^2 \\ &\geqslant C \varepsilon^2 l + l^{-2\beta} \sum_{j=1}^{l} j^{2\beta} x_j^2 \doteq J_\varepsilon(l, x). \end{aligned} \quad (3.18)$$

Легко видеть, что если

$$\sum_{j=1}^{l} j^{2\beta} x_j^2 \to \infty \quad \text{при} \quad l \to \infty, \quad (3.19)$$

то
$$\lim_{\varepsilon\to 0} \varepsilon^{-\frac{4\beta}{1+2\beta}} \inf_l J_\varepsilon(l,x) = \infty.$$
□

Литература

1. K. Bertin, *Asymptotically exact minimax estimation in sup-norm for anisotropic Hölder classes.* — Bernoulli **10** (2004), 873–888.
2. L. Cavalier, *Inverse problems in statistics.* — In: Inverse problems and high-dimensional estimation. Lect. Notes Statist. **203**, pp. 3–96, Springer, 2011.
3. D. L. Donoho, *Asymptotic minimax risk for sup-norm loss: solution via optimal recovery.* — Probab. Theory Relat. Fields **99** (1994), 145–170.
4. D. L. Donoho, I. M. Johnstone, *Minimax estimation via wavelet shrinkage.* — Ann. Statist. **26** (1998), 879–921.
5. P. B. Eggermont, V. N. LaRiccia, *Maximum Penalized Likelyhood Estimation.* II. Springer, 2009.
6. D. Hsu, S. M. Kakade, T. Zang, *A tail inequality for quadratic forms of subgaussian random vector.* — Electronic Commun. Probab. **17** (2012), 1–6.
7. I. Ibragimov, R. Khasminskii, *Statistical Estimation: Asymptotic Theory.* Springer, 1981.
8. Yu. I. Ingster, T. Sapatinas, I. A. Suslina, *Minimax signal detection in ill-posed inverse problems.* — Ann. Statist. **40** (2012), 1524–1549.
9. I. M. Johnstone, *Gaussian Estimation. Sequence and Wavelet Models.* Book Draft http://statweb.stanford.edu/~imj/, 2015.
10. J. M. Loubes, V. Rivoirard, *Review of rates of convergence and regularity conditions for inverse problems.* — Intern. J. Tomography Statist. **11** (2009), 61–82.
11. G. Kerkyacharian, D. Picard, *Density estimation by kernel and wavelets methods: optimality of Besov spaces.* — Statist. Probab. Lett. **18** (1993), 327–336.
12. G. Kerkyacharian, D. Picard, *Minimax or maxisets*? — Bernoulli **8** (2002), 219–253.
13. А. П. Коростелев, *Асимптотически минимаксное оценивание регрессии в равномерной норме с точностью до константы.* — Теория вероятн. и ее примен. **38** (1993), 857–882.
14. Д. А. Кукс, В. Ольман, *Минимаксная линейная оценка коэффициентов регрессии.* — Изв. Акад. наук Эст. ССР **20** (1971), 480–482.
15. O. V. Lepski, A. B. Tsybakov, *Asymptotically exact nonparametric hypothesis testing in supnorm and at a fixed point.* — Probab. Theory Relat. Fields **117** (2000), 17–48.
16. A. S. Nemirovskii, *Nonparametric estimation of smooth regression functions.* — Soviet J. Comput. Syst. Sci. **23** (1985), 1–11.
17. M. Nussbaum, *Spline smoothing in regression models and asymptotic efficiency in* L_2. — Ann. Statist. **13** (1985), 984–997.
18. М. С. Пинскер, *Оптимальная фильтрация квадратично-интегрируемых сигналов на фоне гауссовского шума.* — Пробл. передачи информ. **16** (1980), 52–68.

19. V. Rivoirard, *Maxisets for linear procedures.* — Statist. Probab. Lett. **67** (2004), 267–275.
20. A. Tikhonov, *Regularization of incorrectly posed problems.* — Soviet Math. Dokl. **4** (1963), 1624–1627.
21. A. Tsybakov, *Introduction to Nonparametric Estimation.* Springer Series in Statistics **130**, Springer, 2009.
22. G. Wahba, *Spline Models for Observational Data.* SIAM, 1990.

Ermakov M. S. On minimax nonparametric estimation on maxisets.

For the problem of nonparametric estimation of signal in Gaussian noise we point out the strong asymptotically minimax estimators on maxisets for linear estimators. It turns out that the order of rates of convergence of Pinsker estimator on this maxisets is worse than the order of rates of convergence for the class of linear estimators considered on this maxisets. We show that balls in Sobolev spaces are maxisets for Pinsker estimators.

Институт проблем машиноведения РАН
Большой пр. В.О., 61,
Санкт-Петербург, Россия

Поступило 14 ноября 2017 г.

Санкт-Петербургский государственный Университет
Университетская наб., 7/9
Санкт-Петербург, 199034 Россия
E-mail: erm2512@gmail.com