Math-Net.Ru

Общероссийский математический портал

Л. А. Саханенко, Нижние границы для точности оценивания в методах исследования тензоров диффузии, *Теория вероятн. и ее примен.*, 2009, том 54, выпуск 1, 170–180

DOI: <https://doi.org/10.4213/tvp2553>

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: 44.207.124.84

26 апреля 2023 г., 10:49:35

© 2009 г.

САХАНЕНКО Л. А.*

**НИЖНИЕ ГРАНИЦЫ ДЛЯ ТОЧНОСТИ ОЦЕНИВАНИЯ
В МЕТОДАХ ИССЛЕДОВАНИЯ ТЕНЗОРОВ ДИФФУЗИИ¹⁾**

Векторное поле с аддитивным шумом наблюдается в случайных точках. Соответствующая интегральная кривая должна быть оценена по наблюдениям. Главная цель данной статьи состоит в получении нижних границ для точности статистического оценивания функций от разности между самой интегральной кривой и ее оценкой. В частности, показано, что способы оценивания в работе [4] дают оценки, которые имеют оптимальную скорость сходимости в минимаксном смысле.

С точки зрения приложений, эта работа мотивирована задачей об исследовании тензоров диффузии, которая лежит в основе современной техники получения изображений мозга, где интегральные кривые используются для моделирования аксонных волокон в мозге. В медицинской исследовательской работе важно уметь отслеживать такие волокна и оценивать их месторасположение. В данной статье рассматриваются статистические аспекты, непосредственно относящиеся к этой задаче оценивания.

Ключевые слова и фразы: локальная асимптотическая нормальность, оптимальная скорость сходимости, исследование тензоров диффузии.

1. Введение. Пусть G — ограниченное открытое выпуклое подмножество в $\mathbf{R}^d$. Предположим, что неизвестное векторное поле $v: \mathbf{R}^d \rightarrow \mathbf{R}^d$ наблюдается в случайных точках $X_i \in G$, $i = 1, \dots, n$, со случайными ошибками $\xi_1, \dots, \xi_n$, т.е. наблюдения (X_i, V_i) представимы в виде

$$(X_i, V_i) = (X_i, v(X_i) + \xi_i). \quad (1.1)$$

Для векторного поля v определим интегральную кривую $x(t)$, порождаемую им, с началом в точке $a \in G$, как решение дифференциального уравнения

$$\frac{dx(t)}{dt} = v(x(t)), \quad t \geq 0, \quad x(0) = a \in G. \quad (1.2)$$

Эта кривая $x(t)$ существует, единственна и находится целиком в G , если v — достаточно гладкое векторное поле, обращающееся в нуль вне G (см., например, [2]).

Задача оценивания $x(t)$, $t > 0$, по данным (1.1) естественным образом возникает в так называемых методах *исследования тензоров диффузии (diffusion tensor imaging)*. Эта популярная техника сканирования живых тканей сочетает магнитно-резонансную томографию с измерениями диффузии молекул воды. В этой области науки микроструктуры в мягких тканях моделируются интегральными кривыми $x(t)$, $t > 0$. Более подробное описание приведено в разделе 3.

* Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824-1027; e-mail: luda@stt.msu.edu

¹⁾ Настоящая работа была выполнена при частичной поддержке гранта NSF DMS-0806176.

Недавно в [4] был предложен класс способов оценивания кривых $x(t)$, $t > 0$, по данным (1.1). Процедура работы [4] состоит из двух шагов. Во-первых, строится ядерная оценка для функции $v(x)$ в духе Парзена и Розенблатта

$$\hat{V}_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) V_i,$$

где K — это ядро, а h_n — ширина окна. Во-вторых, по принципу подстановки, оценка $\hat{X}_n(t)$ для $x(t)$ определяется как решение (1.2) в случае, когда v заменено на $\hat{V}_n$.

В [4] были сделаны следующие предположения:

(A1) последовательности $\{X_i: i = 1, 2, \dots\}$ и $\{\xi_i: i = 1, 2, \dots\}$ независимы;

(A2) точки наблюдений $\{X_i: i = 1, 2, \dots\}$ независимы и равномерно распределены в G ;

(A3) погрешности наблюдений $\{\xi_i: i = 1, 2, \dots\}$ независимы и одинаково распределены с нулевым средним, ковариационной матрицей Σ и конечным четвертым моментом $\mathbf{E}|\xi_i|^4 < \infty$;

(A4) G — открытое ограниченное множество с единичной мерой Лебега, которое включает в себя носитель дважды непрерывно дифференцируемого векторного поля v ;

(A5) для любого фиксированного $T > 0$ существует такое $\gamma > 0$, что $|(t - s)^{-1} \int_s^t v(x(\lambda)) d\lambda| \geq \gamma$ для всех $0 \leq s < t \leq T$;

(A6) ширина окна h_n удовлетворяет условию $nh_n^{d+3} \rightarrow \beta$ для некоторого конечного числа β при $n \rightarrow \infty$;

(A7) ядро K дважды непрерывно дифференцируемо и удовлетворяет условиям

$$\int_{\mathbf{R}^d} K(x) dx = 1, \quad \int_{\mathbf{R}^d} K(x)x dx = 0, \quad \int_{\mathbf{R}^d} |K(x)| |x|^2 dx < \infty.$$

При этих сделанных предположениях в [4] показано, что последовательность случайных процессов $n^{2/(d+3)}(\hat{X}_n(t) - x(t))$ при $n \rightarrow \infty$ слабо сходится в пространстве $\mathbf{R}^d$ -значных непрерывных функций на $[0, T]$:

$$n^{2/(d+3)}(\hat{X}_n(t) - x(t)) \xrightarrow{d} \nu(t), \quad (1.3)$$

где $\nu(t)$ — невырожденный гауссовский процесс, вид которого зависит от v , K , Σ , a и β .

