S©MR

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ
Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru
ISSN 1813-3304

Том 22, № 2, стр. ?–? (2025) УДК 519.234
https://doi.org/10.33048/semi.2023.20.? MSC 62G08

УНИВЕРСАЛЬНЫЕ ЛОКАЛЬНО-ЛИНЕЙНЫЕ ЯДЕРНЫЕ ОЦЕНКИ ДЛЯ ПРОИЗВОДНОЙ РЕГРЕССИОННОЙ ФУНКЦИИ

С.С. Петренко, Ю.Ю. Линке

Представлено Н.С. Аркашовым

Abstract: The paper considers the problem of nonparametric regression, which consists in estimating the derivative of the regression function, when the values of the regression function are observed with an accuracy of random errors in some known set of fixed or random points (regressors). The solution of this problem, including methods of kernel smoothing, is devoted to an extensive literature. In the paper, the consistency of a new class of locally linear estimators is studied, while a more general condition on the regressors is used. With respect to the regressors, it is only required that they asymptotically densely fill the domain of the regression function. This condition includes both the case of fixed regressors, without the requirement of regularity, and the situation of random regressors, but without the assumption of some form of weak dependence of quantities or the fulfillment of ergodic properties.

Keywords: nonparametric regression, nonparametric derivative estimators, kernel estimation, universal local linear estimator, consistency, fixed design, random design.

Petrenko, S.S., Linke, Y.Y., Universal local linear kernel estimators for the derivative of a regression function.

Работа выполнена при поддержке программы фундаментальных научных исследований СО РАН (проект FWNF-2024-0001).

Поступила 21 июня 2025 г., опубликована ?? 2025 г.

1 Введение

Рассматривается следующая модель непараметрической регрессии: даны наблюдения (отклики) $X_1, \ldots, X_n$, которые представимы в виде

$$X_i = f(z_i) + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{1}$$

где достаточно гладкая скалярная функция $f(t)$, $t \in [0,1]$, неизвестна, ненаблюдаемые погрешности $\{\varepsilon_i\}$ являются центрированными случайными величинами, регрессоры $\{z_i\}$ могут быть как случайными, так и детерминированными, их значения нам известны. Таким образом, с точностью до случайных погрешностей наблюдаются значения $\{X_i\}$ регрессионной функции f в некотором известном наборе ее аргументов $\{z_i\}$. Задача состоит в том, чтобы по заданным парам $\{(z_i, X_i);\, i = 1, \ldots, n\}$ оценить производную f' регрессионной функции f.

Оценивание производной регрессионной функции (наряду с оцениванием самой функции) играет важную роль в анализе (см. подробности, например, в [1]-[5]). Решению этой задачи посвящена обширная литература (см., например, библиографию работы и ссылки в указанных публикациях). Начиная с 60-х годов, в непараметрической регрессии широко используются методы ядерного сглаживания. Популярны ядерные оценки и в задаче оценивания производных регрессионной функции (см., например, недавние публикации [6]–[19], а также более ранние работы [20]-[29]). Методы ядерного сглаживания относят к одному из основных подходов в задаче оценивания производных регрессионной функции (см., например, [1]-[5]).

В рассматриваемой задаче непараметрической регрессии нас интересуют условия на регрессоры. В известных нам работах, посвященных оцениванию производных регрессионной функции, модели с детерминированным и случайными регрессорами принято рассматривать отдельно (см., например, [1]-[36]). Если регрессоры фиксированы, то предполагается в том или ином смысле регулярное заполнение этими точками области задания регрессионной функции (см., например, [3], [9], [11], [12], [21], [22], [24], [29], [33], [35]). В ситуации случайных регрессоров либо предполагается, что эти величины независимы и одинаково распределены ([2], [5], [8], [10], [13], [16], [19], [25], [26], [27], [31], [32], [36]), либо для задания регрессоров используются те или иные формы слабой зависимости случайных величин или условия эргодичности (см., например, [6], [14], [15], [17], [18], [23], [28]).

В работе исследуются новые универсальные локально-линейные ядерные оценки для производной регрессионной функции, при этом используется более общее и обладающее рядом преимуществ условие на регрессоры, чем известные ранее в этой задаче. В работе семейство регрессоров представляет собой случайные величины в схеме серий, а в качестве параметра серии выступает объем наблюдений. Последнее позволяет включить в рассмотрение в качестве частного случая и модели с детерминированными регрессорами. Доказана состоятельность новых универсальных локально-линейных ядерных оценок для производной регрессионной функции. При этом относительно набора регрессоров требуется лишь, чтобы эти величины с высокой вероятностью образовывали измельчающееся разбиение области задания регрессионной функции. Данное условие в терминах плотных данных является по существу необходимым для восстановления регрессионной функции и ее производных. Оно универсально относительно природы корреляции регрессоров и позволяет в едином подходе рассматривать модели с детерминированными и случайными регрессорами, при этом без требования регулярности или слабой зависимости, что особенно важно для приложений.

