УДК 519.23 МАТЕМАТИЧЕСКИЕ ТРУДЫ
DOI 10.33048/mattrudy.2023.26.209 2023, том 26, № 2, 177-191

ОБ ОДНОМ ПОДХОДЕ К ПОСТРОЕНИЮ ЯВНЫХ ОЦЕНОК В ЗАДАЧАХ НЕЛИНЕЙНОЙ РЕГРЕССИИ

Ю. Ю. Линке, И. С. Борисов

Рассматривается задача построения явных состоятельных оценок конечномерных параметров моделей нелинейной регрессии с помощью различных непараметрических ядерных оценок.

Ключевые слова и фразы: нелинейная регрессия, непараметрическая регрессия, ядерные оценки, равномерная состоятельность, фиксированные регрессоры, случайные регрессоры.

§ 1. Введение и постановка задачи

Рассматривается следующая регрессионная модель:

$$X_i = f_\theta(\mathbf{z}_i) + \xi_i, \qquad i = 1, \dots, n, \tag{1}$$

где $\{f_\theta(\mathbf{t});\, \boldsymbol{\theta} = (\theta_1, \dots, \theta_m) \in \boldsymbol{\Theta}\}$ — параметрическое семейство вещественнозначных непрерывных функций k переменных, k-мерные векторы $\mathbf{t} = (t_1, \dots, t_k)$ принадлежат измеримому по Жордану компактному множеству $\mathcal{P} \subset \mathbb{R}^k$ без изолированных точек, $\boldsymbol{\Theta}$ – открытое множество в $\mathbb{R}^m$. Набор регрессоров $\{\mathbf{z}_i;\, i = 1, \dots, n\}$ состоит из наблюдаемых случайных k-мерных векторов с возможно неизвестными распределениями и со значениями в $\mathcal{P}$, при этом не обязательно независимых или одинаково распределенных. Набор регрессоров может быть рассмотрен в схеме серий, т.е. случайные векторы $\{\mathbf{z}_i;\ i = 1, \dots, n\}$ могут зависеть от n. В частности, данная схема включает в себя модели с фиксированными регрессорами (например, схему эквидистантного плана эксперимента). Погрешности $\{\xi_i;\, i = 1, \dots, n\}$ — центрированные случайные величины, дополнительные ограничения на которые мы сформулируем в конце этого параграфа.

Задача состоит в построении явных оценок для неизвестного значения m-мерного параметра $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ по выборке $(k+1)$-мерных наблюдений $\{(X_i, \mathbf{z}_i);\, i = 1, \dots, n\}$. Отметим, что проблема построения явных оценок в моделях нелинейной регрессии при тех или иных ограничениях на их параметры уже привлекала внимание специалистов (см. [23], [7], [27], [25],

---

Работа выполнена в рамках государственного задания ИМ СО РАН (проект FWNF-2022-0015)

[28], [1], [3], [37], [8], [34], [36], [20], [21], [19], [16]). Прежде всего речь идет о моделях, оценивание в которых может быть сведено к задачам линейной регрессии. Традиционно (см., например, [7], [27], [8], [25]), к *внутренне линейным* (*intrinsically linear*) или *внешне нелинейным* (*nonintrinsically nonlinear*) моделям относят модели регрессии, для которых регрессионное уравнение теми или иными преобразованиями откликов и исходных параметров может быть приведено к линейному. Как правило, речь здесь идет о моделях с мультипликативным шумом (см., например, [27]). В [20] было уточнено и расширено определение внутренне линейных моделей и установлено, что целый ряд известных моделей нелинейной регрессии с аддитивными погрешностями удовлетворяет этому определению. Для таких моделей явные оценки параметров строятся методами линейного регрессионного анализа.

Стоит отметить, что возможность преобразования исходной модели регрессии к линейной (в особенности в случае аддитивных погрешностей) является скорее редким исключением, нежели правилом. В работах [16] и [19] был предложен подход, который в широких условиях позволяет строить явные состоятельные оценки параметров нелинейных регрессионных моделей, не являющихся внутренне линейными. В данной заметке авторы еще более расширяют этот класс нелинейных регрессионных моделей.

Важно отметить, что помимо самостоятельного интереса, построение явных состоятельных оценок для нелинейных моделей регрессии представляет исключительный интерес и для *одношагового оценивания*. Хорошо известно, что в задачах нелинейной регрессии в известном смысле асимптотически оптимальные оценки (например, оценки методов квазиправдоподобия, наименьших квадратов, максимального правдоподобия) зачастую задаются неявно в виде решений тех или иных уравнений (см., например, монографии [3], [7], [9], [11], [25], [30], [34], [36], [37]). Вместе с этим, ситуация, когда имеется несколько корней того или иного уравнения, определяющего оценку, является весьма типичной (см., например, [31], [32], [33]). Данное обстоятельство является главной проблемой, затрудняющей использование численных методов. Дело в том, что при неудачном выборе начального приближения параметра итерационные процедуры обнаруживают лишь корень, ближайший к стартовой точке, а не к параметру. Один из способов обойти указанную проблему состоит в использовании так называемых *одношаговых* оценок. Идея одношагового оценивания, восходящая к работам Р. Фишера, заключается в следующем: в качестве стартовой точки итерационной процедуры ньютоновского типа используется не произвольная точка, а предварительная состоятельная оценка, сходящаяся к параметру с некоторой скоростью. Оказывается, в этом случае зачастую достаточно лишь одного шага итерационной процедуры, чтобы получить явную оценку (так называемую «одношаговую», в англоязычной литературе — *one–step*), имеющую ту же асимптотическую точность, что и искомая статистика (см., например, [35]). По сути, если мы располагаем некоторой предварительной состоятельной оценкой, то мы получаем возможность с помощью ньютоновских процедур выделить тот из корней интересующего нас уравнения, который, собственно, и приближает параметр.