Этот результат об асимптотической нормальности указывает, в частности, на существование оценок интегральной кривой, имеющих скорость сходимости $n^{2/(d+3)}$ в каждой точке. Из нашей первой нижней границы (теорема 1) вытекает, что метод работы [4] дает оценки со скоростью сходимости, оптимальной в том смысле, что не существует оценок интегральной кривой, сходящихся еще быстрее к этой интегральной кривой в фиксированной точке.

Далее, пусть Γ обозначает замкнутое подмножество в G , и пусть $d(x, y)$ — это какое-то расстояние между x и y в $\mathbf{R}^d$. Положим $d(x, \Gamma) = \inf_{y \in \Gamma} d(x, y)$. Естественно рассматривать эту величину как расстояние от точки x до множества Γ . Пусть также m — некоторая строго возрастающая функция на $\mathbf{R}_+$ (например, $m(u) = u^2$ или $m(u) = u$, $u > 0$). Обозначим $\varphi(x) = m(d(x, \Gamma))$. В этом случае в [4] установлена слабая сходимость последовательности случайных величин

$$n^r \left[\min_{t \in [0, T]} \varphi(\hat{X}_n(t)) - \min_{t \in [0, T]} \varphi(x(t)) \right]$$

к некоторой невырожденной случайной величине, которая зависит от v , K , β , m , d и Γ . Наша вторая нижняя граница (теорема 2) показывает, что метод работы [4] дает оценки, которые достигают оптимальной скорости сходимости n^r при $r = 2/(d+3)$ в случае, когда оценивается наименьшее расстояние между интегральной кривой и множеством. Отметим, что в [4] также установлена лучшая скорость сходимости n^r

при $r = 4/(d+3)$ для случая, когда интегральная кривая $x(t)$, $t > 0$, приближается к множеству Γ определенным образом, отличным от условий, налагаемых в нашей теореме 2. Это ни в коей мере не противоречит нашему результату.

Оставшаяся часть статьи организована следующим образом. Основные результаты приводятся в разделе 2. Затем следуют мотивация в разделе 3 и доказательства в разделе 4.

2. Основные результаты. Введем систему обозначений. Для любого вектора $x \in \mathbf{R}^d$ через x_l будет обозначаться его l -я координата, $l = 1, \dots, d$. Будем использовать символ $|\cdot|$ для евклидовой нормы вектора или матрицы, а также для модуля числа. Далее, введем скалярное произведение $\langle x, y \rangle = \sum_{i=1}^d x_i y_i$ для $x, y \in \mathbf{R}^d$. Для функции $f: \mathbf{R}^d \rightarrow \mathbf{R}$ символ ∇f обозначает ее градиент, который представляет собой векторное поле, а $\nabla^2 f$ — это тензорное (матричное) поле частных производных второго порядка от функции f . Наконец, $\mathcal{L}_2(A)$ обозначает класс функций, интегрируемых в квадрате по мере Лебега на множестве A .

Следующие предположения о распределении случайных величин $\{(X_i, \xi_i)\}$ будут использоваться на протяжении всей статьи:

(F0) выполнены условия (A1) и (A2);

(F1) погрешности наблюдений $\{\xi_i: i = 1, 2, \dots\}$ независимы и одинаково распределены с общей плотностью $f: \mathbf{R}^d \rightarrow \mathbf{R}_+$;

(F2) функция $\sqrt{f}$ дважды дифференцируема всюду в $\mathbf{R}^d$;

(F3) $\int |\nabla^2 \sqrt{f}|^2(y) dy < \infty$ и $\int f |\nabla \ln f|^4(y) dy < \infty$.

Поскольку f является плотностью, условие (F2) гарантирует существование открытого множества, которое назовем S , где $\nabla \sqrt{f}$ не равняется нулю. Из условия (F3) и того факта, что f — плотность, пользуясь неравенством Гёльдера, заключаем, что $\int f |\nabla \ln f|^\alpha(y) dy < \infty$ для любого $\alpha \in [0, 4]$.

Мы будем использовать следующие классы. Для фиксированного $T > 0$ через $\mathcal{C}^2(a, G, T)$ обозначается класс векторных полей, которые одновременно удовлетворяют условиям (A4), (A5), равны нулю вне G и не обращаются в нуль в точке a . Пусть $\mathcal{W}$ — класс всех нетривиальных (т.е. не равных тождественно нулю) четных функций $w: \mathbf{R} \rightarrow \mathbf{R}$, которые не убывают на $\mathbf{R}_+$ и равны нулю в нуле. Примерами таких функций служат u^2 , $|u|$, $I(|u| > c)$. Через $\mathcal{E}_n(T)$ обозначим класс всех возможных оценок интегральной кривой $x(t)$, $t \in [0, T]$, из (1.2), построенных по наблюдениям (1.1).

Теорема 1. Пусть $T > 0$, и пусть выполнены условия (F0)–(F3). Тогда для любой точки $a \in G$, любого $t_0 \in (0, T]$, любой функции $w \in \mathcal{W}$ и любого индекса $l = 1, \dots, d$ имеем

$$\liminf_{n \rightarrow \infty} \inf_{\hat{X}_n \in \mathcal{E}_n(T)} \sup_{v \in \mathcal{C}^2(a, G, T)} \mathbf{E} w\left(n^{2/(d+3)} |\hat{X}_{n,l}(t_0) - x_l(t_0)|\right) > 0.$$

Для фиксированного $T > 0$ пусть $\mathcal{M}_n(T)$ обозначает класс всех возможных оценок величины $\inf_{t \in [0, T]} \varphi(x(t))$, построенных по данным (1.1).