Отметим, что исследовать новые универсальные ядерные оценки при близких к минимальным ограничениях на регрессоры во многом удается благодаря специальной структуре этих оценок, содержащей конструкции сумм определенным образом взвешенных наблюдений со структурой интегральных сумм Римана. Конструкции интегральных сумм открывают возможность исследовать асимптотические свойства оценок за счет близости интегральных сумм и соответствующих интегралов, а не предельных теорем теории вероятностей.

Ранее близкие условия в терминах плотных данных были предложены в [37]-[46] в задаче оценивания регрессионной функции, а также функций среднего и ковариации непрерывного случайного процесса. На первый взгляд в случае, когда известна некоторая оценка для регрессионной функции, имеется простейший путь получения оценки для производной этой функции, состоящий в дифференцировании оценки для самой функции. Но такой вариант оценивания годится, по-видимому, только в исключительных случаях, и не гарантирует «хорошей» оценки для производной даже в случае, когда регрессионная функция оценена «очень хорошо» (см. комментарии, например, в [1], [30]). Так что указанный простейший путь не решает поставленной задачи в достаточно широких условиях.

Работа устроена следующим образом: в разделе 2 приведен основной результат, доказательство которого отнесено в раздел 3.

2 Основной результат

Нам потребуется ряд условий на параметры модели.

(**M**) Даны двумерные наблюдения $\{(z_i, X_i);\, i = 1, \ldots, n\}$, представимые в виде

$$X_i = f(z_i) + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{2}$$

где $f(t)$ — неизвестная заданная на $[0, 1]$ дважды непрерывно дифференцируема скалярная функция, случайные погрешности $\{\varepsilon_i;\, i = 1, \ldots, n\}$

ненаблюдаемы, набор регрессоров $\{z_i;\ i = 1, \ldots, n\}$ состоит из случайных величин, вообще говоря, с неизвестными распределениями, не обязательно независимых или одинаково распределенных. Предполагается, что случайные величины $\{z_i;\ i = 1, \ldots, n\}$ могут зависеть от n.

Замечание 1. Условие (M) включает в себя и ситуацию фиксированных регрессоров. Отрезок $[0, 1]$ в качестве области изменения регрессоров мы рассматриваем исключительно с целью простоты изложения.

(**E**) При всех $n \ge 1$ случайные погрешности $\{\varepsilon_i;\ i = 1, \ldots, n\}$ с вероятностью 1 при всех $i, j \le n$, $i \ne j$, удовлетворяют следующим условиям:

$$\mathbb{E}_{\mathcal{F}_n}\varepsilon_i = 0, \qquad \sup_{i\le n}\mathbb{E}_{\mathcal{F}_n}\varepsilon_i^2 \le \sigma^2, \qquad \mathbb{E}_{\mathcal{F}_n}\varepsilon_i\varepsilon_j = 0, \tag{3}$$

где константа $\sigma^2 > 0$ неизвестна и не зависит от n, а символ $\mathbb{E}_{\mathcal{F}_n}$ обозначает условное математическое ожидание при фиксации σ-алгебры $\mathcal{F}_n$, порожденной случайными величинами $\{z_i;\ i = 1, \ldots, n\}$.

(**K**) Ядерная функция $K(t)$, $t \in \mathbb{R}$, является плотностью симметричного распределения с носителем $[-1, 1]$, т.е. $K(t) \ge 0$, $K(t) = K(-t)$ при всех $t \in [-1, 1]$ и $\int_{[-1,1]} K(t)dt = 1$. Кроме того, ядерная функция $K(t)$ удовлетворяет условию Липшица всюду на $\mathbb{R}$ с константой $L \ge 1$.

При некотором $h > 0$ положим $K_h(t) = h^{-1}K(h^{-1}t)$. Понятно, что при выполнении условия (K) функция $K_h(t)$ является плотностью распределения на $[-h, h]$.