В последние годы интерес к одношаговому оцениванию в статистической литературе только нарастает, и библиография в этой области весьма обширна (ряд библиографических ссылок можно найти, например, в [14]). Важность разработки методологии одношагового оценивания именно для задач нелинейной регрессии подчеркивается, например, в монографии [31]. В тех или иных постановках задач нелинейной регрессии одношаговые оценки исследуются, например, в работах [4], [24], [23], [26], [29], [31], но существование предварительной оценки в этих работах (за исключением статьи [23], посвященной дробно–линейным моделям) лишь постулируется. В работах [13], [14] и [15], связанных с одношаговым оцениванием в задачах нелинейной регрессии, при построении предварительных оценок используется прием из [16] и [19]. Новые оценки, предложенные в данной заметке, также могут быть использованы в качестве предварительных в одношаговых процедурах оценивания.

Вернемся к условиям на модель (1). Считаем, что погрешности $\{\xi_i;\, i = 1, \ldots, n\}$ образуют последовательность мартингал-разностей с условием

$$M_p = \sup_{i\le n} \mathbb{E}|\xi_i|^p < \infty \quad \text{при} \quad \text{некотором } p > k \text{ и } p \ge 2,$$

где M_p не зависит от n. Предполагается также, что случайные величины $\{\xi_i\}$ не зависят от $\{\mathbf{z}_i\}$, но могут зависеть от n. Далее, мы предполагаем, что с вероятностью 1 каждая наблюдаемая точка $\mathbf{z}_i$ из набора регрессоров имеет в выборке кратность 1. Отметим, что если бы в выборке $\{\mathbf{z}_i;\, i = 1, \ldots, n\}$ имелись кратные точки, то мы могли бы рассматривать среднее арифметическое откликов X_i с одинаковыми регрессорами и, тем самым, сводить задачу к исходной. Важно отметить, что при наличии растущей (с ростом n) кратности той или иной точки из набора регрессоров задача сводится к классической статистической постановке метода моментов. Поэтому такие случаи мы исключаем из рассмотрения. Так что даже при наличии неслучайных регрессоров мы отвергаем возможность «полного контроля» над ними, когда можно рассматривать растущее число откликов с одной и той же точкой из набора регрессоров.

Для каждого n обозначим через ε_n минимально возможный размер ε-сети, образованной набором регрессоров $\{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$ в компактном множестве $\mathcal{P}$. Итак, единственным ограничением на регрессоры будет следующее условие:

(D) $\varepsilon_n \xrightarrow{p} 0$ *при* $n \to \infty$.

Замечание 1. Если все $\mathbf{z}_i$ не зависят от n, то сходимость по вероятности в приведенном условии будет эквивалентна сходимости почти наверное в силу монотонности последовательности $\{\varepsilon_n\}$. Например, если $\{\mathbf{z}_i;\ i = 1, 2, \ldots\}$ – последовательность одинаково распределенных случайных векторов (не обязательно стационарная) с условием сильного перемешивания и компакт $\mathcal{P}$ является носителем маргинального распределения, то условие (D) будет выполнено. Примеры более сильной корреляции регрессоров, когда не выполняются все известные условия слабой зависимости, но имеет место условие (D), приведены в [12], [17].

§ 2. Методология получения оценок

Для простоты изложения далее полагаем $\mathcal{P} = [0,1]^k$. Обозначим через $C[0,1]^k$ пространство непрерывных функций на $[0,1]^k$. Предположим, что в пространстве $C[0,1]^k$ задана некоторая норма $\|\cdot\|$. Так что далее мы будем рассматривать линейное нормированное пространство $(C[0,1]^k, \|\cdot\|)$, которое предполагается сепарабельным относительно метрики, порожденной нормой $\|\cdot\|$. Нас будут интересовать, главным образом, два случая, когда

$$\|f\| \equiv \|f\|_{sup} = \sup_{\mathbf{t}\in[0,1]^k} |f(\mathbf{t})|, \qquad \|f\| \equiv \|f\|_{pw} = \sum_{i=1}^{\infty} \frac{|f(\mathbf{t}_i)|}{2^i},$$

где суммирование берется по всем точкам $\mathbf{t}_i \in [0,1]^k$ с рациональными координатами, занумерованным произвольным образом. Отметим, что сходимость в норме $\|\cdot\|_{pw}$ эквивалентна поточечной сходимости функций из $C[0,1]^k$.

Предлагаемый подход основан на использовании непараметрических ядерных оценок регрессионной функции. Нам понадобятся следующие предположения:

(I) *существует непрерывное отображение* $\mathbf{G} : (C[0,1]^k, \|\cdot\|) \to \mathbb{R}^m$, *для которого векторная функция* $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{G}(f_\theta)$ *является гомеоморфизмом открытого множества* $\boldsymbol{\Theta}$ *на некоторую область пространства* $\mathbb{R}^m$;

(II) *существует* $\|\cdot\|$*-состоятельная (или сильно* $\|\cdot\|$*-состоятельная) непараметрическая оценка* $f_n^*(\mathbf{t}) \in (C[0,1]^k, \|\cdot\|)$ *для регрессионной функции* $f_{\theta_0}(\mathbf{t})$, *где* $\boldsymbol{\theta}_0$ *— истинное значение параметра в схеме наблюдений* (1).