Теорема 2. Предположим, что выполнены условия (F0)–(F3). Тогда для любого замкнутого подмножества $\Gamma \subset G$ и любой точки $a \in G \setminus \Gamma$, в окрестности которой функция φ непрерывно дифференцируема и $\nabla \varphi(a) \neq 0$, а также для любой функции $w \in \mathcal{W}$ и любого $T > 0$ имеем

$$\liminf_{n \rightarrow \infty} \inf_{\hat{M}_n \in \mathcal{M}_n(T)} \sup_{v \in \mathcal{C}^2(a, G, T)} \mathbf{E} w\left(n^{2/(d+3)} \left| \hat{M}_n - \inf_{t \in [0, T]} \varphi(x(t)) \right| \right) > 0.$$

3. Мотивация. Настоящая работа мотивирована так называемыми методами исследования тензоров диффузии. Эта популярная техника сканирования живых тканей комбинирует магнитно-резонансную томографию с измерениями диффузии молекул воды. Этот неинвазивный метод позволяет изучать микроструктуры в живых органах. Например, он широко применяется в сканировании мозга. Кроме того, у этого метода есть потенциальное применение в области изучения других мягких тканей, таких, как мышцы, связки и сухожилия.

Ниже схематично обрисуем основы методов исследования тензоров диффузии. Под воздействием магнитного поля протон воды ведет себя как гироскоп. Движению многих миллионов таких протонов могут препятствовать волокна в живых тканях. Используя магнитные поля с разными свойствами, соответствующая диффузия воды, характеризуемая 3×3 симметричным положительно определенным тензором диффузии, может быть оценена в любой точке x исходя из затуханий интенсивности сигналов, замеренных в окрестности этой точки x . Компоненты тензора представляют собой степень убывания сигналов для пульсирующего векторного градиента с направлениями вдоль координатных осей. В белом веществе мозга специфическое строение пучков миелинированных аксонных волокон, по-видимому, отвечает за эту анизотропию диффузии (см. [5]). Диффузия воды в большей мере изотропна в сером веществе мозга.

С математической точки зрения, собственные векторы $v(x)$ тензора диффузии, соответствующие максимальным собственным числам, указывают направления аксонных волокон. Поэтому аксонные структуры (в белом веществе) локально можно моделировать с помощью интегральных кривых $x(t)$, $t > 0$, соответствующих собственным векторам $v(x)$, т.е. как решения обыкновенного дифференциального уравнения (1.2), поскольку диффузия воды происходит, в основном, вдоль волокон. Отслеживание $x(t)$, $t > 0$, начиная с точки $x(0)$, известно как *задача об аксонной связности*. Получающиеся карты аксонных волокон очень важны для исследований мозга.

Также необходимы тесты для ответа на вопрос: *соединены ли разные участки мозга при помощи аксонов*, не требующие построения полной карты. Такие тесты и карты могут помочь прояснить вопросы геометрии живого мозга и могут потенциально привести к развитию средств диагностики таких заболеваний, как эпилепсия, рассеянный склероз, опухоль мозга и др. Возможные клинические приложения обсуждаются в работе [5].

Статистические задачи возникают вполне естественно в этих исследованиях. Наблюдения тензорного поля получают на живом мозге, в дискретных точках, с использованием диффузионной магнитно-резонансной томографии. Затем их усредняют по маленьким объемам. Вполне естественно, что тензор вместе с его собственными векторами искажен шумом, возникающим из-за физиологических процессов, эховых эффектов на молекулах воды и ошибок измерений. Все это определяет условия статистической задачи об оценивании месторасположения волокна. В ней исследователей интересуют не только способы отслеживания аксонных волокон, но и способы количественного оценивания неопределенностей в полученных изображениях, а также статистические тесты о том, достигает ли волокно $x(t)$, $t > 0$, определенной области Γ в мозгу.

Продолжает расти количество литературы, посвященной способам отслеживания волокон на основе данных о тензоре диффузии. Большинство предложенных методологий имеет две главные части: сглаживание или регуляризация наблюдаемых дискретных данных о тензоре и пошаговое отслеживание решения (1.2) вдоль собственного вектора, соответствующего максимальному собственному числу для сглаженной версии тензора. Несмотря на растущий интерес, только очень немногие статьи рассматривали статистические вопросы, неизбежно возникающие во всех методах исследования тензоров диффузии. Дополнительные детали и ссылки можно найти в работах [6] и [4]. Отметим, что последняя из упомянутых статей является, по мнению автора, первой и единственной работой, где исследуется точность аппроксимации в задаче об исследовании тензоров диффузии со статистической точки зрения. В настоящей работе получены нижние границы для возможной точности статистического оценивания в случае применения произвольных оценок интегральных кривых и произвольных оценок для минимального расстояния между кривой и подмножеством в проблеме об исследовании тензоров диффузии.

4. Доказательства. Мы воспользуемся идеями и техникой из монографии [3], которые использовались там для оценивания плотностей. Вначале строится небольшой параметрический подкласс в $\mathcal{C}^2(a, G, T)$. Затем с использованием результатов для *локально асимптотически нормальных* семейств распределений, устанавлива-

ются нижние границы для разности между параметром и его оценкой, что является ключевым моментом доказательств. Эта идея оформлена в основной лемме, которая приводится первой в этом разделе и может представлять самостоятельный интерес. Наконец, отклонение оценки для кривой от истинной кривой записывается как разность между этим самым параметром и его оценкой (см. (4.6) и (4.7)).

Мы будем следовать идее доказательства теоремы IV.5.1 из монографии [3]. Без ограничения общности далее будем предполагать, что $l = 1$, $a = 0$ и что G имеет меру Лебега, равную 1.

4.1. Основная лемма. Пусть $v_0: \mathbf{R}^d \rightarrow \mathbf{R}^d$ — векторное поле из класса $\mathcal{C}(0, G, T)$, и пусть $g: \mathbf{R}^d \rightarrow \mathbf{R}^d$ — дважды непрерывно дифференцируемое векторное поле с $\int |g(x)|^4 dx < \infty$. Дополнительно выберем функцию $g(x)$ не ортогональной функциям $\nabla \sqrt{f(y)}$ в $\mathbf{R}^d$ для всех x из открытого подмножества в $[c_1, c_2] \times \mathbf{R}^{d-1}$ и всех $y \in S$, где c_1, c_2 определены ниже, после формулы (4.4).

