Math-Net.Ru

Общероссийский математический портал

Б. С. Дарховский, Новый подход к стохастической задаче восстановления, *Теория вероятн. и ее примен.*, 2004, том 49, выпуск 1, 36–53

DOI: <https://doi.org/10.4213/tvp235>

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: 54.172.240.79

26 апреля 2023 г., 14:04:37

© 2004 г.

ДАРХОВСКИЙ Б. С.*

НОВЫЙ ПОДХОД К СТОХАСТИЧЕСКОЙ ЗАДАЧЕ ВОССТАНОВЛЕНИЯ

Рассматривается задача минимаксного оценивания линейных функционалов на классах функций, известная в литературе как задача восстановления. Предлагается новая формализация исходной прикладной проблемы как в детерминированной, так и в стохастической постановке. В последнем случае предложено введение естественной вероятностной меры на множестве, порождаемом информационным оператором задачи. Приведены примеры решения задач восстановления в новой постановке, установлены статистические свойства решения стохастической задачи восстановления. Рассмотрено применение предложенного подхода к задаче непараметрического минимаксного оценивания регрессионной функции.

Ключевые слова и фразы: задача восстановления, апостериорное распределение, непараметрическое минимаксное оценивание.

1. Введение. Классическая задача восстановления формулируется следующим образом (см. [1], [2] и цитированную там литературу). Пусть X, Y — линейные пространства, $\langle x^*, x \rangle$ — значение заданного линейного функционала x^* на элементе $x \in X$, $F: X \rightarrow Y$ — заданная функция, $\mathcal{M} \subset X$ — заданное множество. Оператор F называется информационным оператором. Требуется минимизировать функционал

$$\sup_{x \in \mathcal{M}, y = F(x)} |\langle x^*, x \rangle - \varphi(y)| \quad (1)$$

по множеству всех функций $\varphi: Y \rightarrow \mathbf{R}$ из некоторого класса.

Обычно (см. [2]) рассматривался оператор $F = \{\langle x_i^*, x \rangle\}_{i=1}^n$, где $\{x_i^*\}$ — заданное семейство линейных непрерывных функционалов, а множество $\mathcal{M}$ предполагалось выпуклым и центрально-симметричным.

Существует большое число публикаций по проблеме восстановления как для детерминистической постановки (1), так и для соответствующего стохастического аналога. В последнем варианте вместо $F(x)$ следует использовать (в простейшем случае) $F(x) + \xi$, где ξ — случайный

* Институт системного анализа РАН, пр-т 60-летия Октября, 9, 117312 Москва, В-312, Россия; e-mail: darbor@isa.ru

элемент со значениями в Y , и, соответственно, оператор математического ожидания перед оператором $\sup$.

Одним из первых и важных результатов для задачи (1) является теорема Смоляка, в которой установлено, что наилучшей функцией φ (наилучшим методом восстановления) среди всех непрерывных в нуле функций является линейная функция. В [3] можно найти достаточно полную информацию о современных результатах для детерминированной задачи восстановления и ее обобщений.

Для стохастической постановки задачи восстановления результат, аналогичный теореме Смоляка, неизвестен. Один из подходов, рассматриваемых в литературе (см., например, [2]) состоит в том, что на множестве $\mathcal{M}$ задается вероятностная мера и операция $\sup$ в (1) заменяется на усреднение по этой вероятностной мере. Такой подход, на наш взгляд, не вполне отвечает прикладной сути исходной задачи: обычно нет оснований полагать, что функции заданного класса суть реализации какого-то случайного процесса, более реалистично (по крайней мере, для многих приложений) считать, что эти функции просто обладают некоторыми априори известными свойствами, например, определенной гладкостью, но измеряются со случайной ошибкой. В частности, именно так обстоит дело, если заданный класс функций реализован как совокупность выходных сигналов динамической системы, возбуждаемой ограниченным по амплитуде и интегрируемым, но неизмеряемым процессом, а наблюдению доступен выходной сигнал такой системы, измеряемый с некоторой случайной погрешностью.

В математической статистике традиционно изучается важный частный случай стохастической задачи восстановления — оценивание регрессионной функции. В этой связи отметим работу [4], где для оценки гладкой вектор-функции в шуме предлагается искать (при помощи квадратичной локальной аппроксимации) кривую, доставляющую максимум апостериорной вероятности в предположении, что эта вероятность задается гауссовой плотностью. Для той же задачи в работах [5], [6] развита теория локального планирования эксперимента, позволяющая выбирать каждую новую точку измерения из соображений максимизации скорости накопления информации.

Весьма интересен результат работы [7], где найдена наилучшая линейная минимаксная оценка (по наблюдениям на всей временной оси) значения и производной в нуле регрессионной функции, допускающей представление в виде полинома заданной степени с ограниченным коэффициентом при остаточном члене. Получены необходимые и достаточные условия того, чтобы некоторая функция задавала линейную минимаксную оценку, и приведены явные формулы для этой функции.

Минимаксное оценивание функции из заданного класса по зашумленным измерениям традиционно рассматривается в статистической те-

ории оценивания, где подобного рода задачи называются задачами минимизации риска (см. [11]; в недавно вышедшей книге [12] содержится много оригинальных результатов в области непараметрического оценивания функций). Задачи минимизации риска обычно изучаются в предположении, что дисперсия шумовой составляющей стремится к нулю (кроме этого, обычно предполагается, что случайная составляющая — это белый гауссовский шум). При таком предположении в ряде случаев (см., например, [8]–[10]) удается получить оценки минимаксного риска и указать асимптотически оптимальное решение. Остается, однако, открытым вопрос о том уровне шума, начиная с которого можно пользоваться асимптотическими соотношениями.

Отметим также, что в [13] автором была предложена неасимптотическая оценка потерь по значению стохастической задачи восстановления при использовании в качестве метода восстановления оптимального метода для детерминированной задачи.

На наш взгляд, общепринятая постановка задачи восстановления как в детерминированной, так и в стохастической ситуации не вполне адекватна тем исходным прикладным проблемам, которые в свое время эти задачи породили.

Действительно, в (1) ищется «наихудшая» пара (x, y) такая, что $x \in \mathcal{M}$, $y = F(x)$. Но в реальной ситуации у нас *уже есть наблюдения* y , т.е. точка y *уже фиксирована*. Поэтому для любого заданного (или измеренного) y мы должны найти (если оставаться в рамках минимаксного подхода) действительное число $u = u(y)$ такое, что функция

$$\Phi(y, u) = \sup_{x \in \mathcal{M}, F(x)=y} |\langle x^*, x \rangle - u| \quad (2)$$

достигает своего минимума, т.е. должны решить задачу

$$\Phi(y, u) \rightarrow \inf_{u \in \mathbf{R}}. \quad (3)$$

На наш взгляд, постановка (2), (3) точно формализует исходную прикладную задачу в детерминированной ситуации. В стохастической ситуации предлагается использовать некоторую специальную конструкцию, которая будет описана ниже и идея которой была предложена в [13]. Эта конструкция в принципе позволяет находить *неасимптотически* оптимальные решения.