Обозначим через $z_{n:1} \le \ldots \le z_{n:n}$ элементы вариационного ряда, построенного по выборке регрессоров $\{z_i;\ i = 1, \ldots, n\}$. Положим

$$z_{n:0} = 0, \quad z_{n:n+1} = 1, \quad \Delta z_{ni} = z_{n:i} - z_{n:i-1}, \quad i = 1, \ldots, n + 1.$$

Отклики из (2), ассоциированный с порядковой статистикой $z_{n:i}$, обозначим через X_{ni}.

Условимся, что все пределы берутся при $n \to \infty$. Относительно регрессоров нам потребуется следующее условие, которое обсудим далее, в замечании 6.

(**R**) Имеет место предельное соотношение $\delta_n \equiv \max_{1\le i\le n+1} \Delta z_{ni} \xrightarrow{p} 0$.

Наконец, для любого $h \in (0, 1)$ введем в рассмотрение следующий класс оценок для производной регрессионной функции f:

$$\widehat{f}'_{n,h}(t) = \sum_{i=1}^{n} \frac{w_{n0}(t)(z_{n:i} - t) + w_{n1}(t)}{w_{n0}(t)w_{n2}(t) - w_{n1}^2(t)} X_{ni}K_h(t - z_{n:i})\Delta z_{ni}, \tag{4}$$

где

$$w_{nj}(t) = \sum_{i=1}^{n}(t - z_{n:i})^j K_h(t - z_{n:i})\Delta z_{ni}, \quad j = 0, 1, 2. \tag{5}$$

Замечание 2. Нетрудно проверить, что ядерная оценка (4) является второй координатой двумерной точки, на которой достигается минимум

$$\min_{(a,b)} \sum_{i=1}^{n} \left(X_{ni} - a - b(z_{n:i} - t)\right)^2 K_h(t - z_{n:i}) \Delta z_{ni}. \tag{6}$$

Таким образом, предлагаемый класс оценок в известном смысле близок к классической локально–линейной оценке для производной, но во взвешенном методе наименьших квадратов используются несколько иные веса, определяемые порядковыми статистиками, построенными по набору регрессоров, а вместо исходных наблюдений X_i участвуют наблюдения X_{ni}, ассоциированные с соответствующими порядковыми статистиками.

Замечание 3. Ядерная оценка для регрессионной функции f, определяемая как первая координата двумерной точки, на которой достигается минимум в (6), была введена и исследована в [38], но лишь в предположении непрерывности функции f. Как мы уже отмечали во введении, оценка для регрессионной функции не позволяет получить оценку для производной регрессионной функции в достаточно широких условиях.

Замечание 4. В случае, когда в наборе $\{z_i;\, i = 1, \ldots, n\}$ имеются кратные точки, некоторые спейсинги Δz_{ni} обращаются в ноль, и мы теряем часть выборочной информации в оценке $\widehat{f}'_{n,h}(t)$, определенной в (4). В этом случае предлагается прежде, чем использовать оценку (4), несколько сократить выборку. А именно, заменить наблюдения X_i с одинаковыми точками z_i их средним арифметическим и оставить в новой выборке лишь одну точку из кратных. В этом случае усредненные наблюдения будут иметь меньшее зашумление, так что, несмотря на меньший объем новой выборки, мы в известной мере не теряем информацию, содержащуюся в исходной выборке.

Основной результат состоит в следующем.

Теорема 1. *Пусть выполнены условия* (M), (E), (K) *и* (R) *и последовательность положительных чисел* $h = h_n \to 0$ *такова, что*

$$\mathbb{E}\delta_n/h^3 \to 0, \qquad \mathbb{E}\delta_n^2/h^4 \to 0. \tag{7}$$

Тогда для любого фиксированного t *имеет место предельное соотношение*

$$\widehat{f}'_{n,h}(t) \xrightarrow{p} f'(t).$$

Замечание 5. Приведенные в теореме 1 достаточные условия (7) позволяют в тех или иных ситуациях оценить порядок скорости сходимости. Например, в случае так называемого эквидистантного плана (т.е. $z_i = i/n$, $i = 1, ..., n$) выполнено $\delta_n = 1/n$. Поэтому условие (7) гарантируется соотношением $nh^3 \to \infty$, что совпадает с порядками сходимости оценок для производных, полученными другими способами (см. [9], [47]). Также нетрудно видеть, что сходимость $\mathbb{E}\delta_n^2/h^6 \to 0$ гарантирует выполнение обоих условий в (7), поскольку $\mathbb{E}\delta_n/h^3 \le \left(\mathbb{E}\delta_n^2/h^6\right)^{1/2}$.