Теперь определим оценку

$$\boldsymbol{\theta}_n^* = \mathbf{g}^{-1}(\mathbf{G}(f_n^*)),$$

где $\mathbf{g}^{-1}$ — обратное преобразование для $\mathbf{g}$; при этом если $\mathbf{G}(f_n^*) \notin \{\mathbf{g}(\boldsymbol{\theta});\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$, то по определению полагаем $\mathbf{g}^{-1}(\mathbf{G}(f_n^*)) = \mathbf{0}$.

Справедливо следующее утверждение.

Теорема 1. *При выполнении предположений* (I) *и* (II) *оценка* θ_n^* *будет состоятельной* (*или сильно состоятельной в соответствии с предположением в п.* (II)) *для* θ_0.

Доказательство этого утверждения достаточно прозрачное. В самом деле, скажем, $\|\cdot\|$-состоятельность непараметрической оценки $f_n^*(\mathbf{t})$ означает, что

$$\|f_n^* - f_{\theta_0}\| \xrightarrow{p} 0 \ \text{ при } \ n \to \infty.$$

Иными словами, с вероятностью, стремящейся к 1, в силу непрерывности преобразования $\mathbf{G}$ в норме $\|\cdot\|$ пространства $C[0,1]^k$ выполнено $\mathbf{G}(f_n^*) \in S_{g(\theta_0)}(\varepsilon)$ для любого наперед заданного $\varepsilon > 0$, где $S_{g(\theta_0)}(\varepsilon)$ — открытый шар радиуса ε с центром в точке $\mathbf{g}(\boldsymbol{\theta}_0)$ пространства $\mathbb{R}^m$. Поскольку образ открытого множества при гомеоморфном преобразовании будет открытым, то для достаточно малого $\varepsilon > 0$ имеет место вложение $S_{g(\theta_0)}(\varepsilon) \subseteq \{\mathbf{g}(\boldsymbol{\theta});\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Остается воспользоваться непрерывностью обратного преобразования $\mathbf{g}^{-1}$, откуда и следует, что $\boldsymbol{\theta}_n^* \xrightarrow{p} \boldsymbol{\theta}_0$ при $n \to \infty$. Аналогичные рассуждения используются и для доказательства сильной $\|\cdot\|$- состоятельности. □

Замечание 2. Предложенная в теореме 1 методика построения оценок конечномерных параметров в задачах нелинейной регрессии близка к методологии метода моментов. В самом деле, теорема 1 по сути предлагает приравнять значения регрессионной функции в тех или иных точках из области ее определения к соответствующим значениям ее состоятельной непараметрической оценки. При этом число таких уравнений (или что то же — указанных точек) должно совпадать с размерностью параметра регрессионной функции. Именно это происходит в методе моментов, когда для построения оценок параметров приравниваются истинные моменты (как функции от рассматриваемого параметра) к соответствующим выборочным моментам, которые, в свою очередь, будут состоятельными оценками для истинных моментов. Так что подход, предложенный в теореме 1, по аналогии с методом моментов может быть изложен следующим образом. Сначала выписывается система уравнений

$$f_\theta(\mathbf{t}_j) = f_n^*(\mathbf{t}_j), \quad j = 1, \ldots, \dim \boldsymbol{\theta} = m. \tag{2}$$

При этом точки $\{\mathbf{t}_j\}$ выбираются так, чтобы эта система была разрешима единственным образом и обратное отображение было бы непрерывным. Здесь мы не планируем обсуждение вопроса оптимального выбора точек $\{\mathbf{t}_j\}$.

Касательно предположения (II) приведем примеры непараметрических ядерных оценок, которые будут $\|\cdot\|_{pw}$-состоятельными при выполнении лишь условия (D).

Прежде всего определим сглаживающую ядерную функцию $K(\mathbf{t})$, $\mathbf{t}\in\mathbb{R}^k$, как плотность центрально-симметричного распределения с носителем в кубе $[-1,1]^k$. Будем предполагать, что функция $K(\mathbf{t})$ удовлетворяет условию Липшица всюду на $\mathbb{R}^k$. Нам также потребуется обозначение $K_h(\mathbf{t}) = h^{-k}K(h^{-1}\mathbf{t})$, $h\in(0,1)$, для плотности распределения с носителем в $[-h,h]^k$.

Наконец, оценку для функции f_{θ_0}, участвующую в условии (II), определим равенством

$$f^*_{n,h}(\mathbf{t}) = \frac{\sum_{i=1}^n X_i K_h(\mathbf{t}-\mathbf{z}_i)\Lambda_k(\mathcal{P}_i)}{\sum_{i=1}^n K_h(\mathbf{t}-\mathbf{z}_i)\Lambda_k(\mathcal{P}_i)}, \tag{3}$$