Организация статьи такова. В разделе 2 рассматривается новая постановка детерминированной задачи восстановления и приводятся результаты для некоторых функциональных классов. В разделе 3 предлагается новая постановка стохастической задачи восстановления, описывается подход к ее решению для произвольного конечного числа наблюдений и рассматриваются статистические свойства предлагаемого решения при специальной организации процесса наблюдений. В разделе 4

рассматривается новая процедура для задачи непараметрического оценивания регрессионной функции в шуме, основанная на предлагаемом подходе к проблеме восстановления.

2. Детерминированная задача.

2.1. Общая постановка. Пусть X, Y, Z — линейные нормированные пространства, $A: X \rightarrow Y, B: X \rightarrow Z$ — линейные операторы, $\mathfrak{M} \subset X$ — выпуклое и центрально-симметричное множество. Пусть $M = \{y \in Y: \exists x \in \mathfrak{M} \text{ такой, что } Ax = y\}$. Нетрудно видеть, что $0 \in M$ и M — выпуклое и центрально-симметричное множество.

Следующая задача будет называться *новой задачей восстановления (НЗВ)*:

$$\forall y \in M \quad \Phi(y, z) = \sup_{x \in \mathfrak{M}, Ax=y} \|Bx - z\| \longrightarrow \inf_{z \in Z}. \quad (4)$$

Положим

$$K(y) = \{z \in Z: z = Bx, x \in \mathfrak{M}, Ax = y\}.$$

Тогда задача (4) может быть переписана следующим образом:

$$\sup_{u \in K(y)} \|u - z\| \rightarrow \inf_{z \in Z}. \quad (5)$$

Из предположений следует, что множество $K(y)$ выпукло для любого $y \in M$. Поэтому для любого $u \in K(y)$

$$\|z - u\| \geq \|Pr_{K(y)} z - u\|,$$

где $Pr_H z$ — проекция точки z на множество H .

Следовательно, задачу (5) можно поставить так:

$$\sup_{u \in K(y)} \|u - z\| \rightarrow \inf_{z \in K(y)}. \quad (6)$$

Решение $z^*(y)$, $y \in M$, задачи (6) мы будем называть *оптимальным методом* для НЗВ.

Пусть $L(z, u): K \times K \rightarrow \mathbf{R}$ — заданная функция. Рассмотрим более общую минимаксную задачу:

$$\lambda(z) \triangleq \sup_{u \in K} L(z, u) \rightarrow \inf_{z \in K}. \quad (7)$$

Теорема 2.1. Пусть выполнены следующие условия:

- 1) множество K имеет центр симметрии;
- 2) функция L зависит только от разности $z - u$, симметрична относительно нуля, непрерывна и квазивыпукла.

Тогда центр симметрии множества K является решением задачи (7).

Доказательство. Будем считать, без потери общности, что $0 \in K$ и 0 является центром симметрии. Нам нужно доказать, что $0 \in \arg \min \lambda(z)$.

Напомним, что функция f называется квазивыпуклой тогда и только тогда, когда ее лебегово множество $\{x: f(x) \leq b\}$ выпукло для любого $b \in \mathbf{R}$. Легко видеть, что если f непрерывная и квазивыпуклая функция, то для любого x существует хотя бы одно полупространство $\mathfrak{R}(x)$ такое, что $f(x+h) \geq f(x)$ для любого $h \in \mathfrak{R}(x)$. Действительно, для любого x существует, в силу выпуклости и замкнутости лебеговых множеств, хотя бы одна разделяющая гиперплоскость, что и доказывает это неравенство (обратное утверждение также справедливо).

Из центральной симметрии ($-K = K$) следует:

$$\begin{aligned} \lambda(z) &= \sup_{u \in K} L(z-u) = \sup_{u \in K} L(z+u) \\ &= \sup_{u \in K} (L(u) + L(z+u) - L(u)) \geq L(u^*) + L(u^*+z) - L(u^*), \end{aligned}$$

где $u^* \in \arg \sup_{u \in K} L(u)$ (для простоты рассуждений мы полагаем, что $\arg \sup$ существует; соответствующие рассуждения нетрудно провести и в общем случае).

Если $z \in \mathfrak{R}(u^*)$, то из этого неравенства и квазивыпуклости L следует, что $\lambda(z) \geq L(u^*)$.

Допустим, что $z \notin \mathfrak{R}(u^*)$. Тогда $L(u^*+z) < L(u^*)$. Используя симметричность L , получаем

$$L(-u^*-z) = L(-(-u^*-z)) < L(u^*) = L(-u^*),$$

и, следовательно, $(-z) \notin \mathfrak{R}(-u^*)$, т.е. $z \in \mathfrak{R}(-u^*)$.

Но, в силу симметрии, $(-u^*) \in \arg \sup_{u \in K} L(u)$, откуда следует, что

$$\lambda(z) \geq L(-u^*+z) - L(-u^*) + L(-u^*) \geq L(-u^*) = L(u^*).$$

Таким образом, для любого $z \in K$ получаем $\lambda(z) \geq L(u^*) = \sup_{u \in K} L(u)$. Но, с другой стороны,

$$\inf_{z \in K} \lambda(z) \leq \lambda(0) = \sup_{u \in K} L(-u) = L(u^*).$$

Поэтому $\inf_{z \in K} \sup_{u \in K} L(z-u) = \sup_{z \in K} L(z)$, т.е. $z = 0$ является оптимальным решением.

Следствие 2.1. Если множество $K(y)$ имеет центр симметрии $u^*(y)$, то $z^*(y) = u^*(y)$ есть оптимальное решение в НЗВ.

В частности, если Bx — скаляр, а множество $K(y)$ ограничено, оптимальным решением НЗВ всегда является центр отрезка $K(y)$.

Заметим, что в общей ситуации оптимальное решение $z^*(y)$ является нелинейной функцией.

2.2. Некоторые функциональные классы. В этом пункте мы рассмотрим несколько конкретных примеров детерминированной задачи восстановления в новой постановке и опишем получающиеся решения. Рассматриваемые далее примеры представляют и самостоятельный интерес для приложений. Всюду далее мы ограничиваемся задачей восстановления для случая, когда $Z = \mathbf{R}^1$ (задача восстановления линейного функционала). В этом случае оператор B представляет собой линейный функционал, который мы будем обозначать x' .