Замечание 6. Таким образом, единственное условие (R) на регрессоры, обеспечивающее поточечную состоятельность новой локально–линейной оценки для производной регрессионной функции, состоит в следующем: регрессоры с высокой вероятностью образуют измельчающееся разбиение области задания регрессионной функции, диаметр которого стремится к нулю по вероятности с увеличением объема выборки. На наш взгляд, условие (R) весьма наглядно и по сути является необходимым для восстановления как регрессионной функции, так и ее производных. Очевидно, что неслучайные регрессоры, регулярно заполняющие $[0,1]$, удовлетворяют условию (R). Если $\{z_i\}$ независимы и одинаково распределены, а отрезок $[0,1]$ является носителем их общего распределения, то условие (R) также выполнено. Если $\{z_i\}$ — стационарная последовательность с условием α-перемешивания и маргинальным распределением с носителем $[0,1]$, то условие (R) также выполнено. Все другие известные в литературе формы слабой зависимости регрессоров также влекут за собой условие вида (R). Но выполнение этого условия вполне возможно и для других типов зависимости, которая может быть более сильной, нежели классические условия слабой зависимости (например, когда не выполнены предельные теоремы типа законов больших чисел, см. подробности и примеры в [37]-[46]).

3 Доказательство теоремы 1

Всюду далее в этом разделе считаем, что выполнены условия теоремы 1. Введем следующие обозначения:

$$\beta_{n,i}(t) = \frac{(z_{n:i}-t)w_{n0}(t) + w_{n1}(t)}{w_{n0}(t)w_{n2}(t) - w_{n1}^2(t)}, \tag{8}$$

$$r_{n,h}(t) = \sum_{i=1}^{n} \beta_{n,i}(t)\big(f(z_{n:i}) - f(t) - f'(t)(z_{n:i}-t)\big)K_h(t-z_{n:i})\Delta z_{ni}, \tag{9}$$

$$\nu_{n,h}(t) = \sum_{i=1}^{n} \beta_{n,i}(t)K_h(t-z_{n:i})\Delta z_{ni}\varepsilon_{ni}, \tag{10}$$

где через $\{\varepsilon_{ni}\}$ обозначены погрешности, ассоциированные с порядковыми статистиками $\{z_{n:i}\}$.

Нам потребуется ряд вспомогательных утверждений. Справедлива

Лемма 1. *Имеет место представление*

$$\widehat{f'}_{n,h}(t) = f'(t) + r_{n,h}(t) + \nu_{n,h}(t).$$

Доказательство. Заметим прежде всего, что с учетом определений (4), (5) и (8), справедливы тождества

$$\widehat{f}'_{n,h}(t) \equiv \sum_{i=1}^{n} \beta_{n,i} X_{ni} K_h(t - z_{n:i}) \Delta z_{ni},$$

$$\sum_{i=1}^{n} \beta_{n,i}(t) K_h(t - z_{n:i})(z_{n:i} - t) \Delta z_{ni} \equiv 1,$$

$$\sum_{i=1}^{n} \beta_{n,i}(t) K_h(t - z_{n:i}) \Delta z_{ni} \equiv 0.$$

Принимая во внимание соотношение

$$X_{ni} = f(z_{n:i}) + \varepsilon_{ni}, \qquad i = 1, \dots, n,$$

утверждение леммы следует из указанных тождеств и определений (9) и (10). □

Замечание 7. Положим

$$A_{n,h}(t) = \{i : |t - z_{n:i}| \le h, 1 \le i \le n\}.$$

Подчеркнем, что ввиду свойств плотности $K_h(\cdot)$, область суммирования в (9) и (10), а также в (5), совпадает с множеством $A_{n,h}(t)$. Отмеченный факт является принципиальным для дальнейшего анализа. Условимся далее использовать символ $\sum$ без индексов только в случае, когда суммирование ведется по множеству индексов $i \in A_{n,h}(t)$.

Положим

$$c_* \equiv c_*(K) = \frac{\kappa_2 - \kappa_1^2}{96L(6L + \kappa_2 + \kappa_1/2)} < \frac{1}{864L}, \tag{11}$$

где символ κ_j, $j = 1, 2$, обозначает j-й абсолютный момент распределения с плотностью $K(t)$:

$$\kappa_j = \int_{-1}^{1} |u|^j K(u) du. \tag{12}$$

Легко видеть, что разность $\kappa_2 - \kappa_1^2$ представляет собой дисперсию невырожденного распределения, тем самым она строго положительна.

Нам потребуется следующее вспомогательная лемма, которая является частным случаем леммы 1 в [38]. Отметим, что в основе этого утверждения лежит близость между интегральными суммами Римана и соответствующими интегралами.