Для произвольного $M > 0$ определим параметрическое семейство векторных полей

$$v_n(x; \theta) = v_0(x) + \theta n^{-\alpha} g_n(x), \quad |\theta| < M, \quad (4.1)$$

где

$$g_n(x) = g(x_1, n^\gamma x_2, \dots, n^\gamma x_d), \quad \alpha = 2/(d+3), \quad 0 \leq \gamma \leq \alpha/2. \quad (4.2)$$

Заметим, что для случайной величины (наблюдения) $(X_i, V_i) = (X_i, v_n(X_i; \theta) + \xi_i)$ плотность можно записать как

$$\phi_n(x, y; \theta) = f(y - v_0(x) - \theta n^{-\alpha} g_n(x)) I(x \in G). \quad (4.3)$$

Следующая замена переменных будет неоднократно использоваться в доказательстве:

$$\begin{aligned} \tilde{x}_1 &= x_1, & \tilde{x}_i &= n^\gamma x_i, \quad i = 2, \dots, d, & \tilde{G}_n &= \{\tilde{x}: x \in G\}, \\ \tilde{y} &= y - v_0(x) - \theta n^{-\alpha} g_n(x). \end{aligned} \quad (4.4)$$

Поскольку G есть выпуклое множество, можно показать, что определенные в (4.4) множества $\tilde{G}_n$ сходятся к $[c_1, c_2] \times \mathbf{R}^{d-1}$.

Введем

$$I_0 := 4 \int_{\mathbf{R}^d} \int_{[c_1, c_2] \times \mathbf{R}^{d-1}} \langle \nabla \sqrt{f(y)}, g(x) \rangle^2 dx dy.$$

Лемма 1. Для любой оценки U_n параметра θ в параметрическом семействе (4.3), и любой функции потерь $w \in \mathcal{W}$ и любого числа $0 < c \leq \sqrt{I_0} M$ справедливо неравенство

$$\liminf_{n \rightarrow \infty} \sup_{|\theta| \leq c I_0^{-1/2}} \mathbf{E}_\theta^n w(\sqrt{I_0} |U_n - \theta|) \geq \frac{1}{2\sqrt{2\pi}} \int_{|y| \leq c/2} w(y) e^{-y^2/2} dy,$$

где символ $\mathbf{E}_\theta^n$ обозначает математическое ожидание по семейству мер

$$\int \cdots \int_A \prod_{i=1}^n \phi_n(x_i, y_i; \theta) dx_i dy_i, \quad (4.5)$$

порождаемому выборкой $\{(X_i, V_i) = (X_i, v_n(X_i; \theta) + \xi_i), i = 1, \dots, n\}$.

4.2. Доказательство основной леммы. Легко заметить, что функция ϕ_n дифференцируема по θ для всех x и y . Пользуясь этим фактом, ниже мы докажем четыре вспомогательные леммы. Так, из лемм 2 и 3 будет вытекать, что мы имеем дело с регулярным экспериментом, понимаемым в том же смысле, что в определении I.7 из монографии [3]. Далее, из лемм 2, 4 и 5 будет следовать, что условия теоремы II.3.1' в [3] выполнены, из чего, в свою очередь, можно заключить, что семейство мер (4.5) является локально асимптотически нормальным при $n \rightarrow \infty$. Наконец, мы воспользуемся теоремой II.12.1 и замечанием II.12.3 в [3] (см. также формулу (II.12.19)) и получим заключение леммы 1. Заметим, что наша основная лемма похожа на лемму IV.5.1 для оценок плотностей. Также отметим, что эта лемма может быть получена из хорошо известного результата Гаека [1].

Лемма 2. Пусть

$$\Psi^2(n, \theta) = n \int \int \left| \frac{d \ln \phi_n}{d \theta} \right|^2 \phi_n(x, y; \theta) dx dy.$$

Тогда $\Psi^2(n, \theta)/n$ положительна и ограничена для всех θ . Кроме того,

$$\lim_{n \rightarrow \infty} \Psi^2(n, \theta) = I_0 > 0.$$

Доказательство. Заметим, что $\Psi^2(n, \theta)/n$ — это информационное количество Фишера. Поэтому, воспользовавшись заменой переменных (4.4), функцию $\Psi^2(n, \theta)$ можно записать в виде

$$\begin{aligned} \Psi^2(n, \theta) &= n^{1-2\alpha} \int \int \left| \langle f^{-1/2} \nabla f(y - v_0(x) - \theta n^{-\alpha} g_n(x)), g_n(x) \rangle \right|^2 I(x \in G) dx dy \\ &= n^{1-2\alpha-(d-1)\gamma} \int \int \left| \langle f^{-1/2} \nabla f(\tilde{y}), g(\tilde{x}) \rangle \right|^2 I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \\ &=: n^{1-2\alpha-(d-1)\gamma} I_{0,n}. \end{aligned}$$

Заметим, что $\tilde{G}_n$ сходятся к $[c_1, c_2] \times \mathbf{R}^{d-1}$ и $I_{0,n}$ сходятся к I_0 . Поскольку выполнено неравенство $\langle f^{-1/2} \nabla f, g \rangle^2 \leq |f^{-1/2} \nabla f|^2 |g|^2$, то предположения $g \in \mathcal{L}_2(\mathbf{R}^d)$ и $\nabla \sqrt{f} \in \mathcal{L}_2(\mathbf{R}^d)$ гарантируют ограниченность $\Psi^2(n, \theta)/n$. Условие $\nabla \sqrt{f} \in \mathcal{L}_2(\mathbf{R}^d)$ эквивалентно условию $\int f |\nabla \log f|^2(y) dy < \infty$, которое, в свою очередь, вытекает из (F3).

