ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2019 Управление, вычислительная техника и информатика № 49

УДК 519.2
DOI: 10.17223/19988605/49/9

Е.А. Пчелинцев, С.С. Перелевский

АДАПТИВНОЕ ЭФФЕКТИВНОЕ ОЦЕНИВАНИЕ ФУНКЦИИ В ГЕТЕРОСКЕДАСТИЧНОЙ РЕГРЕССИИ

Работа выполнена при финансовой поддержке Российского научного фонда, грант № 17-11-01049. Работа второго автора частично поддержана Министерством науки и высшего образования РФ (государственное задание № 2.3208.2017/4.6).

Изучаются асимптотические свойства адаптивной улучшенной процедуры выбора модели для оценивания неизвестной функции в гетероскедастичной регрессии. Установлено, что процедура является асимптотически эффективной в смысле среднеквадратического риска, т.е. асимптотический среднеквадратический риск процедуры совпадает с соответствующей константой Пинскера, обеспечивающей точную нижнюю границу риска по всем возможным оценкам. Приводятся результаты численного моделирования.
Ключевые слова: непараметрическая гетероскедастичная регрессия; улучшенное оценивание; среднеквадратический риск; оракульные неравенства; асимптотическая эффективность.

Рассматривается задача адаптивного асимптотически эффективного оценивания в непараметрической гетероскедастичной регрессионной модели

$$y_j = S(x_j) + \sigma_j \xi_j\,,\ 1 \le j \le n\,, \tag{1}$$

где $x_j = j/n$ – точки разбиения сегмента [0, 1], S – неизвестная 1-периодическая функция, которую требуется оценить, $(\xi_j)_{1\le j\le n}$ – последовательность центрированных независимых одинаково распределенных случайных величин с единичной дисперсией, $(\sigma_j)_{1\le j\le n}$ – неизвестные коэффициенты волатильности, зависящие от точек x_j и функции S, n – число наблюдений.

Гетероскедастичные регрессионные модели широко применяются в эконометрических исследованиях, в частности при анализе инвестиционного поведения фирм, для задачи потребления и др. [1, 2]. Также регрессионная модель типа (1) используется в стохастических дифференциальных уравнениях для аппроксимации диффузионных процессов с непрерывным временем путем применения последовательных ядерных оценок [3, 4]. Важной является задача качественной статистической идентификации таких моделей [5].

Цель данной работы – доказать свойство асимптотической (при $n \to \infty$) эффективности улучшенной процедуры выбора модели, предложенной авторами в [6], для оценивания функции S в модели (1). Рассматривается адаптивная постановка задачи, т.е. в условиях отсутствия информации о гладкости неизвестной функции S.

Понятие асимптотической эффективности связано с оптимальной скоростью сходимости минимаксного риска, т.е. важным вопросом в результатах оптимальности является изучение точной асимптотики минимаксного риска [7]. Проблема асимптотического непараметрического оценивания в модели гетероскедастичной регрессии изучались в работах [8–10]. Для доказательства асимптотической эффективности процедуры необходимо показать, что ее асимптотический квадратичный риск совпадает с нижней границей, определяемой константой Пинскера [11]. В данной работе поставленная задача решается с использованием подхода, основанного на методах выбора модели и точных оракульных неравенств, разработанных в статьях [10, 12]. В отличие от указанных работ, в статье применяется метод улучшенного адаптивного непараметрического оценивания, предложенный в [6, 13].

Статья состоит из четырех разделов. В разд. 1 описывается построение адаптивной улучшенной процедуры выбора модели и формулируется точное оракульное неравенство для среднеквадратического риска процедуры оценивания. В разд. 2 доказывается теорема об асимптотической эффективности процедуры выбора модели в минимаксном смысле. В разд. 3 изучается задача оценивания непериодической функции в модели (1). В разд. 4 приводятся результаты численного моделирования.