где h – так называемый размер окна наблюдений, который с ростом n стремится к нулю с определенной скоростью, $\Lambda_k(\cdot)$ — мера Лебега в $\mathbb{R}^k$, измеримые подмножества $\{\mathcal{P}_i, i=1,\ldots,n\}$ образуют конечное разбиение куба $[0,1]^k$, при этом каждое из этих подмножеств содержит ровно по одной регрессионной точке из набора $\{\mathbf{z}_i\}$ и максимальный диаметр всех элементов разбиения $\{\mathcal{P}_i\}$ должен стремиться к нулю с ростом объема наблюдений n (при выполнении условия (D) такое разбиение, очевидно, существует, см. [17]). С практической точки зрения указанное разбиение с отмеченными точками можно организовать, например, методом последовательных покоординатно-медианных сечений или с помощью мозаики Вороного (см. подробности в [17]). Приведем одну из одномерных версий этой оценки [2]:

$$f^*_{n,h}(t) = \frac{\sum_{i=1}^n X_{ni} K_h(t-z_{n:i})\Delta z_{ni}}{\sum_{i=1}^n K_h(t-z_{n:i})\Delta z_{ni}}, \tag{4}$$

где $z_{n:0}=0$, $z_{n:1}\le\ldots\le z_{n:n}$ — элементы вариационного ряда, построенного по выборке $\{z_i;\ i=1,\ldots,n\}$, $\Delta z_{ni} = z_{n:i}-z_{n:i-1}$, $i=1,\ldots,n$; величины X_{ni} — отклики из регрессионного уравнения (1), ассоциированные с порядковой статистикой $z_{n:i}$.

Введем также в рассмотрение классическую оценку Надарая–Ватсона

$$\widehat{f}_{n,h}(\mathbf{t}) = \frac{\sum_{i=1}^n X_i K_h(\mathbf{t}-\mathbf{z}_i)}{\sum_{i=1}^n K_h(\mathbf{t}-\mathbf{z}_i)}. \tag{5}$$

Известно (см. [17], [2], [18], [12]), что при выполнении лишь условия (D) существует последовательность $h \equiv h_n \to 0$, для которой вышеприведенные три оценки будут $\|\cdot\|_{pw}$-состоятельными, а оценки (3) и (4) — еще и $\|\cdot\|_{sup}$-состоятельными. Далее по умолчанию предполагается, что при использовании ядерных оценок (3) и (5) размер окна $h \equiv h_n$ с ростом n стремится к нулю с известной скоростью. В параграфе 3 настоящей работы формула (11) определяет оптимальный размер окна для ядерной оценки (3) (см. также нижеследующее замечание 5).

Замечание 3. Заметим, что имеют место соотношения

$$f^*_{n,h}(\mathbf{t}) = \arg\min_a \sum_{i=1}^n (X_i - a)^2 K_h(\mathbf{t} - \mathbf{z}_i)\Lambda(\mathcal{P}_i),$$
$$\widehat{f}_{n,h}(\mathbf{t}) = \arg\min_a \sum_{i=1}^n (X_i - a)^2 K_h(\mathbf{t} - \mathbf{z}_i),$$

т.е. ядерная оценка $f^*_{n,h}(t)$, как и классическая оценка Надарая-Ватсона $\widehat{f}_{n,h}(t)$, является оценкой взвешенного метода наименьших квадратов и принадлежит классу локально-постоянных оценок, но с некоторыми иными весами, нежели в классическом варианте оценок Надарая–Ватсона.

Приведем примеры оценок, построенных с помощью вышеприведенной теоремы.

Пример 1. Рассмотрим модель (1) с регрессионной функцией

$$f_\theta(\mathbf{t}) = \theta_1 t_1^{\theta_2} t_2^{\theta_3},$$

где $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2_+$ и $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3_+$ — векторы с положительными координатами. Это так называемая модель Кобба–Дугласа, достаточно популярная в эконометрике (см., например, [10]). Рассмотрим отображение из $(C[0,1]^2, \|\cdot\|_{pw})$ в $\mathbb{R}^3_+$

$$G(f) = (f(2^{-1}, 2^{-1}), f(2^{-1}, 3^{-1}), f(3^{-1}, 3^{-1})),$$

которое, очевидно, будет непрерывным. Теперь покажем, что суперпозиция $\mathbf{g}(\theta) \equiv G(f_\theta)$ осуществляет гомеоморфизм $\mathbb{R}^3_+$ в коническую область $C^3_+ = \{(r_1, r_2, r_3) \in \mathbb{R}^3_+ : 0 < r_3 < r_2 < r_1\}$. Нам достаточно доказать, что $\mathbf{g}(\boldsymbol{\theta})$ — биекция, так как непрерывность отображения очевидна. В самом деле, рассмотрим систему уравнений

$$\begin{cases} f_\theta(2^{-1}, 2^{-1}) \equiv \theta_1 2^{-\theta_2} 2^{-\theta_3} = r_1, \\ f_\theta(2^{-1}, 3^{-1}) \equiv \theta_1 2^{-\theta_2} 3^{-\theta_3} = r_2, \\ f_\theta(3^{-1}, 3^{-1}) \equiv \theta_1 3^{-\theta_2} 3^{-\theta_3} = r_3, \end{cases}$$

где (r_1, r_2, r_3) – произвольная точка из C^3_+. Логарифмируя эту систему уравнений, приводим ее к эквивалентной форме

$$\begin{cases} \widetilde{\theta}_1 - \theta_2 \log 2 - \theta_3 \log 2 = s_1, \\ \widetilde{\theta}_1 - \theta_2 \log 2 - \theta_3 \log 3 = s_2, \\ \widetilde{\theta}_1 - \theta_2 \log 3 - \theta_3 \log 3 = s_3, \end{cases}$$