Сформулируем сначала одно простое общее свойство оптимального решения в этой ситуации. Допустим, что множество $\mathfrak{M}$ ограничено (во всех рассматриваемых далее примерах это условие выполняется). Тогда для любого $y \in M$ множество $K(y)$ представляет собой отрезок действительной прямой и, в силу теоремы 2.1, оптимальное решение $z^*(y)$ НЗВ является центром этого отрезка. Это означает, что

$$z^*(y) = \frac{1}{2} [\phi(y) + \psi(y)], \quad (8)$$

где

$$\phi(y) = \sup_{x \in \mathfrak{M}, Ax=y} \langle x', x \rangle, \quad \psi(y) = \inf_{x \in \mathfrak{M}, Ax=y} \langle x', x \rangle.$$

Соотношение (8) можно переписать в виде

$$z^*(y) = \frac{1}{2} [\phi(y) - \phi(-y)]. \quad (9)$$

Из ограниченности множества $\mathfrak{M}$ следует, что $\phi(y)$ — собственная вогнутая функция. Но тогда из общих свойств выпуклых функций (см., например, [3]) следует, что $z^*(y)$ — непрерывная функция на $\text{int } M$. Наконец, из (9) видно, что $z^*(0) = 0$ и что $z^*(y) = -z^*(-y)$. Таким образом, имеет место следующее предложение.

Предложение 2.1. *Если множество $\mathfrak{M}$ ограничено, то оптимальное решение $z^*(y)$ в НЗВ является нечетной функцией, непрерывной на $\text{int } M$.*

Перейдем теперь к рассмотрению примеров.

а) Пусть $X = W_2^r([-\pi, \pi])$, $r \geq 1$, — соболевское пространство 2π -периодических функций, имеющих $(r-1)$ -ю абсолютно непрерывную производную и r -ю производную из $L_2([-\pi, \pi])$. Пусть $\mathfrak{M}$ — подмножество функций с нулевым средним на $[-\pi, \pi]$, удовлетворяющих неравенству $\|x^{(r)}(\cdot)\|_2^2 \leq C^2$. Оператор A ставит в соответствие каждой функции первые $2n$ коэффициентов ее ряда Фурье, а функционал x' есть значение функции в некоторой точке $t^* \in [\pi, \pi]$.

Функция $x \in \mathfrak{M}$ представляется равномерно сходящимся рядом Фурье

$$x(t) = \sum_{k=1}^{\infty} (a_k \sin kt + b_k \cos kt).$$

Положим $y_k = (a_k, b_k)$, $y = (y_1, \dots, y_n)$ и

$$\alpha(y) = \sum_{k=1}^n (a_k \sin kt^* + b_k \cos kt^*), \quad \beta(y) = \sum_{k=1}^n (a_k^2 + b_k^2) k^{2r}.$$

Тогда из неравенства Коши–Буняковского получаем

$$\begin{aligned} |x(t^*) - \alpha(y)| &= \left| \sum_{k=n+1}^{\infty} (a_k \sin kt^* + b_k \cos kt^*) \right| = \left| \sum_{k=n+1}^{\infty} A_k \cos(kt^* + \phi_k) \right| \\ &\leq \left(\frac{C^2}{\pi} - \beta(y) \right)^{1/2} \left(\sum_{k=n+1}^{\infty} (\cos^2(kt^* + \phi_k)) k^{-2r} \right)^{1/2} \triangleq H(y), \end{aligned}$$

причем из теоремы Парсеваля следует, что это неравенство точное (равенство достигается при $A_k = C(k^{-2r}/\sqrt{\pi \sum_{k=1}^{\infty} k^{-2r}})$, $\phi_k = -kt^*$, $k = 1, 2, \dots$). Поэтому множество $K(y)$ представляет собой отрезок с центром в точке $\alpha(y)$ и длиной $2H(y)$.

Таким образом, мы установили, что справедливо следующее предложение.

Предложение 2.2. *Оптимальным в НЗВ для примера а) является линейный метод $z^* = \alpha(y)$.*

З а м е ч а н и е 2.1. Очевидно, что в качестве значений оператора A не обязательно рассматривать именно первые n коэффициентов ряда Фурье. Тот же результат (с очевидными изменениями обозначений) справедлив и для набора любых n коэффициентов Фурье с фиксированными номерами.

б) Пусть $X = L_2([0, 1])$ и $\{\phi_k(t)\}$ — некоторая система непрерывных и ортонормированных функций. В качестве множества $\mathcal{M}$ рассмотрим совокупность функций, представимых равномерно сходящимися рядами Фурье по системе $\{\phi_k\}$ и удовлетворяющих неравенству $\sum_{k=1}^{\infty} m_k^2 a_k^2 \leq C^2$, где $\{m_k\}$ — числовая последовательность такая, что $\sum_{k=1}^{\infty} m_k^{-2} \phi_k^2(t) < \infty$ для любого t , $\{a_k\}$ — коэффициенты ряда Фурье. Так же, как в примере а), оператор A ставит в соответствие каждой функции первые n коэффициентов ее ряда Фурье (т.е. $y = (a_1, \dots, a_n)$), а функционал x' есть значение функции в некоторой точке $t^* \in [0, 1]$. Полагая

$$\alpha(y) = \sum_{k=1}^n a_k \phi_k(t^*), \quad \beta(y) = \sum_{k=1}^n m_k^2 a_k^2$$

и пользуясь неравенством Коши–Буняковского, получаем

$$|x(t^*) - \alpha(y)| \leq (C^2 - \beta(y))^{1/2} \left(\sum_{k=n+1}^{\infty} m_k^{-2} \phi_k^2(t^*) \right)^{1/2} \triangleq H(y),$$

причем из теоремы Парсеваля следует, что это неравенство точное (равенство достигается при выборе $a_k = C(m_k^{-2} \phi_k(t^*)/\sqrt{\sum_{k=1}^{\infty} m_k^{-2} \phi_k^2(t^*)})$).

Поэтому множество $K(y)$ представляет собой отрезок с центром в точке $\alpha(y)$ и длиной $2H(y)$.

Таким образом, мы установили, что справедливо следующее предложение.

Предложение 2.3. *Оптимальным в НЗВ для примера б) является линейный метод $z^* = \alpha(y)$.*

Как и в примере а), здесь также можно рассматривать любой набор n коэффициентов Фурье с фиксированными номерами.