1. Адаптивная процедура выбора модели. Оракульное неравенство

Для оценивания неизвестной функции S в модели (1) предлагается использовать семейство $(S^*_\lambda)_{\lambda\in\Lambda}$ улучшенных взвешенных оценок наименьших квадратов, определенное в [6], т.е.

$$S^*_\lambda(x)=\sum_{j=1}^{n}\lambda_j\theta^*_{j,n}\phi_j(x), \qquad (2)$$

где вектор весовых коэффициентов $\lambda=(\lambda_1,...,\lambda_n)$ принадлежит некоторому конечному множеству $\Lambda\subset[0,1]^n$ с $n\geq 3$, $(\phi_j)_{1\leq j\leq n}$ – система ортонормированных тригонометрических функций в $L_2[0,1]$,

$$\theta^*_{j,n}=\left(1-\frac{c_n}{|\tilde{\theta}_n|}\mathbf{1}_{\{1\leq j\leq d\}}\right)\hat{\theta}_{j,n},\ \ |\tilde{\theta}_n|^2=\sum_{j=1}^{d}\hat{\theta}^2_{j,n},\ \ \hat{\theta}_{j,n}=\frac{1}{n}\sum_{l=1}^{n}y_l\phi_j(x_l).$$

Здесь коэффициент $c_n\approx d/n$, $d\approx n^\delta/\ln n,\ 0<\delta<1$, $\mathbf{1}_{\mathrm{A}}$ – индикатор множества A. Заметим, что $c_n\to 0$ при $n\to\infty$. В данной работе используется множество весов Λ, введенное в статье [12], т.е.

$$\Lambda=\{\lambda_\alpha,\ \alpha\in\mathrm{A}\},\ \ \mathrm{A}=\{1,...k\}\times\{t_1,...,t_m\},$$

где $t_i=i\varepsilon$ и $m=\left[\varepsilon^{-2}\right]$ – целая часть числа ε^{-2}. Считаем, что параметры $k\geq 1$ и $0<\varepsilon<1$ – функции от n, т.е. $k=k_n$ и $\varepsilon=\varepsilon_n$ такие, что

$$\begin{cases}\lim\limits_{n\to\infty}k_n=+\infty, & \lim\limits_{n\to\infty}\dfrac{k_n}{\ln n}=0,\\ \lim\limits_{n\to\infty}\varepsilon_n=0, & \lim\limits_{n\to\infty}n^\delta\varepsilon_n=+\infty\end{cases}$$

для любого $\delta>0$. Например, можно взять для $n\geq 3$

$$\varepsilon_n=\frac{1}{\ln n},\ \ k_n=\bar{k}+\sqrt{\ln n},$$

где $\bar{k}$ – некоторая неотрицательная постоянная. Для любого $\alpha=(\beta,t)\in\mathrm{A}$ положим, что весовой вектор $\lambda_\alpha=(\lambda_{\alpha,1},...,\lambda_{\alpha,n})$ с компонентами

$$\lambda_{\alpha,j}=\mathbf{1}_{\{1\leq j\leq d\}}+(1-(j/\omega_\alpha)^\beta)\mathbf{1}_{\{d<j\leq\omega_\alpha\}}.$$

Здесь $d=d(n)=[\omega_\alpha/\ln n]$,

$$\omega_\alpha=\varpi+(\kappa_\beta tn)^{1/(2\beta+1)},\ \ \kappa_\beta=(\beta+1)(2\beta+1)/(\pi^{2\beta}\beta),$$

и ϖ – некоторая неотрицательная постоянная.

Метод выбора модели заключается в описании правила выбора процедуры из семейства оценок (2), т.е. нужно выбрать вектор весовых коэффициентов $\lambda\in\Lambda$ в (2). Для этого естественно минимизировать эмпирическую квадратическую ошибку вида:

$$Err_n(\lambda)=\|S_\lambda^*-S\|_n^2=\frac{1}{n}\sum_{l=1}^{n}(S_\lambda^*(x_l)-S(x_l))^2=\sum_{j=1}^{n}\lambda_j^2\theta^{*2}_{j,n}-2\sum_{j=1}^{n}\lambda_j\theta^*_{j,n}\theta_j+\|S\|^2.$$