Лемма 2. *При $h < 1/2$ на подмножестве элементарных исходов, определяемых соотношением $\delta_n \le c_* h$, справедливы неравенства*

$$\sup_{t \in [0,1]} |w_{nj}(t)| \le 3Lh^j, \quad j = 0, 1, 2, \tag{13}$$

$$\inf_{t \in [0,1]} (w_{n0}(t) w_{n2}(t) - w_{n1}^2(t)) \ge \frac{1}{8}(\kappa_2 - \kappa_1^2) h^2. \tag{14}$$

Доказательство. Это утверждение доказано в [38] (см. лемму 1). □

Обозначим через $\omega_{f'}(h)$ модуль непрерывности функции f', т.е.

$$\omega_{f'}(h) = \sup_{x,y\in[0;1]:|x-y|\le h} |f'(x) - f'(y)|.$$

Справедливы также следующие два утверждения.

Лемма 3. *Для любого положительного $h < 1/2$ на подмножестве элементарных исходов, определяемых соотношением $\delta_n \le c_* h$, имеет место оценка*

$$|r_{n,h}(t)| \le \omega_{f'}(h)\frac{144L^2}{\varepsilon(\kappa_2 - \kappa_1^2)}.$$

Доказательство. По теореме Лагранжа о среднем значении, для любого $i \in A_{n,h}$ выполнено

$$f(z_{n:i}) - f(t) = f'(\xi_i)(z_{n:i} - t),$$
$$\xi_i \in [\min\{z_{n:i}, t\},\ \max\{z_{n:i}, t\}], \quad |\xi_i - t| \le h.$$

Кроме того, имеют место оценки

$$\sum \Delta z_{ni} \le (2h + \delta_n) \le 3h, \qquad K_h(t - z_{n:i}) \le Lh^{-1}, \tag{15}$$

где в первом случае мы учли, что $\delta_n \le c_* h$ и $c_* < 1$. Следовательно, с учетом определений (9), (8), (5), замечания 7 и утверждений леммы 2, справедлива следующая цепочка соотношений:

$$|r_{n,h}(t)| = \Big|\sum \beta_{n,i}(t)(f'(\xi_i) - f'(t))(z_{n:i} - t)K_h(t - z_{n:i})\Delta z_{ni}\Big| \le$$
$$\le w_{f'}(h)\sum |\beta_{n,i}(t)(z_{n:i} - t)|K_h(t - z_{n:i})\Delta z_{ni} \le$$
$$\le \frac{w_{f'}(h)}{w_{n0}(t)w_{n2}(t) - w_{n1}^2(t)}\Big(w_{n0}(t)w_{n2}(t) +$$
$$+ |w_{n1}(t)|\sum |z_{n:i} - t|K_h(t - z_{n:i})\Delta z_{ni}\Big) \le$$
$$\le \frac{w_{f'}(h)}{8^{-1}(\kappa_2 - \kappa_1^2)h^2}(3L\cdot 3Lh^2 + 3Lh\cdot h\cdot Lh^{-1}\cdot 3h) = w_{f'}(h)\frac{144L^2}{\kappa_2 - \kappa_1^2}.$$

При выводе этого соотношения, которое и доказывает лемму, мы также учли, что $w_{n0}(t)w_{n2}(t) - w_{n1}^2(t) \ge 0$ в силу неравенства Коши-Буняковского. □

Лемма 4. *При $h < 1/2$ на подмножестве элементарных исходов, определяемых соотношением $\delta_n \le c_* h$, для любого $\varepsilon > 0$ имеет место оценка*

$$\mathbb{P}_{\mathcal{F}_n}(|\nu_{n,h}(t)| > \varepsilon) \le \frac{4608\ L^4\sigma^2}{(\kappa_1 - \kappa_2)^2\varepsilon^2}\left(\frac{\delta_n}{h^3} + \frac{\delta_n^2}{h^4}\right),$$

где символ $\mathbb{P}_{\mathcal{F}_n}$ обозначает условную вероятность при фиксации σ-алгебры $\mathcal{F}_n$, введенной в условии (E).