З а м е ч а н и е 2.2. Из рассмотрения приведенных примеров следует, что минимаксная оценка $\hat{x}(t)$ всей функции $x(t)$ в равномерной метрике линейна и имеет вид

$$\hat{x}(t) = \langle y, R(t) \rangle,$$

где $R(t)$ — некоторая вектор-функция, $\langle \cdot, \cdot \rangle$ — символ скалярного произведения в $\mathbf{R}^n$. Так, в примере б) $R(t) = (\phi_1(t), \dots, \phi_n(t))$.

с) Пусть L — невырожденный линейный дифференциальный оператор k -го порядка с непрерывными коэффициентами. Рассмотрим на отрезке $[0, 1]$ множество функций $\mathfrak{M}$, которые являются решениями уравнения

$$Lx = u(t), \quad (10)$$

когда $u(t)$ пробегает множество всех ограниченных (почти всюду) заданной константой C и интегрируемых на $[0, 1]$ функций. Предполагается, что начальные условия для уравнения (10) заданы в форме Коши и принадлежат некоторому заданному выпуклому центрально-симметричному множеству в $\mathbf{R}^k$.

Очевидно, $\mathfrak{M}$ — выпуклое и центрально-симметричное множество в соболевском пространстве $W_1^{(k)}([0, 1]) = X$.

Пусть $\{t_1, \dots, t_n\}$ — заданный (упорядоченный по возрастанию) набор точек отрезка $[0, 1]$ и оператор A ставит в соответствие каждой функции $x(t) \in \mathfrak{M}$ набор ее значений $y_1 = x(t_1), \dots, y_n = x(t_n)$. Функционал x' задает значение функции $x(t)$ в точке τ .

Пользуясь явным выражением для решения линейного дифференциального уравнения через функцию Грина, можно показать, что построение множества допустимых значений вектора $y = (y_1, \dots, y_n)$, $y_1 = x(t_1), \dots, y_n = x(t_n)$, сводится к решению последовательности экстремальных задач ляпуновского типа (соответствующее определение см., например, в [14]) и их решение в нашем случае легко находится. При этом если точки $\tilde{y} = (y_1, \dots, y_{n-1})$ принадлежат допустимому множеству, то множество допустимых точек y_n имеет вид $\{\langle l, \tilde{y} \rangle - h \leq y_n \leq \langle l, \tilde{y} \rangle + h, h > 0\}$, где линейная функция l и число h определяются видом оператора L и набором точек $(t_1, \dots, t_n)$.

Отсюда и из теоремы 2.1 сразу следует, что оптимальным решением в рассматриваемой задаче при условии, что $\tau > t_n$, является линейная функция наблюдений y . Тот же результат верен и если $\tau < t_1$, что следует из аналогичных рассмотрений в обращенном времени. Если же точка $t_s < \tau < t_{s+1}$, $1 \leq s \leq n - 1$, то соответствующий отрезок получается пересечением отрезков, каждый из которых определяется экстремальной задачей уже рассмотренного типа при разных способах задания краевых условий через уже известные допустимые значения функции слева и справа от точки τ . При этом оптимальное решение будет кусочно-линейной функцией наблюдений (так как границы множества возможных значений оцениваемого функционала получаются в результате выбора максимальной и минимальной из соответствующих множеств аффинных функций). Итак, имеет место следующее предложение.

Предложение 2.4. *Оптимальным в НЗВ для примера с) является кусочно-линейный метод.*

3. Стохастическая задача. Для стохастической задачи мы используем формализацию, предложенную в [13], и рассмотрим статистические свойства соответствующего оптимального решения.

Далее мы будем предполагать, что оператор A в (4) имеет *конечномерный* образ (заметим, что это так для всех рассмотренных в предыдущем разделе примеров).

Для удобства чтения напомним предложенную формализацию. Пусть $\text{Im } A\mathcal{M} \equiv \Theta \subset \mathbf{R}^k$ и набор конечномерных векторов $Y = \{y_i\}$, $y_i = \theta + \xi_i$, $i = 1, \dots, n$, представляет собой (для простейшей схемы наблюдений) имеющуюся информацию после получения n наблюдений. Допустим, что набор случайных векторов $(\xi_1, \dots, \xi_n)$ имеет плотность $p(\cdot)$ по мере Лебега. Тогда на конечномерном пространстве можно ввести вероятностную меру с плотностью $q(\cdot | Y)$, $q(\theta | Y) = 0$, если $\theta \notin \Theta$, и

$$q(\theta | Y) = p(y_1 - \theta, \dots, y_n - \theta) \left[\int_{\Theta} p(y_1 - \theta, \dots, y_n - \theta) d\theta \right]^{-1}, \quad (11)$$

если $\theta \in \Theta$.

Рассмотрим теперь для каждого вектора $\theta \in \Theta$ множество на прямой $K(\theta) = \{v \in \mathbf{R}: \langle x', x \rangle = v, Ax = \theta, x(\cdot) \in \mathcal{M}\}$. В силу предположения выпуклости множество $K(\theta)$ представляет собой отрезок (быть может, бесконечный). Это совокупность всех возможных значений оцениваемого функционала при условии, что информационный образ множества $\mathcal{M}$ при отображении A попадает в точку θ .

Теперь можно рассмотреть такую игру с «природой»: «природа» с вероятностью, задаваемой плотностью $q(\cdot | Y)$, выбирает вектор $\theta \in \Theta$, а статистик должен в качестве своего решения о значении функционала $\langle x', x \rangle$ выбрать некоторое число u .

Так как при $\theta \in \Theta$ истинные значения оцениваемого функционала содержатся в множестве $K(\theta)$, то потери от принятого решения естественно оценивать некоторой функцией $h(\cdot)$ от расстояния между этим решением и точками множества $K(\theta)$, т.е. величиной $\sup_{t \in K(\theta)} h(|u - t|)$. Поскольку теперь на множестве Θ введена вероятностная мера, то получаем следующую задачу:

$$\inf_{u \in \mathbf{R}} \mathbf{E} \left(\sup_{t \in K(\theta)} h(|u - t|) \mid Y \right), \quad (12)$$

где $\mathbf{E}(\cdot \mid Y)$ — символ математического ожидания по плотности $q(\cdot \mid Y)$.

Нетрудно видеть, что если $h(t) = t$, а $K(\theta)$ есть отрезок длины $2a(\theta)$ с центром $b(\theta)$, то

$$\sup_{t \in K(\theta)} |t - u| = a(\theta) + |b(\theta) - u|. \quad (13)$$

Отсюда получаем, что решением задачи (12) является *медиана* случайной величины $b(\theta)$.

Перейдем теперь к статистическому анализу оптимального решения для рассматриваемой ситуации. Мы предполагаем, что $\{\xi_i\}$ — независимые и одинаково распределенные случайные векторы с плотностью распределения $\varphi(\cdot)$.