Однако поскольку во втором слагаемом коэффициенты θ_j неизвестны, то минимизация этого выражения не приводит к нахождению коэффициентов λ. Необходимо величины $\theta^*_{j,n}\theta_j$ заменить некоторыми оценками. В качестве таких оценок предлагаются

$$\overline{\theta}_{j,n} = \theta^*_{j,n}\hat{\theta}_{j,n} - \frac{1}{n}\hat{\varsigma}_n,$$

где $\hat{\varsigma}_n$ – некоторая оценка интегрированной дисперсии шума

$$\varsigma_n = \frac{1}{n}\sum_{j=1}^{n}\sigma_j^2.$$

При осуществлении этой замены в эмпирической ошибке нужно заплатить «штраф». Определим платежную функцию как

$$J_n(\lambda) = \sum_{j=1}^{n}\lambda_j^2\theta^{*2}_{j,n} - 2\sum_{j=1}^{n}\lambda_j\overline{\theta}_{j,n} + \rho\hat{P}_n(\lambda). \tag{3}$$

Здесь $0<\rho<1$ – некоторый положительный коэффициент, зависящий от n и такой, что $\rho(n)\to 0$ при $n\to\infty$ и пенализационное слагаемое

$$\hat{P}_n(\lambda) = \frac{|\lambda|^2\hat{\varsigma}_n}{n},\ |\lambda|^2 = \sum_{j=1}^{n}\lambda_j^2.$$

Отметим, что в случае, когда последовательность $(\sigma_j)_{1\le j\le n}$, известна, то

$$P_n(\lambda) = \frac{|\lambda|^2\varsigma_n}{n}.$$

Полагая теперь

$$\lambda^* = \operatorname*{arg\,min}_{\lambda\in\Lambda} J_n(\lambda),$$

определим процедуру выбора модели равенством

$$S^* = S^*_{\lambda^*}. \tag{4}$$

Заметим, что λ^* существует, поскольку множество Λ конечно. В случае, когда λ^* не единственное, берем любое из них.

Обозначим далее через $\mathbf{P}_n$ семейство распределений p в $\mathbf{R}^n$ векторов $(\xi_1,...,\xi_n)$ в модели (1), таких что компоненты ξ_j совместно независимы, центрированы с единичной дисперсией и

$$\max_{1\le k\le n}\mathbf{E}(\xi_k^4)\le l_n,$$

где последовательность $(l_n)_{n\ge 1}$ удовлетворяет условиям $\lim_{n\to\infty} l_n = \infty$, $\lim_{n\to\infty} l_n / n^{\delta} = 0$, для любого $\delta>0$. Легко видеть, что для любого $n\ge 1$ центрированное гауссовское распределение в $\mathbf{R}^n$ с единичной ковариационной матрицей принадлежит семейству $\mathbf{P}_n$. Для оценки S^* определим следующий робастный среднеквадратический риск

$$\mathfrak{R}_n(S^*,S) = \sup_{p\in\mathbf{P}_n}\mathbf{E}_{S,p}\left\|S^*-S\right\|_n^2, \tag{5}$$

где $\mathbf{E}_{S,p}$ – математическое ожидание относительно распределения $\mathbf{P}_{S,p}$ по наблюдениям $(y_1,...,y_n)$ с фиксированной функцией S и фиксированным распределением $p\in\mathbf{P}_n$ случайных величин $(\xi_j)_{1\le j\le n}$ в модели (1).

Замечание 1. Если оценка S^* определена только в точках $(x_j)_{1\le j\le n}$, то продолжаем ее ступенчатой функцией на сегмент [0, 1] следующим образом:

$$S^*(x) = S^*(x_1)\mathbf{1}_{[0,x_1]}(x) + \sum_{l=2}^{n} S^*(x_l)\mathbf{1}_{(x_{l-1},x_l]}(x) \text{ для любого } 0\le x\le 1.$$

Тогда риск, не зависящий от разбиения сегмента [0, 1], определим равенством

$$Q_n(S^*,S) = \sup_{p\in\mathbf{P}_n}\mathbf{E}_{S,p}\left\|S^*-S\right\|^2, \tag{6}$$

где $\|\cdot\|$ обозначает норму в $L_2[0,1]$.