где $\widetilde{\theta}_1 = \log \theta_1$, $s_j = \log r_j$, $j = 1, 2, 3$. Легко проверить, что матрица $\mathbb{A}$ этой системы линейных уравнений невырождена. Единственное ее решение легко находится методом последовательного исключения переменных:

$$\widetilde{\theta}^*_{n1} = \frac{s_1 \log 3 - s_3 \log 2}{\log(3/2)}, \quad \theta^*_{n2} = \frac{s_1 - s_2}{\log(3/2)}, \quad \theta^*_{n3} = \frac{s_2 - s_3}{\log(3/2)}. \tag{6}$$

Теперь для наших целей можно использовать, например, ядерную оценку $f^*_{n,h}(\mathbf{t})$, определенную в (3). Положим в (6)

$$s_1 = \log f^*_{n,h}(2^{-1}, 2^{-1}), \quad s_2 = \log f^*_{n,h}(2^{-1}, 3^{-1}), \quad s_3 = \log f^*_{n,h}(3^{-1}, 3^{-1}).$$

В силу сказанного на множестве элементарных исходов асимптотически полной меры для всех достаточно больших n выполнено двойное неравенство $s_1 > s_2 > s_3$, т.е. трехмерная оценка $(\widetilde{\theta}^*_{n1}, \theta^*_{n2}, \theta^*_{n3})$ корректно задана, и на основании теоремы 1 она будет состоятельной для трехмерного параметра $(\log \theta_1, \theta_2, \theta_3)$. Тем самым, оценка $\boldsymbol{\theta}^*_n = (\exp\{\widetilde{\theta}^*_{n1}\}, \theta^*_{n2}, \theta^*_{n3})$ будет состоятельной для исходного параметра $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$.

Пример 2. Рассмотрим модель (1) с регрессионной функцией

$$f_\theta(t) = \frac{\theta_1 t}{t + \theta_2},$$

где $t > 0$ и $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \mathbb{R}^2_+$. Это так называемая модель Михаэлиса–Ментен, хорошо известная в биохимии (см., например, [5], [6]). Рассмотрим непрерывное отображение из $(C[0,1], \|\cdot\|_{pw})$ в $\mathbb{R}^2_+$

$$\mathbf{G}(f) = (f(1), f(1/2)).$$

Теперь покажем, что суперпозиция $\mathbf{g}(\boldsymbol{\theta}) \equiv \mathbf{G}(f_\theta)$ — это гомеоморфизм открытого положительного квадранта $\mathbb{R}^2_+$ в открытый конус $C^2_+ = \{(r_1, r_2);\ r_2 > 0,\ r_2 < r_1 < 2r_2\}$. Нам достаточно доказать, что $\mathbf{g}(\boldsymbol{\theta})$ — биекция. В самом деле, рассмотрим систему уравнений

$$\begin{cases} \frac{\theta_1}{1+\theta_2} = r_1, \\ \frac{2^{-1}\theta_1}{2^{-1}+\theta_2} = r_2 \end{cases}$$

при любом векторе $(r_1, r_2) \in C_+^2$. Эта система очевидным образом сводится к следующей системе двух линейных уравнений с двумя неизвестными θ_1 и θ_2:

$$\begin{cases} \theta_1 - r_1\theta_2 = r_1, \\ \theta_1 - 2r_2\theta_2 = r_2, \end{cases}$$

матрица которой невырождена всюду в вышеуказанном открытом конусе. В результате при известном векторе (r_1, r_2) получаем

$$\theta_{n1}^* = r_1\left(1 + \frac{r_1 - r_2}{2r_2 - r_1}\right), \quad \theta_{n2}^* = \frac{r_1 - r_2}{2r_2 - r_1}. \tag{7}$$

Очевидно, построенное взаимно-однозначное отображение двусторонне непрерывно, т.е. является гомеоморфизмом указанных выше областей.

Теперь можно рассмотреть ядерную оценку $f_{n,h}^*(t)$, определенную в (4), или оценку Надарая–Ватсона $\widehat{f}_{n,h}(t)$, определенную в (5). Как уже было доказано в [2] и [18], при выполнении условия (D) обе эти оценки будут состоятельными в норме $\|\cdot\|_{pw}$ для некоторого $h \equiv h_n \to 0$. Так что, например, можно положить в (7)

$$r_1 = f_{n,h}^*(t)(1), \quad r_2 = f_{n,h}^*(t)(1/2).$$

В этом случае двумерная оценка $\boldsymbol{\theta}_n^* = (\theta_{n1}^*, \theta_{n2}^*)$ в (7) будет корректно определена на множестве элементарных исходов асимптотически полной меры с ростом n и, более того, будет состоятельной в силу теоремы 1.

Замечание 4. Отметим, что явные оценки для параметров модели Михаэлиса–Ментен были известны и ранее (см. [22]). В частности, оценки в работе [22] построены, по сути, за счет внутренней линейности этой модели и выполнение условия типа (D) в [22] не требуется. Однако принципиальным отличием этой работы от результатов настоящей заметки состоит в том, что в [22] регрессоры $\{\mathbf{z}_i\}$ неслучайны, а случайные погрешности $\{\xi_i\}$ независимы.

Пример 3. Рассмотрим так называемую логистическую регрессию. В этом случае

$$f_\theta(\mathbf{t}) = \left(1 + e^{-(\mathbf{t}, \boldsymbol{\theta})}\right)^{-1},$$

где $\dim \mathbf{t} = \dim \boldsymbol{\theta} = m$, а $(\cdot, \cdot)$ – стандартное евклидово скалярное произведение в $\mathbb{R}^m$.