Вновь воспользуемся заменой переменных (4.4), тогда легко видеть, что $\Psi^2(n, \theta)$ положительна для всех θ , если функция $\nabla \sqrt{\phi_n}$ не обращается в нуль на некотором открытом множестве из $\mathbf{R}^{2d}$. Это условие равносильно тому, что функция $g(x)$ не ортогональна к $\nabla \sqrt{f(y)}$ для всех x из открытого подмножества в $[c_1, c_2] \times \mathbf{R}^{d-1}$ и всех y из открытого подмножества в S , что гарантируется предположением (F2) и выбором функции g .

Отметим, что $I_0 = 0$ тогда и только тогда, когда функция g ортогональна к $\nabla \sqrt{f}$ в $\mathbf{R}^d$ для всех $x \in [c_1, c_2] \times \mathbf{R}^{d-1}$ и всех y . Таким образом, предположение (F2) вместе с выбором функции g ведет к неравенству $I_0 > 0$, что завершает доказательство леммы 2.

Лемма 3. Функция $\frac{d\sqrt{\phi_n}}{d\theta}(x, y; \theta)$ непрерывна по θ в $\mathcal{L}_2(G \times \mathbf{R}^d)$.

Доказательство. Покажем, что

$$\int \int \left(\frac{d\sqrt{\phi_n}}{d\theta}(x, y; \theta) \right)_{\theta}^{\theta+h} I(x \in G) dx dy \leq h^2 n^{-4\alpha-(d-1)\gamma} L,$$

где L — конечная положительная константа. Аналогично выводу для $\Psi^2(n, \theta)$ воспользуемся заменой переменных (4.4) и получим

$$\begin{aligned} &\int \int \left(\frac{d\sqrt{\phi_n}}{d\theta}(x, y; \theta) \right)_{\theta}^{\theta+h} I(x \in G) dx dy \\ &= n^{-2\alpha-(d-1)\gamma} \int \int \left\langle \nabla \sqrt{f}(\tilde{y} - h n^{-\alpha} g(\tilde{x})) - \nabla \sqrt{f}(\tilde{y}), g(\tilde{x}) \right\rangle^2 I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \\ &= n^{-2\alpha-(d-1)\gamma} \int \int \left\langle \int_0^1 \nabla^2 \sqrt{f}(\tilde{y} - \tau h n^{-\alpha} g(\tilde{x})) (-h) n^{-\alpha} g(\tilde{x}) d\tau, g(\tilde{x}) \right\rangle^2 \\ &\quad \times I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \\ &\stackrel{(*)}{\leq} n^{-2\alpha-(d-1)\gamma} \int \int \int_0^1 \left\langle \nabla^2 \sqrt{f}(\tilde{y} - \tau h n^{-\alpha} g(\tilde{x})) (-h) n^{-\alpha} g(\tilde{x}), g(\tilde{x}) \right\rangle^2 d\tau \end{aligned}$$

$$\begin{aligned} & \times I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \\ & \stackrel{(**)}{=} h^2 n^{-4\alpha-(d-1)\gamma} \iint \left\langle \nabla^2 \sqrt{f}(y) g(\tilde{x}), g(\tilde{x}) \right\rangle^2 I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y}, \end{aligned}$$

где в (*) использовано простое неравенство $(\int_0^1 p(\tau) d\tau)^2 \leq \int_0^1 p^2(\tau) d\tau$, а в (**) — замена переменных $y = \tilde{y} - \tau h n^{-\alpha} g(\tilde{x})$. Отметим, что

$$\iint \left\langle \nabla^2 \sqrt{f}(y) g(\tilde{x}), g(\tilde{x}) \right\rangle^2 I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \leq \int \left| \nabla^2 \sqrt{f}(y) \right|^2 dy \int |g(x)|^4 dx = L < \infty,$$

поскольку $\int |\nabla^2 \sqrt{f}(y)|^2 dy < \infty$ и $\int |g(x)|^4 dx < \infty$, что завершает доказательство леммы 3.

Лемма 4. Для любого $k > 0$

$$\lim_{n \rightarrow \infty} \sup_{|u| < k} \frac{n}{I_{0,n}} \iint \left(\frac{d\sqrt{\phi_n}}{d\theta}(x, y; \theta) \Big|_{\theta}^{\theta + u I_{0,n}^{-1/2}} \right)^2 dx dy = 0.$$

Доказательство. Положим $1 - 2\alpha - (d-1)\gamma = 0$ и $h = u I_{0,n}^{-1/2}$ в доказательстве леммы 3. Тогда для фиксированного $k > 0$

$$\lim_{n \rightarrow \infty} \sup_{|u| < k} \frac{n^{-2\alpha} u^2}{I_{0,n}^2} \iint \left\langle \nabla^2 \sqrt{f}(y) g(\tilde{x}), g(\tilde{x}) \right\rangle^2 I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} = \lim_{n \rightarrow \infty} O(n^{-2\alpha}) = 0.$$

Еще раз отметим, что требуется условие $1 - 2\alpha - (d-1)\gamma = 0$, которое с легкостью записывается как $\gamma = (1 - 2\alpha)/(d-1) = 1/(d+3) = \alpha/2$ и из которого вытекает лемма 4.

Лемма 5. Для некоторого $\delta \in (0, 2)$ выполнено условие Ляпунова

$$\lim_{n \rightarrow \infty} \Psi(n, \theta)^{-(2+\delta)} n \iint \left| \phi_n^{-1} \frac{d\phi_n}{d\theta} \right|^{2+\delta} \phi_n(x, y; \theta) dx dy = 0.$$

Доказательство. Используя замену переменных (4.4), получаем

$$\begin{aligned} & \lim_{n \rightarrow \infty} \Psi(n, \theta)^{-(2+\delta)} n \iint \left| \phi_n^{-1} \frac{d\phi_n}{d\theta} \right|^{2+\delta} \phi_n(x, y; \theta) dx dy \\ &= \lim_{n \rightarrow \infty} n^{1-\alpha(2+\delta)-\gamma(d-1)} I_{0,n}^{-1-\delta/2} \iint \left\langle f^{-1} \nabla f(\tilde{y}), g(\tilde{x}) \right\rangle^{2+\delta} f(\tilde{y}) I(\tilde{x} \in \tilde{G}_n) d\tilde{x} d\tilde{y} \\ &= \lim_{n \rightarrow \infty} O(n^{-\alpha\delta}) = 0, \end{aligned}$$

где требуется, чтобы $|g|^2 \in L_{1+0.5\delta}(\mathbf{R}^d)$, что выполнено для $\delta \leq 2$, а также необходимо условие $\int |\nabla f|^{2+\delta} f^{-1-\delta}(y) dy < \infty$, которое равносильно $\int f |\nabla \log f|^{2+\delta}(y) dy < \infty$, что гарантируется условием (F3). Этим завершается доказательство леммы 5.