В работе [6] установлено, что оценка (2) превосходит по среднеквадратической точности процедуру выбора модели, построенную на основе взвешенных оценок наименьших квадратов $\hat{S}_\lambda$ из [12], получено выражение для минимального выигрыша в точности, т.е.

$$\Re_n(S^*_\lambda, S) - \Re_n(\hat{S}_\lambda, S) \le -c_n^2. \qquad (7)$$

Также в работе [6] получено следующее оракульное неравенство, определяющее неасимптотическую точную верхнюю границу для среднеквадратического риска (5) процедуры оценивания (4).

Теорема 1. *Пусть наблюдения описываются уравнением* (1). *Тогда для любого* $0 < \rho < 1/3$ *среднеквадратический риск* (5), *предложенной процедуры выбора модели* (4) *для оценивания функции S удовлетворяет следующему оракульному неравенству*:

$$\Re_n(S^*, S) \le \frac{1-\rho}{1-2\rho} \min_{\lambda\in\Lambda} \Re_n(S^*_\lambda, S) + \frac{1}{n}\Psi_n(\rho),$$

где $\lim_{n\to\infty} \Psi_n(\rho)/n^\delta = 0$ *для любого* $\delta > 0$.

2. Асимптотическая эффективность

Пусть $C^k_{per,1}(\mathbf{R})$ – множество 1-периодических k раз дифференцируемых $\mathbf{R}\to\mathbf{R}$ функций и пусть функция S в модели (1) принадлежит следующему соболевскому шару:

$$W_r^k = \left\{ f \in C^k_{per,1}(\mathbf{R}) : \sum_{i=0}^k \left\| f^{(i)} \right\|^2 \le r \right\},$$

где $k \ge 1$ и $r > 0$ – неизвестные параметры.

Обозначим через Σ_n множество всех оценок $\hat{S}_n$, измеримых относительно наблюдений $(y_1,\ldots,y_n)$. Предположим, что коэффициенты волатильности $(\sigma_j)_{j\ge1}$ удовлетворяют следующим условиям.

C₁) $\sigma_j(S) = g(x_j, S)$ для некоторой неизвестной функции $g : [0,\,1] \times L_2[0,\,1] \to \mathbf{R}_+$, которая квадратично интегрируема по x и такая, что

$$\lim_{n\to\infty} \sup_{W_r^k} \left| \frac{1}{n} \sum_{j=1}^n g^2(x_j, S) - \varsigma(S) = 0 \right|,$$

где $\varsigma(S) = \int_0^1 g^2(x,S)dx$. Кроме того, $g_* = \inf_{0\le x\le1} \inf_{W_r^k} g^2(x,S) > 0$ и $\varsigma(S) < \infty$.

C₂) Для некоторого $x \in [0,\,1]$ оператор $g^2(x,\cdot) : C[0,\,1] \to \mathbf{R}$ дифференцируем в смысле Фреше при любой фиксированной функции f_0 из $C[0,\,1]$, т.е. для любой f из некоторой окрестности f_0 в $C[0,\,1]$

$$g^2(x,f) = g^2(x,f_0) + \mathrm{L}_{x,f_0}(f-f_0) + \Upsilon(x,f_0,f),$$

где производная по Фреше $\mathbf{L}_{x,f_0} : C[0,1] \to \mathbf{R}$ – ограниченный линейный оператор, а остаточный член $\Upsilon(x,f_0,f)$ для каждого $x \in [0,\,1]$, удовлетворяет следующему свойству:

$$\lim_{\|f-f_0\|_\infty\to0} \frac{\|\Upsilon(x,f_0,f)\|}{\|f-f_0\|_\infty} = 0, \quad \|f\|_\infty = \sup_{0\le t\le1}|f(t)|.$$

C₃) Существует некоторая положительная постоянная C^*, такая что для любой функции S из $C[0,1]$, оператор $\mathrm{L}_{x,S}$, определенный в **C₂**), удовлетворяет следующему неравенству для любой функции f из $C[0,1]$:

$$\left|\mathrm{L}_{x,S}(f)\right| \le C^*\left(|S(x)f(x)| + |f|_1 + \|S\|\|f\| \right), \quad |f|_1 = \int_0^1 |f(t)|dt.$$