Далее, для произвольного набора точек $\{\mathbf{t}_j;\ j = 1, \ldots, m\}$ из единичного m-мерного куба система уравнений (2) очевидным образом приводится к системе линейных уравнений вида

$$\begin{cases} (\mathbf{t}_1, \boldsymbol{\theta}) = r_1, \\ \dots\dots\dots\dots \\ (\mathbf{t}_m, \boldsymbol{\theta}) = r_m. \end{cases}$$

Единственное ограничение на векторы $\{\mathbf{t}_j\}$ — их линейная независимость, т.е. невырожденность матрицы $\mathbb{T}$ приведенной системы. Теперь положим

$$r_j = \log \frac{f_n^*(\mathbf{t}_j)}{1 - f_n^*(\mathbf{t}_j)},$$

где $f_n^*(\mathbf{t})$ — любая из оценок (3) или (5). Отметим, что в силу $\|\cdot\|_{pw}$-состоятельности указанных оценок при выполнении условия (D) имеет место неравенство $f_n^*(\mathbf{t}_j) < 1$ с вероятностью, стремящейся к 1 с ростом n для любого фиксированного $\mathbf{t}_j$. Если указанное неравенство нарушено, полагаем $r_j = 0$. Окончательно получаем следующую конструкцию состоятельной оценки для логистической модели:

$$\boldsymbol{\theta}_n^* = \mathbb{T}^{-1} \cdot (r_1, \dots, r_m)^\top. \tag{8}$$

§ 3. Анализ α_n-состоятельности оценок

Теперь обсудим вопрос об уточнении результатов теоремы 1. Напомним, что оценка $\boldsymbol{\theta}_n^*$ параметра $\boldsymbol{\theta} \in \mathbb{R}^m$ называется α_n-состоятельной, если $\alpha_n(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}_0) \xrightarrow{p} \mathbf{0}$ и $\alpha_n \to \infty$ при $n \to \infty$. Это определение без труда переносится на случай бесконечномерных параметров.

Определение. Пусть $(\mathcal{C}, d)$ – сепарабельное метрическое пространство. Последовательность случайных элементов $g_n^* \in \mathcal{C}$ называется α_n-состоятельной оценкой случайного элемента $g \in \mathcal{C}$, если $\alpha_n d(g_n^*, g) \xrightarrow{p} 0$ и $\alpha_n \to \infty$ при $n \to \infty$.

Приведем теперь уточнение теоремы 1 при соответствующей детализации ограничений на входящие в нее параметры.

Теорема 2. *Пусть в условиях теоремы* 1 *дополнительно предполагается, что отображения* $\mathbf{G}$ *и* $\mathbf{g}^{-1}$ *удовлетворяют условию Липшица в своих пространствах и существует непараметрическая* α_n*-состоятельная* (*в норме линейного пространства* $(C[0,1]^k, \|\cdot\|)$) *оценка* f_n^* *для неизвестной регрессионной функции* f_{θ_0}*. Тогда оценка* $\boldsymbol{\theta}_n^*$ *будет* α_n*-состоятельной для параметра* $\boldsymbol{\theta}_0$.

Доказательство сразу следует из представления оценки $\boldsymbol{\theta}_n^* = \mathbf{g}^{-1}(\mathbf{G}(f_n^*))$ в теореме 1 и вышеприведенного определения α_n-состоятельной непараметрической оценки f_n^*. $\square$

Как нетрудно видеть, основное условие в теореме 2 — это существование α_n-состоятельной непараметрической оценки f_n^*. Следующее утверждение дает пример такой непараметрической оценки.

Теорема 3. *Пусть в модели* (1) *для каждого* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *регрессионная функция* $f_\theta(\mathbf{t})$ *удовлетворяет условию Липшица. Тогда при выполнении условия* (D) *ядерная оценка* (3) *будет* α_n*-состоятельной, если только* $\alpha_n = o(h_n^{-1})$, *где*

$$h_n = \left(\mathbb{E}(\varepsilon_n^{kp/2})\right)^{\frac{1}{p(k/2+1)+k}}.$$

Доказательство. В [17] показано, что в условиях рассматриваемой модели (1) для ядерной оценки (3) справедливо неравенство

$$\|f_{n,h}^* - f_{\theta_0}\|_{sup} \le \omega_{f_{\theta_0}}(h) + \zeta_n(h) \tag{9}$$

почти наверное, где $\omega_{f_{\theta_0}}(h)$ — модуль непрерывности регрессионной функции, а случайная величина $\zeta_n(h)$ имеет порядок

$$\zeta_n(h) = O_p\left(\left(h^{-k(p/2+1)}\,\mathbb{E}(\varepsilon_n^{kp/2})\right)^{1/p}\right);$$

здесь верхний предел в определении символа $O_p(\cdot)$ зависит от m, p, M_p и ядра K. Величина $h \equiv h_n$, выравнивающая порядки обоих слагаемых в правой части (9), находится как решение $h \equiv h_n$ уравнения

$$\mathbb{E}(\varepsilon_n^{kp/2}) = h^{k(p/2+1)}\omega_{f_{\theta_0}}^p(h). \tag{10}$$

Для липшицевой регрессионной функции в этом случае получаем представление для оптимального (по порядку малости) размера окна:

$$h_n = \left(\mathbb{E}(\varepsilon_n^{kp/2})\right)^{\frac{1}{p(k/2+1)+k}}, \tag{11}$$