Мы будем предполагать, что множество Θ — компакт и для плотности выполнены стандартные условия регулярности (см. [15, с. 238–239]) в некоторой окрестности множества Θ . В целях экономии места мы не приводим здесь точной формулировки этих условий, отсылая читателя к [15], и напомним только, что, грубо говоря, условия регулярности означают достаточную гладкость плотности $\varphi(y_i - \theta)$ по параметру θ и существование соответствующей матрицы Фишера. В приложениях такие условия обычно выполнены.

В соответствии с предлагаемым подходом, следует найти «апостериорное» распределение вектора θ при условии, что имеются наблюдения $Y = (y_1, \dots, y_n)$. Плотность этого «апостериорного» распределения $q(\theta \mid Y)$ определяется (на некоторой окрестности множества Θ) формулой (11) с учетом независимости векторов ξ_i .

В силу (13) оптимальное решение стохастической задачи восстановления имеет вид $u^*(Y) = \text{Med}(b(\theta) \mid Y)$, где $\text{Med}(\cdot \mid Y)$ — «апостериорная» медиана.

Пусть $P_{\theta^*}(\mathbf{E}_{\theta^*})$ — вероятностная мера (математическое ожидание) на выборочном пространстве, соответствующая истинному (неизвестному) значению параметра $\theta^* \in \Theta$. Пусть $\hat{\theta}$ — оценка максимального правдоподобия (ОМП) параметра θ . Хорошо известно (см., например, [16], [17]), что при условиях регулярности апостериорная плотность случайного вектора $\sqrt{n}(\theta - \hat{\theta})$ равномерно на любом компакте сходится

по вероятности к плотности нормального распределения с нулевым средним и некоторой конечной матрицей ковариаций при $n \rightarrow \infty$.

Так как для произвольной случайной величины η справедливо неравенство $|\text{Med}(\eta) - \mathbf{E}\eta| \leq 2\sqrt{\mathbf{D}\eta}$, а дисперсия «апостериорного» распределения θ имеет порядок $O(1/n)$, то получаем, что

$$\text{Med}(b(\theta) | Y) = \mathbf{E}(b(\theta) | Y) + O\left(\frac{1}{\sqrt{n}}\right). \quad (14)$$

Введем функцию (информацию Кульбака–Лейблера)

$$\mathbf{K}(\theta^*, \theta) = \int \left(\ln \frac{\varphi(y - \theta^*)}{\varphi(y - \theta)} \right) \varphi(y - \theta^*) dy$$

и будем предполагать, что она дважды непрерывно дифференцируема по θ .

Далее мы будем использовать одну и ту же букву C для обозначения, вообще говоря, разных положительных констант, не зависящих от n .

Из результатов работы [18] следует, что в нашей ситуации условие дифференцируемости функции $\mathbf{K}(\cdot)$ гарантирует существование предела

$$\lim_{n \rightarrow \infty} n^{-1} \ln \left(\int_{\|\theta - \hat{\theta}\| > z} q(\theta | Y) d\theta \right) = - \min_{\|h\|=1} \{ \mathbf{K}(\theta^*, \hat{\theta} + zh) \} \quad \mathbf{P}_{\theta^*}\text{-п.н.} \quad (15)$$

(на самом деле в [18] рассматривается скалярный параметр, однако это связано с тем, что, усиливая свои прежние результаты, авторы изучают некомпактное параметрическое множество, из-за чего потребовалось использовать негладкую нормализующую функцию; в компактном случае нам достаточен прежний результат, цитируемый авторами в [18] и справедливый, как подтвердили авторы в частном сообщении, для многомерного параметра).

Известно (см. [15, с. 201]), что в условиях регулярности имеет место оценка ($\|h\| = 1$):

$$\mathbf{K}(\theta^*, \hat{\theta} + zh) \geq C \|(\hat{\theta} - \theta^*) + zh\|^2 \geq C \|\hat{\theta} - \theta^*\| - z|^2.$$

Поэтому из (15) получаем

$$\int_{\|\theta - \hat{\theta}\| > z} q(\theta | Y) d\theta \leq \exp(-Cn \|\hat{\theta} - \theta^*\| - z|^2),$$

откуда следует, что для $z > 0$

$$\int_{\Theta} \|\theta - \hat{\theta}\| q(\theta | Y) d\theta \leq z + \exp(-Cn \|\hat{\theta} - \theta^*\| - z|^2) \quad \mathbf{P}_{\theta^*}\text{-п.н.} \quad (16)$$

Выбирая в (16) z из условия $\|\hat{\theta} - \theta^*\| - z = \sqrt{(\ln n)/(Cn)}$, получаем, что

$$\int_{\Theta} \|\theta - \hat{\theta}\| q(\theta | Y) d\theta \leq C\sqrt{n^{-1} \ln n} + \|\hat{\theta} - \theta^*\| \quad \mathbf{P}_{\theta^*}\text{-п.н.} \quad (17)$$

Оценим величину $\Delta = \mathbf{E}_{\theta^*} |u^*(Y) - b(\theta^*)|$, предполагая, что функция $b(\theta)$ удовлетворяет условию Липшица на компакте Θ . Тогда получим

$$\begin{aligned} |\mathbf{E}\{b(\theta) - b(\theta^*) \mid Y\}| &\leq C \mathbf{E}\{\|\theta - \theta^*\| \mid Y\} \\ &\leq C (\mathbf{E}\{\|\theta - \hat{\theta}\| \mid Y\} + \mathbf{E}\{\|\hat{\theta} - \theta^*\| \mid Y\}). \end{aligned} \quad (18)$$

Усредняя левую и правую части (18) по мере $\mathbf{P}_{\theta^*}$ и учитывая (17), получим

$$\Delta \leq C (\sqrt{n^{-1} \ln n} + \mathbf{E}_{\theta^*} \|\hat{\theta} - \theta^*\|). \quad (19)$$

Из свойств ОМП (см. [15, с. 243]) следует, что второе слагаемое в (19) допускает оценку сверху величиной $O(1/\sqrt{n})$. Поэтому окончательно получаем $\Delta \leq O(\sqrt{(\ln n)/n})$, что с учетом (14) дает

$$\mathbf{E}_{\theta^*} |u^*(Y) - b(\theta^*)| \leq O(\sqrt{n^{-1} \ln n}). \quad (20)$$

Таким образом, мы установили, что справедлива следующая теорема.