C₄) Функция $g_0(\cdot) = g(\cdot, S_0)$, соответствующая $S_0 \equiv 0$, является непрерывной на интервале [0,1]. Кроме того,

$$\lim_{\delta\to 0} \sup_{0\le x\le 1} \sup_{\|S\|_\infty \le\delta} |g(x,S) - g(x,S_0)| = 0.$$

Замечание 2. Из условия **C**$_1$) следует, что функция g равномерно интегрируема по первой переменной, отделена от нуля и ограничена на классе W_r^k. Отделимость от нуля равносильна невырожденности распределения наблюдений, а ограниченность означает, что интенсивность шума конечна. Из условий **C**$_2$) и **C**$_3$) вытекает свойство регулярности функция g по второй переменной. Условие **C**$_4$) обеспечивает непрерывность функции в точке S_0.

Например, в качестве функции, удовлетворяющей перечисленным условиям **C**$_1$)–**C**$_4$), можно взять предложенную в [1]

$$g^2(x,S) = c_0 + c_1 x + c_2 S^2(x) + c_3 \int_0^1 S^2(t)dt,$$

где $c_0 > 0$, $c_i \ge 0$, $1 \le i \le 3$, – некоторые неизвестные постоянные. В этом случае

$$\varsigma(S) = c_0 + 0{,}5c_1 + (c_2 + c_3)\int_0^1 S^2(t)dt$$

и производная Фреше

$$\mathrm{L}_{x,S}(f) = 2S(x)f(x) + 2\int_0^1 S(t)f(t)dt.$$

Хорошо известно, что оптимальная скорость сходимости равна $n^{\frac{2k}{2k+1}}$, когда риск берется равномерно по классу W_r^k. Справедливы следующие асимптотические результаты.

Теорема 2. *Пусть для модели* (1) *выполнены условия* **C**$_2$)–**C**$_4$). *Тогда робастный риск* (5) *для всех* $k \ge 1$ *и* $r > 0$ *удовлетворяет следующему асимптотическому неравенству*:

$$\liminf_{n\to\infty} n^{\frac{2k}{2k+1}} \inf_{\hat S_n \in \Sigma_n} \sup_{S\in W_r^k} \frac{\mathfrak{R}_n(\hat S_n, S)}{\gamma_k(S)} \ge 1,$$

где $\gamma_k(S) = \Gamma_k^*(\varsigma(S))^{2k/(2k+1)}$, $\Gamma_k^* = ((2k+1)r)^{1/(2k+1)}(k/(\pi(k+1)))^{2k/(2k+1)}$ *– константа Пинскера.*

Доказательство этой теоремы аналогично доказательству Теоремы 4.3 из работы [10].

Теорема 3. *Пусть для модели* (1) *выполнено условие* **C**$_1$). *Тогда для робастного риска* (5) *процедуры выбора модели* S^* *всех* $k \ge 1$ *и* $r > 0$ *справедливо неравенство*

$$\limsup_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S\in W_r^k} \frac{\mathfrak{R}_n(S^*_{\lambda^*}, S)}{\gamma_k(S)} \le 1. \tag{8}$$

Доказательство. Для доказательства теоремы необходимо найти оценку из семейства (2), для которой возможно получить асимптотическую верхнюю границу (8). Положим

$$\tilde l_n = \min(\inf\{i \ge 1 : i\varepsilon \ge \bar r(S)\}, m),\ \bar r(S) = r/\varsigma(S)$$

и выберем $\tilde\alpha = (k, \tilde t_n) \in \mathrm{A}$, где $k \ge 1$ – неизвестный параметр множества W_r^k и $\tilde t_n = \tilde l_n \varepsilon$. Определим

$$\tilde S = S^*_{\tilde\lambda},\ \tilde\lambda = \lambda_{\tilde\alpha}.$$

Ясно, что эта оценка принадлежит семейству (2) и зависит от неизвестных параметров k, r и $\bar r(S)$, следовательно, не может быть явно вычислена и применяться для решения задачи оценивания в адаптивной постановке. Однако для риска этой оценки, принимая во внимание неравенство (7) и Теорему 5.1 из [10], имеем

$$\limsup_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S\in W_r^k} \frac{\mathfrak{R}_n(\tilde S, S)}{\gamma_k(S)} \le \limsup_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S\in W_r^k} \frac{\mathfrak{R}_n(\hat S_{\tilde\lambda}, S)}{\gamma_k(S)} \le 1. \tag{9}$$

Теперь, применяя оракульное неравенство из Теоремы 1 и полученную верхнюю границу (9), приходим к неравенству (8) для робастного риска процедуры (4). Теорема 3 доказана.