что и требовалось показать. □

Таким образом, если положить $\alpha_n = o(h_n^{-1})$, то во всех вышеприведенных примерах при подстановке в полученные формулы непараметрической ядерной оценки f_{n,h_n}^* мы с помощью теорем 2 и 3 получим α_n-состоятельные оценки для многомерных параметров рассматриваемых моделей нелинейной регрессии. Например, для оценки (8) это утверждение следует из того факта, что функция $\phi(z) = \log\frac{z}{1-z}$ в окрестности каждой точки открытого интервала $(0,1)$ удовлетворяет условию Липшица, и, кроме того, в силу (8) на множестве элементарных исходов асимптотически полной меры при всех достаточно больших n имеет место неравенство

$$\|\boldsymbol{\theta}_n^* - \boldsymbol{\theta}_0\| \le C\max_{j\le m}|f_n^*(\mathbf{t}_j) - f_{\theta_0}(\mathbf{t}_j)|, \tag{12}$$

где неслучайная постоянная C зависит от набора $\{\mathbf{t}_j\}$ и m.

Замечание 5. Если минимальный радиус ε-сети в условии (D) допускает детерминированную оценку сверху $\varepsilon_n \le \hat{\varepsilon}_n$, то формула (11) для размера окна может быть преобразована как

$$h_n \le \widehat{h}_n = \hat{\varepsilon}_n^{\frac{k}{k+2+2/p}},$$

при этом при достаточно большом p (например, когда шум гауссов) порядок малости размера окна можно сделать сколь угодно близким к величине $\hat{\varepsilon}_n^{k/(k+2)}$. В качестве примера можно рассмотреть одномерный ($k = 1$) эквидистантный план эксперимента, когда в оценке (4) нужно положить $z_{n:i} = i/n$ и $\hat{\varepsilon}_n = \Delta z_{ni} = 1/n$. При этом $\hat{h}_n = \hat{\varepsilon}_n^{1/3} = n^{-1/3}$. Стоит также отметить, что в случае независимых одинаково распределенных случайных регрессоров $\{\mathbf{z}_i;\ i = 1, \dots, n\}$ в модели (1) нетрудно установить факт асимптотической нормальности рассмотренных выше оценок, используя известные результаты об асимптотической нормальности ядерных оценок (3)–(5). Например, при $k = 1$ нормировка невязки $|f_n^*(\mathbf{t}_j) - f_{\theta_0}(\mathbf{t}_j)|$ в широких условиях будет иметь порядок $\sqrt{nh_n}$ для некоторой стремящейся к нулю со степенной скоростью по n последовательности h_n (размер окна), что является верхней границей (недостижимой) для нормирующего множителя α_n в определении α_n-состоятельности.

Замечание 6. Предложенный здесь подход к построению α_n-состоятельных оценок параметров нелинейной регрессии несколько проигрывает в точности оценкам работ [16], [19], где был рассмотрен иной метод построения α_n-состоятельных и асимптотически нормальных оценок без использования ядерных непараметрических оценок. В широких условиях величина α_n в упомянутых работах может быть сколь угодно близкой к $\sqrt{n}$, в то время как с использованием оценки (12) возможно получить α_n-состоятельные оценки с порядком роста α_n не более, чем $n^{1/3}$ (см. замечание 5). Однако явные оценки, полученные методом теоремы 1, во многих моделях выглядят значительно проще, нежели оценки в этих же моделях, предложенные в [16], [19]. Кроме того, важно отметить, что с помощью одношаговых процедур точность тех или иных оценок может быть существенно улучшена, при этом в качестве предварительных в широких условиях могут выступать $n^{1/4}$-состоятельные оценки (см., например, [13], [14], [15]). Так что оценки, предложенные в настоящей работе, могут быть использованы как предварительные (начальные) для многошаговых процедур ньютоновского типа.

Список литературы

1. *Bates D.M., Watts D.G. Nonlinear regression analysis and its applications.* New York: Wiley, 1988.

2. *Borisov I.S., Linke Yu.Yu. and Ruzankin P.S.* Universal weighted kernel-type estimators for some class of regression models // *Metrika.* 2021. V. 84, N 2. P. 141–166.

3. *Chatterjee S., Hadi A.S. Regression analysis by example.* Hoboken: Wiley, 2006.

4. *Chung I.H. and Kim K.H.* Asymptotic properties of the one-step M-estimators in nonlinear regression model // *Commun. Korean Math. Soc.* 1992. V. 7, N 2. P. 293-306.

5. *Dette H., Melas V. and Pepelyshev A.* Standardized E-optimal designs for the Michaelis–Menten model // *Statist. Sinica.* 2003. V. 13, P. 1147–1163.

6. *Dette H., Melas V. and Wong W.K.* Optimal design for goodness-of-fit of the Michaelis–Menten enzyme kinetic function // *J. Amer. Statist. Assoc.* 2005. V. 100, N. 472, P. 1370–1381.

7. *Draper N.R., and Smith H. Applied regression analysis.* New York: Wiley, 1998.

8. *Freund R.J., Wilson W.J., and Sa P. Regression analysis: statistical modeling of a response variable.* Cambridge: Academic Press, 2006.

9. *Ghilagaber G., Midi H., and Riazoshams H. Robust nonlinear regression: with applications using R.* New York: Wiley, 2019.