Теорема 3.1. Пусть для стохастической задачи восстановления выполнены следующие условия:

- 1) множество $\Theta = \{\theta \in \mathbf{R}^k: \theta = Ax, x(\cdot) \in \mathfrak{M}\}$ — компакт;
- 2) в принятой схеме наблюдений случайные векторы $\{\xi_i\}$ независимы, одинаково распределены и имеют плотность (по мере Лебега), которая удовлетворяет стандартным условиям регулярности;
- 3) информация Кульбака–Лейблера $\mathbf{K}(\theta^*, \theta)$ дважды непрерывно дифференцируема по θ для любых $\theta^*, \theta \in \Theta$;
- 4) оптимальное решение детерминированной НЗВ удовлетворяет условию Липшица на множестве Θ .

Тогда

а) для любого n (неасимптотическое) оптимальное решение $u^*(Y)$ стохастической задачи восстановления имеет вид

$$u^*(Y) = \text{Med}(b(\theta) \mid Y),$$

где $b(\theta)$ — оптимальное решение НЗВ при безошибочных наблюдениях;

б) для любого $\theta^* \in \Theta$ оптимальное решение $u^*(Y)$ сходится $\mathbf{P}_{\theta^*}$ -п.н. к оптимальному решению детерминированной НЗВ $b(\theta^*)$ при $n \rightarrow \infty$;

в) имеет место оценка

$$\sup_{\theta^* \in \Theta} \mathbf{E}_{\theta^*} |u^*(Y) - b(\theta^*)| \leq O(\sqrt{n^{-1} \ln n}).$$

4. Применение к задаче непараметрического оценивания регрессионной функции. В этом разделе мы рассмотрим возможные применения развитого выше подхода к интересной и важной для приложений задаче непараметрического оценивания функции регрессии. Мы не будем здесь формулировать точные утверждения, а ограничимся иллюстративным примером, который дает представление о новых возможностях.

Рассмотрим задачу оценки регрессионной функции для стандартной схемы наблюдений:

$$y(t_i) = x(t_i) + \xi_i, \quad i = 1, \dots, n, \quad (21)$$

где $\{y(t_i)\}$ — набор наблюдений, $x(t)$, $t \in [0, 1]$, — неизвестная функция регрессии, $\{\xi_i\}$ — последовательность независимых и одинаково распределенных случайных величин с нулевым математическим ожиданием, плотностью (по мере Лебега) и дисперсией σ^2 , $\{t_i\}$ — разбиение отрезка $[0, 1]$ (без ограничения общности далее считаем, что это разбиение равномерно и имеет шаг $1/n$).

Мы предполагаем, что функция $x(t)$ принадлежит некоторому заданному классу $\mathcal{M}$, который будет описан ниже, и требуется построить ее оценку по наблюдениям (21).

Начнем с детерминированной задачи (шум в наблюдениях отсутствует, $\sigma = 0$). Пусть $\{\psi_k(t)\}$, $k = 1, 2, \dots$, — некоторая совокупность непрерывно дифференцируемых функций, образующих в $L_2([0, 1])$ ортонормированную систему. В качестве множества $\mathcal{M}$ рассмотрим совокупность функций, представимых равномерно сходящимися рядами Фурье по системе $\{\psi_k\}$ и таких, что:

1)

$$\sup_{x \in \mathcal{M}} \sum_{k=s+1}^{\infty} a_k^2 \triangleq F(s) \rightarrow 0 \quad \text{при} \quad s \rightarrow \infty,$$

где $\{a_k\}$ — коэффициенты Фурье функции $x(t)$ (равномерная конечномерная аппроксимируемость);

2) для любого s множество, пробегаемое первыми s коэффициентами ряда Фурье функций из $\mathcal{M}$, является компактом в $\mathbf{R}^s$.

Кроме того, будем считать, что для множества $\mathcal{M}$ известно оптимальное решение детерминированной НЗВ значения функции $x(\cdot)$ в заданной точке по заданным нескольким ее коэффициентам Фурье.

Все эти условия выполняются, в частности, для множества $\mathcal{M}$ из примера б), который далее мы и будем иметь в виду.

Допустим, что заданы первые s коэффициентов Фурье неизвестной функции $x(t) \in \mathcal{M}$. Тогда (для примера б)) минимаксная оценка значения неизвестной функции $x(\cdot)$ в произвольной точке t есть решение НЗВ

и имеет вид

$$\chi(t) = \sum_{i=1}^s a_i \psi_i(t), \quad (22)$$

где $\{a_i\}$ — коэффициенты Фурье.

Построим по наблюдениям (21) (т.е., в детерминированном варианте, по значениям неизвестной функции в точках $\{t_i = i/n\}$, $i = 0, \dots, n-1$) оценки коэффициентов Фурье, пользуясь формулами численного интегрирования. Тогда получим

$$\hat{a}_k = n^{-1} \sum_{i=1}^n x\left(\frac{i}{n}\right) \psi_k\left(\frac{i}{n}\right) \quad (23)$$

(мы здесь пользуемся, для простоты, формулой интегрирования по правилу прямоугольников, что несущественно с точки зрения предлагаемой идеи).

Используя характеристики заданной системы $\{\psi_k\}$ и компакта $\mathfrak{M}$, можно вычислить величину

$$\sup_{x(\cdot) \in \mathfrak{M}} |a_k - \hat{a}_k| = n^{-1} \rho(k), \quad (24)$$

где положительная функция $\rho(k)$ вычисляется через константы Липшица функции и ее производной.

Нетрудно понять, что минимаксное решение при заданных неточных коэффициентах Фурье $\{\hat{a}_k\}$ имеет вид

$$\hat{x}(t) = \sum_{i=1}^s \hat{a}_i \psi_i(t).$$

Из (22), (23) и (24) получаем

$$\sup_{x(\cdot) \in \mathfrak{M}} \|\chi(t) - \hat{x}(t)\|_2^2 = n^{-2} \sum_{k=1}^s \rho^2(k) = n^{-2} R(s), \quad R(s) \rightarrow \infty \quad (25)$$

при $s \rightarrow \infty$.

В силу предположения 1)

$$\sup_{x(\cdot) \in \mathfrak{M}} \|x(t) - \chi(t)\|_2^2 = \sup_{x(\cdot) \in \mathfrak{M}} \sum_{k=s+1}^{\infty} a_k^2 = F(s) \rightarrow 0 \quad \text{при } s \rightarrow \infty. \quad (26)$$

Поэтому

$$\sup_{x(\cdot) \in \mathfrak{M}} \|x(t) - \hat{x}(t)\|_2^2 = F(s) + n^{-2} R(s) = H(n, s). \quad (27)$$

Из (27), учитывая (25) и (26), заключаем, что для любого n (начиная с некоторого, достаточного для вычисления первого коэффициента) можно указать *оптимальное число* $s^* = \varphi(n)$ коэффициентов

Фурье (равное, очевидно, единственной точке минимума по s функции $H(n, s)$), которое следует вычислять по наблюдениям вида (21) для получения $\epsilon^*(n)$ -минимаксной оценки неизвестной функции регрессии, где $(\epsilon^*)^2(n) = H(n, s^*)$. Аналогичные рассуждения можно провести и в общем случае, если для множества $\mathcal{M}$ выполнены предположения 1) и 2) и известно оптимальное решение детерминированной НЗВ через коэффициенты Фурье.