Доказательство. Оценим величину $\mathbb{E}_{\mathcal{F}_n}\nu^2_{n,h}(t)$ и воспользуемся затем неравенством

$$\mathbb{P}_{\mathcal{F}_n}(|\nu_{n,h}(t)| \geq \varepsilon) \leq \varepsilon^{-2}\mathbb{E}_{\mathcal{F}_n}\nu^2_{n,h}(t). \qquad (16)$$

Нетрудно видеть, что погрешности $\{\varepsilon_{ni}\}$, ассоциированные с порядковыми статистиками $\{z_{n:i}\}$, удовлетворяют соотношениям (3) из условия (E). Следовательно, с учетом определения (10), выполнено

$$\mathbb{E}_{\mathcal{F}_n}\nu^2_{n,h}(t) = \mathbb{E}_{\mathcal{F}_n}\Big(\sum \beta_{n,i}(t)K_h(t - z_{n:i})\Delta z_{ni}\varepsilon_{ni}\Big)^2 \leq$$
$$\sigma^2 \sum \beta^2_{n,i}(t)K^2_h(t - z_{n:i})(\Delta z_{ni})^2 \leq$$
$$\leq 2\sigma^2\delta_n \sum \frac{w^2_{n0}(t)(z_{n:i} - t)^2 + w^2_{n1}(t)}{\big(w_{n0}(t)w_{n2}(t) - w^2_{n1}(t)\big)^2} K^2_h(t - z_{n:i})\Delta z_{ni} \leq$$
$$\leq \frac{2304\ L^2\sigma^2\delta_n}{(\kappa_1 - \kappa_2)^2h^2} \sum K^2_h(t - z_{n:i})\Delta z_{ni}. \qquad (17)$$

При выводе этой оценки мы воспользовались также определением (8), утверждениями леммы 2, замечания 7 и элементарным неравенством $(a + b)^2 \leq 2a^2 + 2b^2$. Остается оценить сумму $\sum K^2_h(t - z_{n:i})\Delta z_{ni}$, в которой суммирование, с учетом замечания 7, ведется по множеству индексов $i : |z_{n:i} - t| \leq h$. Отметим, что в силу условия (K) выполнено $\sup_{y\in[-1,1]} K(y) \leq L$, а потому с учетом первого соотношения в (15), имеем

$$\sum K^2_h(t - z_{n:i})\Delta z_{ni} = \frac{1}{h^2}\sum K^2\left(\frac{t - z_{n:i}}{h}\right)\Delta z_{ni} \leq$$
$$\leq \frac{L^2}{h^2}\sum \Delta z_{ni} \leq \frac{L^2}{h^2}(2h + \delta_n).$$

Эта оценка вместе с (17) и (16) доказывает утверждение леммы. □

Лемма 5. *В условиях теоремы 1 имеют место предельные соотношения*

$$r_{n,h}(t) \xrightarrow{p} 0, \qquad \nu_{n,h}(t) \xrightarrow{p} 0.$$

Доказательство. В силу неравенства Маркова с первым моментом и утверждения леммы 3, для любого фиксированного $\varepsilon > 0$ выполнено

$$\mathbb{P}(|r_{n,h}(t)| > \varepsilon) = \mathbb{P}(|r_{n,h}(t)| > \varepsilon,\ \delta_n \leq c_*h) + \mathbb{P}(|r_{n,h}(t)| > \varepsilon,\ \delta_n > c_*h) \leq$$
$$\leq \mathbb{P}(c\, w_{f'}(h) > \varepsilon) + \mathbb{P}(\delta_n > c_*h) \leq \mathbb{P}(c\, w_{f'}(h) > \varepsilon) + (c_*h)^{-1}\mathbb{E}\delta_n,$$

где $c = 144L^2(\kappa_2 - \kappa_1^2)^{-1}$. Но поскольку $h \to 0$ с ростом n, то начиная с некоторого n выполнено $\mathbb{P}(cw_{f'}(h) > \varepsilon) = 0$, а соотношение $\mathbb{E}\delta_n/h \to 0$ следует из (7). Отмеченные факты вместе с приведенной оценкой доказывают первое утверждение леммы.

Аналогично вышесказанному, с учетом леммы 4 для любого фиксированного $\varepsilon > 0$ справедливо соотношение

$$\mathbb{P}(|\nu_{n,h}(t)| > \varepsilon) = \mathbb{P}(|\nu_{n,h}(t)| > \varepsilon,\ \delta_n \le c_* h) + \mathbb{P}(|\nu_{n,h}(t)| > \varepsilon,\ \delta_n > c_* h) \le$$
$$\le \mathbb{E}\big(I(\delta_n \le c_* h)\mathbb{P}_{\mathcal{F}_n}(|\nu_{n,h}(t)| > \varepsilon)\big) + \mathbb{P}(\delta_n > c_* h) \le$$
$$\le \frac{4608\ L^4\sigma^2}{(\kappa_1-\kappa_2)^2\varepsilon^2}\left(\frac{\mathbb{E}\delta_n}{h^3} + \frac{\mathbb{E}\delta_n^2}{h^4}\right) + \frac{\mathbb{E}\delta_n}{c_* h},$$

где $I(\cdot)$ — индикаторная функция. Второе утверждение леммы следует теперь из условия (7). □

Для завершения доказательства теоремы 1 нам остается воспользоваться леммами 1 и 5. Теорема доказана.