4.3. Доказательство теоремы 1. Для начала опишем векторные поля v_0 и g , которые далее будут использоваться в качестве полей из определения параметрического семейства (4.1), следующим образом: $v_0(x) = b(x)e_1 = (b(x), 0, \dots, 0)$, $x \in \mathbf{R}^d$, где $b: \mathbf{R}^d \rightarrow \mathbf{R}$ — дважды непрерывно дифференцируемая функция, которая обращается в нуль вне G , причем $b(0) \neq 0$. Соответствующая интегральная кривая, определяемая дифференциальным уравнением

$$\frac{d}{dt} x_0(t) = v_0(x_0(t)), \quad x_0(0) = 0,$$

может быть записана как $x_{0,1}(t)$, $x_{0,i}(t) = 0$, $i = 2, \dots, d$. Заметим, что $x_0(t) \in G$ для всех $t > 0$. Дополнительно потребуем, чтобы v_0 удовлетворяло условию (A5) для некоторого положительного γ_0 . Тогда $v_0 \in \mathcal{C}^2(0, G, T)$.

Векторное поле $g: \mathbf{R}^d \rightarrow \mathbf{R}^d$ выбирается так, чтобы $g(x) = h(x)e_1 = (h(x), 0, \dots, 0)$, $x \in \mathbf{R}^d$, где $h: \mathbf{R}^d \rightarrow \mathbf{R}$ — дважды непрерывно дифференцируемая функция, $h(0) \neq 0$ и, более того, $h(x) = 0$ для всех x вне некоторой окрестности нуля, точнее, для всех x таких, что $|x| \geq \rho$ для некоторого фиксированного положительного ρ , выбранного так, чтобы множество $\{x: |x| < \rho\}$ лежало целиком в G . Дополнительно потребуем, чтобы для всех $t \in [0, T]$ выполнялось неравенство $h(x_{0,1}(t), 0, \dots, 0) > 0$.

Очевидно, что все вторые частные производные векторного поля $v_n = v_0 + \theta n^{-\alpha} g_n$ в (4.1), а именно,

$$\begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} v_n(x) &= \frac{\partial^2}{\partial x_i \partial x_j} v_0(x) + \theta n^{-\alpha+2\gamma} \frac{\partial^2}{\partial x_i \partial x_j} g(x_1, n^\gamma x_2, \dots, n^\gamma x_d), \\ &\quad i \neq 1, \quad j \neq 1, \\ \frac{\partial^2}{\partial x_1 \partial x_j} v_n(x) &= \frac{\partial^2}{\partial x_1 \partial x_j} v_0(x) + \theta n^{-\alpha+\gamma} \frac{\partial^2}{\partial x_1 \partial x_j} g(x_1, n^\gamma x_2, \dots, n^\gamma x_d), \\ &\quad j \neq 1, \\ \frac{\partial^2}{\partial x_1^2} v_n(x) &= \frac{\partial^2}{\partial x_1^2} v_0(x) + \theta n^{-\alpha} \frac{\partial^2}{\partial x_1^2} g(x_1, n^\gamma x_2, \dots, n^\gamma x_d), \end{aligned}$$

ограничены. Более того, $v_n(0) \neq 0$ и $v_n(x) = 0$ для $x \notin G$ и всех достаточно больших n . Заметим, что соответствующая интегральная кривая, определяемая дифференциальным уравнением

$$\frac{d}{dt} x_n(t; \theta) = v_n(x_n(t; \theta); \theta), \quad x_n(0; \theta) = 0,$$

может быть записана как $x_{n,1}(t; \theta)$, $x_{n,i}(t; \theta) = 0$, $i = 2, \dots, d$. Ниже будет показано, что векторное поле v_n удовлетворяет условию (A5) и, как следствие, принадлежит классу $\mathcal{C}^2(0, G, T)$ для всех достаточно больших n .

Но сначала свяжем $x_n(t; \theta)$ и θ . Положим $\Delta_n(t) := x_{n,1}(t; \theta) - x_{0,1}(t)$. Поскольку $x_0(t) = \int_0^t v_0(x_0(s)) ds$ и $x_n(t; \theta) = \int_0^t v_n(x_n(s; \theta); \theta) ds$, то разность Δ_n удовлетворяет следующему уравнению:

$$\Delta_n(t) = \int_0^t [b(x_n(s; \theta)) - b(x_0(s))] ds + \theta n^{-\alpha} \int_0^t h(x_{0,1}(s) + \Delta_n(s), 0, \dots, 0) ds.$$

Так как g ограничена на $[0, T]$, то, используя неравенство Гронуолла–Беллмана, получаем, что $|\Delta_n(t)| \leq C|\theta|n^{-\alpha}$ для фиксированного $t \in [0, T]$ и некоторой константы $C > 0$ (см., например, [2]).

Таким образом, имеем

$$\begin{aligned} \frac{1}{t-s} \left| \int_s^t v_n(x_n(\lambda; \theta)) d\lambda \right| &= \frac{1}{t-s} \left| \int_s^t [b(x_{0,1}(\lambda) + \Delta_n(\lambda), 0, \dots, 0) \right. \\ &\quad \left. + \theta n^{-\alpha} h(x_{0,1}(\lambda) + \Delta_n(\lambda), 0, \dots, 0)] d\lambda \right| \\ &\geq \left| \frac{1}{t-s} \int_s^t v_0(x_0(\lambda; \theta)) d\lambda - MCn^{-\alpha} \right| \geq |\gamma_0 - MCn^{-\alpha}| > 0 \end{aligned}$$

для всех достаточно больших n и некоторой константы $C > 0$, что означает выполнение условия (A5) для векторного поля v_n .