10. *Grigoriev Yu. D. Melas V. B. and Shpilev P.V.* Excess of locally D-optimal designs for Cobb–Douglas model // *Statist. Papers* 2018. V.259. P. 1425–1439.

11. *Heyde C.C. Quasi-likelihood and its application: a general approach to optimal parameter estimation.* Springer, 1997.

12. *Linke Yu.Yu.* Towards insensitivity of Nadaraya–Watson estimators to design correlation // *Theory Probab. Appl.* 2023. V. 68, N. 2. — P .198–210.

13. Linke Yu. Yu. Asymptotic properties of one-step weighted M-estimators with application to some regression problems // *Theory Probab. Appl.* 2018. V. 62, N. 3. P. 373–398.

14. *Linke Yu.Yu.* Asymptotic properties of one-step M-estimators // *Communications in Statistics — Theory and Methods.* 2019. V. 48. N 16. P. 4096–4118.

15. *Linke Yu.Yu.* Asymptotic normality of one-step M-estimators based on non-identically distributed observations // *Statist. Probab. Lett.* 2017. V. 129, P. 216–221.

16. *Linke Yu.Yu., Borisov I.S.* Constructing explicit estimators in nonlinear regression models // *Theory Probab. Appl.* 2018. V. 63, N. 1. P .22–44.

17. *Linke Yu.Yu., Borisov I.S. and Ruzankin P.S.* Universal kernel-type estimation of random fields. // *Statistics.* 2023. V. 57, N. 4. P. 785–810.

18. *Linke Yu.Yu. and Borisov I.S.* Insensitivity of Nadaraya–Watson estimators to design correlation // *Communications in Statistics – Theory and Methods.* 2022. V. 51, N. 19. P. 6909–6918.

19. *Linke Yu.Yu. and Borisov I.S.* Constructing initial estimators in one-step estimation procedures of nonlinear regression // *Statist. Probab. Lett.* 2017. V. 120, P. 87–94.

20. *Linke Yu.Yu. and Borisov I.S.* Toward the notion of intrinsically linear models in nonlinear regression // *Siberian Adv. Math.* 2019. V. 29, N. 3. P. 210–216.

21. *Linke Yu. Yu. and Sakhanenko A.I.* Asymptotically normal estimation of a multidimensional parameter in the linear-fractional regression problem. // *Sibirsk. Mat. Zh.* 2001. V. 42, N. 2. P. 372–388.

22. *Linke Yu. Yu. and Sakhanenko A.I.* Asymptotically normal explicit estimation of parameters in the Michaelis–Menten equation // *Sibirsk. Mat. Zh.* 2001. V. 42, N. 3. P. 610–633.

23. *Mu B., Bai E.-W., Zheng W.X. and Zhu Q.* A globally consistent nonlinear least squares estimator for identification of nonlinear rational systems // *Automatica J. IFAC.* 2017. V. 77. P. 322–335.

24. *Müller Ch. H.* Asymptotic behaviour of one-step-M-estimators in contaminated nonlinear models.//*Asymptotic Statistics, Physica.* Heidelberg: Verlag, 1994. P.395–404.

25. *Panik M.J. Regression modeling: methods, theory, and computation with SAS.* Boca Raton: CRC Press, 2009.

26. *Poetscher B.M. and Prucha I.R.* A class of partially adaptive one-step M-estimators for the non-linear regression model with dependent observations // *J. Econometrics.* 1986. V. 32. P. 219–251.

27. *Rawlings J.O., Pantula S.G., Dickey D.A. Applied regression analysis: a research tool.* Springer, 2001.

28. *Sakhanenko A. I.* On existence of explicit asymptotically normal estimators in nonlinear regression problems // *Analytical Methods in Statistics (AMISTAT 2015)*/ Springer Proceed. in Math. and Statist. 193, ed. J. Antoch, J. Jureckova, M. Maciak, M. Pesta. Springer, 2017. P. 159–187.

29. *Savinkina E. N. and Sakhanenko A. I.* On improvement of statistical estimators in a power regression problem // *Sib. Electron. Mat. Izv.* 2019. V. 16. P. 1901–1912. [Russian, English abstract]

30. *Shults J. and Hilbe J.M. Quasi-least squares regression.* CRC Press, 2014.

31. *Small C.G., and Wang J. Numerical methods for nonlinear estimating equations.* — Oxford: Clarendon press, 2003.

32. *Small C.G., Wang J., Yang Z.* Eliminating multiple root problems in estimation // *Statistical Science.* 2000. V. 15, N. 4. P.313–341.

33. *Small C.G. and Yang Z.* Multiple roots of estimating functions // *Canad. J. Statist.* 1999. V.27, N. 3. P.585–598.

34. *Seber G.A.F., and Wild C.J. Nonlinear regression.* New York: Wiley, 2003.

35. *Van der Vaart A. W. Asymptotic statistics.* Cambridge University Press, and 2000.

36. *Wakefield J. Bayesian and frequentist regression methods.* Springer, 2013.

37. *Young D.S. Handbook of regression methods.* Chapman and Hall, 2017.

Линке Юлиана Юрьевна
Борисов Игорь Семёнович
Институт математики
им. С.Л.Соболева СОРАН,
просп. Академика Коптюга, 4,
Новосибирск, 630090 РОССИЯ.
E-mail: linke@math.nsc.ru
E-mail: sibam@math.nsc.ru

Поступила в редакцию
13 июля 2023 г.
Получена после доработки
1 октября 2023 г.
Принята к публикации
5 октября 2023 г.