References

[1] De Brabanter K., De Brabanter J., De Moor B., Gijbels I. *Derivative estimation with local polynomial fitting*, J. Mach. Learn. Res., **14** (2013), 281–301.

[2] Liu Z., Li M. *On the estimation of derivatives using plug-in kernel ridge regression estimators*, J. Mach. Learn. Res., **24** (2023), 1–37.

[3] Wang W. W., Lin L. *Derivative estimation based on difference sequence via locally weighted least squares regression*, J. Mach. Learn. Res., **16** (2015), 2617–2641.

[4] Wang W., Yu P., Lin L., Tong T. *Robust estimation of derivatives using locally weighted least absolute deviation regression*, J. Mach. Learn. Res., **20** (2019), 1–49.

[5] Comte F., Marie N. *On a projection estimator of the regression function derivative*, J. Nonparametr. Stat., **35** (2023), 773–819.

[6] Delecroix M., Rosa A.C. *Nonparametric estimation of a regression function and its derivatives under an ergodic hypothesis*, J. Nonparametr. Stat., **6** (2007), 367–382.

[7] Li D., Lu Z., Linton O. *Local linear fitting under near epoch dependence: uniform consistency with convergence rates*, Econometric Theory, **28** (2012), 935–958.

[8] Zheng Q., Gallagher C., Kulasekera K. B. *Adaptively weighted kernel regression*, J. Nonparametr. Stat., **25**, (2013), 855–872.

[9] Chen, J., Zhang, L.X. *Local linear M-estimation for spatial processes in fixed-design models*, Metrika, **71** (2010), 319–340.

[10] Yao W., Lindsay B. G., Li R. *Local modal regression*, J. Nonparametr. Stat., **24** (2012), 647–663.

[11] Liu S., Yang J. *Kernel regression for estimating regression function and its derivatives with unknown error correlations*, Metrika, **87** (2024), 1–20.

[12] Boente G., Rodriguez D. *Robust estimators of high order derivatives of regression functions*, Statist. Probab. Lett., **76**, 1335–1344.

[13] Xie Q., Sun Q., Liu J. *Local weighted composite quantile estimation and smoothing parameter selection for nonparametric derivative function*, Econom. Rev., **39**, (2020), 215–233.

[14] Lu Z., Linton O. *Local linear fitting under near epoch dependence*, Econom. Theory, **23**, 37–70.

[15] El Machkouri M., Es-Sebaiy K., Ouassou I. *On local linear regression for strongly mixing random fields*, J. Multivar. Anal., **156** (2017), 103–115.

[16] Bercu B., Capderou S., Durrieu G. *Nonparametric recursive estimation of the derivative of the regression function with application to sea shores water quality*, Stat. Inference Stoch. Process, **22** (2019), 17–40.