Перейдем теперь к стохастической задаче оценки функции регрессии при тех же предположениях относительно этой функции, что и выше. Фиксируем число s коэффициентов ряда Фурье и разобьем наблюдения на $p = \min(\varphi^{-1}(s), n)$ групп по $m = n/p$ наблюдений в каждой (без ограничения общности будем считать, что число m целое). Положим $\delta = m/n$ и рассмотрим совокупность независимых между собой и с $\{\xi_i\}_{i=1}^n$ случайных величин ν_j , $j = 1, \dots, m$, имеющих равномерное распределение на $[0, m]$. Положим $([\cdot])$ — символ целой части числа)

$$\tilde{y}(i, j) = (1 + \nu_j - [\nu_j]) y\left(\frac{[\nu_j] + im}{n}\right) + ([\nu_j] - \nu_j) y\left(\frac{[\nu_j] + 1 + im}{n}\right),$$

$$\tilde{x}(i, j) = (1 + \nu_j - [\nu_j]) x\left(\frac{[\nu_j] + im}{n}\right) + ([\nu_j] - \nu_j) x\left(\frac{[\nu_j] + 1 + im}{n}\right),$$

$$\tilde{\xi}(i, j) = (1 + \nu_j - [\nu_j]) \xi_{([\nu_j] + im)/n} + ([\nu_j] - \nu_j) \xi_{([\nu_j] + 1 + im)/n},$$

$$i = 0, \dots, p-1, \quad j = 1, \dots, m,$$

$$w_{kj} = p^{-1} \sum_{i=0}^{p-1} \tilde{y}(i, j) \psi_k\left((2i+1) \frac{\delta}{2}\right), \quad k = 1, \dots, s,$$

$$w_j = (w_{1j}, \dots, w_{sj}), \quad j = 1, \dots, m,$$

$$\eta_{kj} = p^{-1} \sum_{i=0}^{p-1} \tilde{\xi}(i, j) \psi_k\left((2i+1) \frac{\delta}{2}\right), \quad k = 1, \dots, s,$$

$$\eta_j = (\eta_{1j}, \dots, \eta_{sj}), \quad j = 1, \dots, m,$$

$$\theta_k = a_k, \quad \theta = (\theta_1, \dots, \theta_s),$$

$$\hat{\theta}_k = p^{-1} \sum_{i=0}^{p-1} \mathbf{E} \tilde{x}(i, j) \psi_k\left((2i+1) \frac{\delta}{2}\right), \quad \hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_s),$$

где, напомним, a_k — коэффициент Фурье. Тогда в нашем распоряжении оказывается набор новых векторов наблюдений w_j , $j = 1, \dots, m$, — набор оценок первых s коэффициентов Фурье, имеющий вид

$$w_j = \hat{\theta} + \eta_j + \varepsilon_j, \quad j = 1, \dots, m, \quad (28)$$

где, по построению, $(\{\eta_j\}, \{\varepsilon_j\})$ — независимые случайные векторы с нулевым математическим ожиданием. Векторы η_j одинаково распределены, и их матрица ковариаций имеет вид $\sigma^2 p^{-1} E(1 + o(1))$ (при

$p \rightarrow \infty$). Распределение векторов ε_j тоже одинаково и является равномерным, причем, в силу (24) и (25), $\|\varepsilon_j\| \leq 2p^{-1}\sqrt{R(s)}$.

Истинный параметр θ , т.е. набор точных коэффициентов Фурье, пробегает некоторое компактное множество Θ в $\mathbf{R}^s$, определяемое по заданному множеству $\mathfrak{M}$. При этом, в силу (24) и (25),

$$\sup_{x(\cdot) \in \mathfrak{M}} \|\theta - \hat{\theta}\|_2^2 = p^{-2} \sum_{k=1}^s \rho^2(k) = p^{-2} R(s). \quad (29)$$

Таким образом, задача оценивания регрессионной функции (в предположении, что по заданным наблюдениям вычисляются первые s коэффициентов Фурье) теперь может быть поставлена как стохастическая задача восстановления с наблюдениями (28) и связью между точными коэффициентами Фурье и их детерминированными оценками по дискретным наблюдениям (21), даваемой формулой (29), причем мы находимся в условиях применимости теоремы 3.1.

В соответствии с общим подходом, для решения этой задачи следует построить на множестве $\hat{\Theta}$, которое представляет собой $p^{-1}\sqrt{R(s)}$ -окрестность множества Θ , «апостериорное» распределение для нового параметра $\hat{\theta}$ при условии заданных новых наблюдений w , и в качестве оптимальной оценки взять «апостериорное» среднее детерминированного оптимального решения (на самом деле следует брать «апостериорную» медиану, однако ее отличие от среднего имеет более высокий порядок малости, чем отличие решения от детерминированного). В частности, при справедливости (22) (например, для примера б)) это решение — минимаксная оценка функции регрессии $\tilde{x}(t)$ при вычислении первых s коэффициентов Фурье — имеет вид

$$\tilde{x}(t) = \sum_{k=1}^s \tilde{a}_k \psi_k(t), \quad (30)$$

где $\tilde{a}_k = \mathbf{E}\{\hat{\theta}_k | (w_1, \dots, w_m)\}$ — «апостериорное» среднее для нового параметра.

Теперь следует выбрать оптимальное количество коэффициентов Фурье s . Для этого воспользуемся соотношением (29) и теоремой 3.1, учитывая, что в нашем случае (для наблюдений вида (28)) оценка (20) допускает уточнение $\Delta \leq C\sigma\sqrt{(s/p)(\ln m)/m} = C\sigma\sqrt{(s/n)\ln(n/p)}$, где константа C может быть оценена достаточно точно. Тогда получим (см. теорему 3.1 и (27)), что

$$\sup_{\theta \in \Theta} \mathbf{E}_{\theta^*} \|x(t) - \tilde{x}(t)\|_2 \leq C\sigma \sqrt{\frac{s}{n} \ln \frac{n}{p}} + \sqrt{H(p, s)}. \quad (31)$$

Из соотношения (31) минимизацией по s выбираем оптимальное число $s^{**} < s^*$ коэффициентов Фурье для любого (достаточно большого) числа наблюдений n . В результате получаем $\epsilon^{**}(n)$ -минимаксную

оценку функции регрессии, где $(\epsilon^{**})(n) = \{C\sigma\sqrt{(s^{**}/n)\ln(n/p(s^{**}))} + \sqrt{H(p(s^{**}), s^{**})}\}$.