Также заметим, что

$$\begin{aligned} \Delta_n(t) &= \int_0^t [b(x_n(s; \theta)) - b(x_0(s))] ds + \theta n^{-\alpha} \int_0^t h(x_{0,1}(s), 0, \dots, 0) ds \\ &\quad + \theta n^{-\alpha} \int_0^t \int_0^1 h'_1(x_{0,1}(s) + \tau \Delta_n(s), 0, \dots, 0) d\tau \Delta_n(s) ds, \end{aligned}$$

где $h'_1(u_1, \dots, u_d) = \frac{\partial h}{\partial u_1}(u_1, \dots, u_d)$. Далее, для некоторой константы $C > 0$ имеем

$$\left| \int_0^t [b(x_n(s; \theta)) - b(x_0(s))] ds - \int_0^t b'_1(x_{0,1}(s), 0, \dots, 0) \Delta_n(s) ds \right| \leq C n^{-2\alpha}.$$

Значит, для некоторой константы $C > 0$ получаем

$$\left| \Delta_n(t) - \int_0^t b'(x_0(s)) \Delta_n(s) ds - \theta n^{-\alpha} \int_0^t h(x_{0,1}(s), 0, \dots, 0) ds \right| \leq C n^{-2\alpha}.$$

Решая интегральное уравнение, получающееся приравнованием левой части последнего неравенства к нулю, имеем

$$\begin{aligned} x_{n,1}(t_0; \theta) - x_{0,1}(t_0) &= \Delta_n(t_0) = \exp \left\{ \int_0^{t_0} b'(x_0(s)) ds \right\} \\ &\times \left[\theta n^{-\alpha} \int_0^{t_0} \exp \left\{ - \int_0^s b'(x_0(\tau)) d\tau \right\} h(x_{0,1}(s), 0, \dots, 0) ds + O(n^{-2\alpha}) \right] \\ &=: \theta n^{-\alpha} C_0 (1 + o(1)), \end{aligned}$$

где константа C_0 положительна в силу выбора g . (Отметим, что мы воспользовались условием (A5) для v_0 .)

Заметим, что существуют $r_n(t_0; \theta)$, сходящиеся к нулю, такие, что

$$\theta = \frac{n^\alpha}{C_0} (x_{n,1}(t_0; \theta) - x_{0,1}(t_0)) (1 + r_n(t_0; \theta)), \quad n \rightarrow \infty.$$

По оценке $\hat{X}_n(t_0)$ интегральной кривой $x_n(t_0; \theta)$ можно, в частности, сконструировать оценку параметра θ следующим образом: $\hat{\theta}_n := \frac{n^\alpha}{C_0} (\hat{X}_{n,1}(t_0) - x_{0,1}(t_0)) (1 + R_n(t_0))$, где $R_n(t_0)$ сходятся к 0 при $n \rightarrow \infty$ и, кроме того,

$$\sup_{|\theta| \leq M} |R_n(t_0) - r_n(t_0; \theta)| \leq \sup_{|\theta| \leq M} \frac{n^\alpha}{C_0} |\hat{X}_{n,1}(t_0) - x_{n,1}(t_0; \theta)|.$$

Таким образом, имеем

$$\begin{aligned} |\hat{\theta}_n - \theta| &= \frac{n^\alpha}{C_0} |\hat{X}_{n,1}(t_0) - x_{n,1}(t_0; \theta)| \\ &+ \frac{n^\alpha}{C_0} |(\hat{X}_{n,1}(t_0) - x_{0,1}(t_0)) R_n(t_0) - (x_{n,1}(t_0; \theta) - x_{0,1}(t_0)) r_n(t_0; \theta)| \\ &\leq \frac{n^\alpha}{C_0} |\hat{X}_{n,1}(t_0) - x_{n,1}(t_0; \theta)| |1 + R_n(t_0)| + \frac{|\theta|}{|1 + r_n(t_0; \theta)|} |R_n(t_0) - r_n(t_0; \theta)|. \end{aligned}$$

Поскольку $R_n(t_0)$ было выбрано особым образом, для достаточно больших n получаем

$$\sup_{|\theta| \leq M} \mathbf{E} w(\sqrt{I_0} |\hat{\theta}_n - \theta|) \leq \sup_{|\theta| \leq M} \mathbf{E} w \left(\sqrt{I_0} (2 + 2M) \frac{n^\alpha}{C_0} |\hat{X}_{n,1}(t_0) - x_{n,1}(t_0; \theta)| \right).$$

Теперь для любого $y \in \mathbf{R}^d$ определим функцию потерь $\tilde{w}(I_0^{1/2} (2 + 2M) y / C_0) := w(y)$. Очевидно, что $\tilde{w} \in \mathcal{W}$. Заметим, что для любой оценки $\hat{X}_n$ интегральной кривой и любой константы $0 < c \leq \sqrt{I_0 M}$ имеем

$$\sup_{v \in \mathcal{C}^2(0, G, T)} \mathbf{E} w \left(n^\alpha |\hat{X}_{n,1}(t_0) - x_1(t_0)| \right) \geq \sup_{|\theta| \leq c I_0^{-1/2}} \mathbf{E} w \left(n^\alpha |\hat{X}_{n,1}(t_0) - x_{n,1}(t_0; \theta)| \right). \quad (4.6)$$

Из основной леммы, примененной к функции $\tilde{w}$ при $c = \sqrt{I_0}M$, вытекает, что

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \inf_{\hat{X}_n \in \mathcal{E}_n(T)} \sup_{v \in \mathcal{C}^2(0, G, T)} \mathbf{E} w \left(n^\alpha |\hat{X}_{n,1}(t_0) - x_1(t_0)| \right) \\ & \geq \liminf_{n \rightarrow \infty} \sup_{|\theta| \leq cI_0^{-1/2}} \mathbf{E} \tilde{w}(\sqrt{I_0}(\hat{\theta}_n - \theta)) \geq \frac{1}{2\sqrt{2\pi}} \int_{|x| \leq \sqrt{I_0}M/2} \tilde{w}(x) e^{-x^2/2} dx. \end{aligned} \quad (4.7)$$

Таким образом, теорема 1 доказана, так как всегда можно число M выбрать настолько большим, чтобы правая часть в (4.7) стала положительной.