[17] Martins-Filho C., Saraiva P. *On asymptotic normality of the local polynomial regression estimator with stochastic bandwidths*, Commun. Stat. Theory Methods, **41** (2012), 1052–1068.
[18] Prasangika K. D., W Tang W., Yao Z. *Double smoothing local linear estimation in nonlinear time series*, Commun. Stat. Theory Methods, **52** (2023), 1385–1399.
[19] Calonico S., Cattaneo M. D., Farrell M. H. *Coverage error optimal confidence intervals for local polynomial regression*, Bernoulli, **28** (2022), 2998–3022.
[20] Muller H.-G., Stadtmuller U., Schmitt T. *Bandwidth choice and confidence intervals for derivatives of noisy data*, Biometrika, **74** (1987), 743–749.
[21] Gasser T., Muller H.-G. *Estimating regression functions and their derivatives by the kernel method*, Scand. J. Statist., **11** (1984), 171–185.
[22] Hardle W., Gasser T. *On robust kernel estimation of derivatives of regression functions*, Scand. J. Statist., **12** (1985), 233–240.
[23] Jiang J., Mack Y. P. *Robust local polynomial regression for dependent data*, Stat. Sin., **11** (2001), 705-722.
[24] Beran, J., Feng, Y. *Local polynomial fitting with long-memory, short-memory and antipersistent errors*, Ann. Inst. Stat. Math., **54** (2002), 291–311.
[25] Fan J., Gasser T., Gijbels I., Brockmann M., Engel J. *Local polynomial regression: optimal kernels and asymptotic minimax efficiency*, Ann. Inst. Stat. Math., **49** (1997), 79–99.
[26] Mack Y. P., Muller H.-G. *Derivative estimation in nonparametric regression with random predictor variable*, Sankhya Indian J. Stat. Ser. A, **51** (1989), 59–72.
[27] Fan J., Gijbels I. *Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation*, J. R. Stat. Soc. (Series B), **57** (1995), 371–394.
[28] Masry E., Fan J. *Local polynomial estimation of regression functions for mixing processes*, Scand. Stat. Theory Appl, **24** (1997), 165–179.
[29] Beran J., Feng Y. *Local polynomial estimation with a FARIMA-GARCH error process*, Bernoulli, **7** (2001), 733–750.
[30] Newell J., Einbeck J. *A comparative study of nonparametric derivative estimators*. Proc. of the 22nd International Workshop on Statistical Modelling, (2007), 453–456.
[31] Zhou S., Wolfe D. A. *On derivative estimation in spline regression*, Statistica Sinica, **10** (2000), 93–108.
[32] Cattaneo M. D., Farrell M. H., Feng Y. *Large sample properties of partitioning-based series estimators*, Ann. Stat., **48** (2020), 1718–1741.
[33] Dai W., Tong T., Genton M.G. *Optimal estimation of derivatives in nonparametric regression*, J. Mach. Learn. Res., **17** (2016), 1–25.
[34] Liu Y., De Brabanter K. *Smoothed nonparametric derivative estimation using weighted difference quotients*, J. Mach. Learn. Res., **21** (2020), 1–45.
[35] Song Q. *Non parametric derivative estimation with confidence bands*, Commun. Stat. Theory Methods, **45** (2016), 277–290.
[36] Qin G., Tsao M. *Empirical likelihood based inference for the derivative of the nonparametric regression function*, Bernoulli, **11** (2005), 715–735.
[37] Borisov I.S., Linke Yu.Yu., Ruzankin P.S. *Universal weighted kernel-type estimators for some class of regression models*, Metrika, **84** (2021), 141–166.
[38] Linke Y., Borisov I., Ruzankin P., Kutsenko V., Yarovaya E., Shalnova S. *Universal local linear kernel estimators in nonparametric regression*, Mathematics, **10** (2022), 2693.
[39] Linke Yu. Yu., Borisov I. S., Ruzankin P. S. *Universal kernel-type estimation of random fields*, Statistics, **57** (2023), 785–810.
[40] Linke Y., Borisov I., Ruzankin P., Kutsenko V., Yarovaya E., Shalnova S. *Multivariate universal local linear kernel estimators in nonparametric regression: uniform consistency*, Mathematics, **12** (2024), 1890.

[41] Linke Yu. Yu., Borisov I. S. *Insensitivity of Nadaraya–Watson estimators to design correlation*, Commun. Stat. Theory Methods, **51** (2022), 6909–6918.
[42] Linke Yu. Yu. *Towards insensitivity of Nadaraya–Watson estimators to design correlation*, Theory Probab. Appl., **68** (2023), 198–210.
[43] Linke Yu. Yu. *On sufficient conditions for the consistency of local linear kernel estimators*, Math. Notes, **114** (2023), 283–296.
[44] Linke Yu. Yu. *Kernel estimators for the mean function of a stochastic process under sparse design conditions*, Siberian Adv. Math., **32** (2022), 269–276.
[45] Linke Yu. Yu., Borisov I.S. *Universal nonparametric kernel-type estimators for the mean and covariance functions of a stochastic process*, Theory Probab. Appl., **69** (2024), 35–58.
[46] Linke Yu. Yu. *Mean function estimation for a noisy random process under a sparse data condition*, Chebyshevskii Sbornik, **24** (2023), 112–125 (In Russ.).
[47] Nyssbaum M. hrefhttps://doi.org/10.1137/1131010 *Nonparametric estimation of a regression function that is smooth in a domain in R^k*, Theory Probab. Appl., **31** (1987), 108–115.

Petrenko Sergey Sergeevich
Sobolev Institute of Mathematics,
pr. Koptyuga, 4,
630090, Novosibirsk, Russia
Email address: s.petrenko@g.nsu.ru

Yuliana Yurievna Linke
Sobolev Institute of Mathematics,
pr. Koptyuga, 4,
630090, Novosibirsk, Russia
Email address: linke@math.nsc.ru