З а м е ч а н и е 4.1. Все проведенные рассуждения практически дословно переносятся на случай, когда аргумент регрессионной функции является конечномерным вектором.

З а м е ч а н и е 4.2. Вышеизложенный подход можно применить также к непараметрическому оцениванию регрессионной функции, обладающей известными свойствами гладкости. В этом случае для каждого n надо сначала найти оптимальную степень $m(n) < n$ полинома наилучшего приближения (по подходящей системе функций). В детерминированном варианте $m(n)$ зависит лишь от класса функций и метода вычисления коэффициентов, в стохастической задаче следует воспользоваться соотношением типа (31). После этого задача сведется к проблеме восстановления вектора коэффициентов полинома наилучшего приближения, т.е. к вычислению соответствующей «апостериорной» медианы. Соответствующая (приближенная) минимаксная оценка регрессионной функции будет получена через статистически восстановленный полином наилучшего приближения.

З а м е ч а н и е 4.3. После того, как эта статья была написана, автору стала известна весьма интересная работа [19]. В этой работе обстоятельно анализируются методы непараметрического минимаксного оценивания регрессионной функции и предлагается новая идея построения такой оценки. Она состоит в том, что по наблюдениям вида (21) строятся эмпирические оценки коэффициентов wavelet-преобразования неизвестной функции, а затем те коэффициенты, которые по модулю меньше, чем $\sigma\sqrt{2(\ln n)/n}$, отбрасываются и оценка представляется в виде обратного wavelet-преобразования по оставшимся коэффициентам. Доказывается, что такая оценка асимптотически минимаксна (в общепринятом смысле). Таким образом, в [19] предлагается оставлять только «статистически значимые» коэффициенты wavelet-разложения, которые существенно превосходят уровень шума. Это хорошо согласуется с нашей идеей, состоящей в подсчете оптимального количества эмпирических коэффициентов Фурье (отметим, что среднеквадратичное отклонение эмпирических коэффициентов Фурье от их детерминированных значений имеет тот же порядок — $\sigma\sqrt{(\ln n)/n}$). Однако авторы [19] не учитывают погрешность, связанную с дискретностью наблюдений, что с необходимостью должно приводить к коррекции выбора оптимального числа коэффициентов wavelet-разложения. Отметим также, что для детерминированной задачи оптимальное число коэффициентов Фурье s^* в нашем методе может быть определено точно (оно зависит только от способа численного интегрирования и системы функций $\{\psi_k\}$). При этом для любого заранее заданного количества коэффициентов Фурье решение (30) является, в нашей постановке задачи восстановления, неасим-

птотически минимаксным, а асимптотические соображения используются только при подсчете оптимального числа коэффициентов s^{**} . Было бы очень интересно провести численное сравнение нашего метода и метода [19] (теоретически, по-видимому, оба метода дают сходный порядок сходимости), что автор надеется предпринять в будущем.

Автор глубоко благодарен В. М. Тихомирову за внимание к работе и рецензенту, замечания которого позволили существенно улучшить изложение.

СПИСОК ЛИТЕРАТУРЫ

1. Трауб Дж., Вожьяняковский Х. Общая теория оптимальных алгоритмов. М.: Мир, 1983, 382 с.
2. Micchelli C. A., Rivlin T. J. A survey of optimal recovery. — Optimal Estimation in Approximation Theory. Ed. by C. A. Micchelli and T. J. Rivlin. New York: Plenum, 1977, p. 1–54.
3. Магарил-Ильяев Г. Г., Тихомиров В. М. Выпуклый анализ и его приложения. М.: Эдиториал УРСС, 2000, 174 с.
4. Клепиков Н. П., Соколов С. Н. Нелинейный конфлюентный анализ. — Теория вероятн. и ее примен., 1957, т. 2, в. 4, с. 472–475.
5. Соколов С. Н. Непрерывное планирование регрессионных экспериментов. I. — Теория вероятн. и ее примен., 1963, т. 8, в. 1, с. 95–101.
6. Соколов С. Н. Непрерывное планирование регрессионных экспериментов. II. — Теория вероятн. и ее примен., 1963, т. 8, в. 3, с. 318–324.
7. Легостаева И. Л., Ширяев А. Н. Минимаксные веса в задаче выделения тренда случайного процесса. — Теория вероятн. и ее примен., 1971, т. 16, в. 2, с. 339–345.
8. Lepski O. V., Spokoiny V. G. Optimal pointwise adaptive methods in nonparametric estimation. — Ann. Statist., 1997, v. 25, № 6, p. 2512–2546.
9. Коростелев А. П. Асимптотически минимаксная оценка регрессии в равномерной норме с точностью до постоянной. — Теория вероятн. и ее примен., 1993, т. 38, в. 4, с. 875–882.
10. Leonov S. L. On the solution of an optimal recovery problem and its application in nonparametric regression. — Math. Methods Statist., 1997, v. 6, № 4, p. 476–490.
11. Ибрагимов И. А., Хасьминский Р. З. Асимптотическая теория оценивания. М.: Наука, 1979, 527 с.
12. Efromovich S. Nonparametric Curve Estimation. Methods, Theory, and Applications. New York: Springer-Verlag, 1999, 411 p.
13. Дарховский Б. С. О стохастической задаче восстановления. — Теория вероятн. и ее примен., 1998, т. 43, в. 2, с. 357–364.
14. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. М.: Наука, 1974, 480 с.
15. Боровков А. А. Математическая статистика. М.: Наука, 1984, 472 с.
16. Walker A. M. On the asymptotic behavior of posterior distributions. — J. Roy. Statist. Soc. Ser. B, 1969, v. 31, p. 80–88.
17. Schervish M. J. Theory of Statistics. New York: Springer-Verlag, 1995, 702 p.
18. Fu J. C., Kass R. E. The exponential rates of convergence of posterior distributions. — Ann. of Inst. Statist. Math., 1988, v. 40, № 4, p. 683–691.
19. Donoho D. L., Johnstone I. M., Kerkycharian G., Picard D. Wavelet shrinkage: asymptopia?—J. Roy. Statist. Soc. Ser. B, 1995, v. 57, № 2, p. 301–369.

Поступила в редакцию
10.XII.2001

Исправленный вариант
30.VII.2002