4.4. Доказательство теоремы 2. Мы будем следовать структуре доказательства теоремы 1. Пусть $t_0 = T$. Поскольку $\nabla \varphi(a) \neq 0$, то, не уменьшая общности, можно считать, что $a = 0$ и $\varphi'_1(0) \neq 0$. В этом случае существует окрестность нуля $\mathcal{N}_\rho = \{x \in \mathbf{R}^d: |x| \leq 2\rho\}$, в которой $\varphi'_1(x) \neq 0$. Выберем функцию b в определении $v_0(x) = b(x)e_1$ таким образом, чтобы соответствующая интегральная кривая $x_0(t)$ при $t \in [0, T]$ оставалась внутри $\mathcal{N}_\rho$. При сделанных предположениях получаем, что $\varphi'_1(x_0(t)) \neq 0$ и $(\varphi(x_0(t)))' = \nabla \varphi(x_0(t))v_0(x_0(t)) = \varphi'_1(x_0(t))b(x_0(t)) \neq 0$ для всех $t \in [0, T]$. Можно еще предположить, что $\varphi(x_0(0)) = \varphi(0) < \varphi(x_0(T))$. Если это не так, то надо заменить функцию b на $-b$. В результате находим, что

$$\varphi(x_0(T)) = \inf_{t \in [0, T]} \varphi(x_0(t)).$$

Поскольку $|\Delta_n(t)| \leq |\theta|n^{-\alpha}C_0$ для всех $t \in [0, T]$ и $|\theta| \leq M$, то существует такое n_0 , зависящее от ρ и M , что при каждом $n \geq n_0$ справедливо неравенство $|x_n(t; \theta)| \leq 2\rho$ для всех $t \in [0, T]$ и $|\theta| \leq M$. Поэтому $\varphi'_1(x_n(t; \theta))b(x_n(t; \theta)) < 0$ для всех $t \in [0, T]$, $|\theta| \leq M$ и всех $n \geq n_0$.

Таким образом, для всех $t \in [0, T]$, $|\theta| \leq M$ и всех $n \geq n_1$, где n_1 зависит от ρ , M и g , получаем

$$\nabla \varphi(x_n(t; \theta))v_n(x_n(t; \theta)) = \varphi'_1(x_n(t; \theta)) \left[b(x_n(t; \theta)) + \theta n^{-\alpha} g(x_n(t; \theta)) \right] < 0,$$

поскольку g и φ'_1 являются ограниченными функциями на $\mathcal{N}_\rho$. Значит,

$$\varphi(x_n(T; \theta)) = \inf_{t \in [0, T]} \varphi(x_n(t; \theta)).$$

Как результат, имеем

$$\begin{aligned} & \inf_{t \in [0, T]} \varphi(x_n(t; \theta)) - \inf_{t \in [0, T]} \varphi(x_0(t)) = \varphi(x_n(T; \theta)) - \varphi(x_0(T)) \\ & = \varphi'_1(x_0(T))(x_n(T) - x_0(T))(1 + o(1)) = \varphi'_1(x_0(T))\theta n^{-\alpha} C_0(T)(1 + o(1)). \end{aligned}$$

Положим $C_1 = \varphi'_1(x_0(T))C_0(T)$. Эта константа положительна по построению поля v_n и в силу условий на функцию φ . Как и в доказательстве теоремы 1, получим

$$\theta = \frac{n^\alpha}{C_1} \left(\inf_{t \in [0, T]} \varphi(x_n(t; \theta)) - \inf_{t \in [0, T]} \varphi(x_0(t)) \right) (1 + r_n(T; \theta))$$

и можем определить

$$\hat{\theta}_n = \frac{n^\alpha}{C_1} \left(\widehat{M}_n - \inf_{t \in [0, T]} \varphi(x_0(t)) \right) (1 + R_n(T))$$

для произвольной оценки $\widehat{M}_n$ минимального расстояния и правильно подобранных последовательностей $r_n(T; \theta)$ и $R_n(T)$.

Оставшаяся часть доказательства практически дословно повторяет доказательство теоремы 1.

СПИСОК ЛИТЕРАТУРЫ

1. *Hájek J.* Local asymptotic minimax and admissibility in estimation. — Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, v. 1. Berkeley: Univ. California Press, 1972, p. 175–194.
2. *Hille E.* Lectures on Ordinary Differential Equations. Reading: Addison-Wesley, 1969, 723 p.
3. *Ибрагимов И. А., Хасьминский Р. З.* Асимптотическая теория оценивания. М.: Наука, 1979, 527 с.
4. *Koltchinskii V., Sakhanenko L., Cai S.* Integral curves of noisy vector fields and statistical problems in diffusion tensor imaging: nonparametric kernel estimation and hypotheses testing. — *Ann. Statist.*, 2997, v. 35, № 4, p. 1576–1607.
5. *Le Bihan D., Mangin J.-F., Poupon C., Clark C., Pappata S., Molko N., Chabriet H.* Diffusion tensor imaging: concepts and applications. — *J. Magnetic Resonance Imaging*, 2001, v. 13, p. 534–546.
6. *Mori S., Barker P. B.* Diffusion magnetic resonance imaging: its principle and applications. — *The Anatomical Record*, 1999, v. 257, p. 102–109.

Поступила в редакцию
28.VIII.2008