Следствие 1. *Из Теорем 2 и 3 следует, что процедура выбора модели* $S^*_{\lambda^*}$, *определенная в* (4), *является асимптотически эффективной, т.е.*

$$\lim_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S\in W_r^k} \frac{\Re(S^*_{\lambda^*},S)}{\gamma_k(S)} = 1.$$

Замечание 3. Утверждения Теорем 2 и 3 также справедливы для робастного риска (6), поскольку нетрудно видеть, что для любого $0<\delta<1$ и любой оценки $\hat{S}_n\in\Sigma_n$ функции $S\in W_r^k$ справедливо неравенство

$$\left\|\hat{S}_n - S\right\|_n^2 \geq (1-\delta)\left\|T(\hat{S}_n)-S\right\|^2 - (\delta^{-1}-1)r/n^2,$$

где $T(f)(x)=f(x_1)\mathbf{1}_{[0,x_1]}(x)+\sum_{l=2}^{n} f(x_l)\mathbf{1}_{(x_{l-1},x_l]}(x)$. Данное неравенство есть оценка погрешности от аппроксимации нормы в $L_2[0,1]$ эмпирической нормой.

3. Оценивание непериодической функции

Рассмотрим задачу оценивания непериодической функции S в модели (1). В этом случае будем оценивать функцию S на любом внутреннем сегменте $[a, b]$ из $[0,1]$, т.е. для $0 < a < b < 1$. Следует отметить, что в граничных точках $x = 0$ и $x = 1$ необходимо использовать ядерные оценки, предложенные в [14].

Пусть $\boldsymbol{\chi}:[0,1]\to\mathbf{R}^n_+$ – бесконечно дифференцируемая функция и такая, что $\chi(x) = 1$ для $a\leq x\leq b$ и $\chi^{(k)}(0)=\chi^{(k)}(1)=0$ для всех $k\geq 0$. Умножая уравнение (1) на функцию $\chi(\cdot)$ и моделируя последовательность $(\zeta_j)_{1\leq j\leq n}$ независимых стандартных нормальных случайных величин, приходим к задаче оценивания периодической функции $S_1(x)=S(x)\chi(x)$ по наблюдениям

$$\tilde{y}_j = S_1(x_j)+\tilde{\sigma}_j(S)\tilde{\xi}_j,\ 1\leq j\leq n,$$

где $\tilde{\sigma}_j(S)=\sqrt{\sigma_j^2(S)\chi^2(x)+\varepsilon^2}$,

$$\tilde{\xi}_j = \frac{\sigma_j(S)}{\tilde{\sigma}_j(S)}\xi_j + \frac{\varepsilon}{\tilde{\sigma}_j(S)}\zeta_j$$

и $\varepsilon>0$ – некоторый достаточно малый параметр. Заметим, что если последовательность $\left(\sigma_j(S)\right)_{1\leq j\leq n}$ удовлетворяет условиям **C₁)**–**C₄)**, то и последовательность $\left(\tilde{\sigma}_j(S)\right)_{1\leq j\leq n}$ удовлетворяет этим условиям с функцией $\tilde{g}(x,S)=\sqrt{g^2(x,S)\chi^2(x)+\varepsilon^2}$.

4. Численное моделирование

В этом разделе проиллюстрируем теоретически установленные результаты с помощью численного моделирования в среде MatLab. В качестве функции S в модели (1) выберем

$$S(x)=x\sin(2\pi x)+x^2(1-x)\cos(4\pi x).$$

Последовательность $(\xi_j)_{1\leq j\leq n}$ – независимые стандартные гауссовские случайные величины, коэффициенты волатильности $\sigma_j^2 = 2+x_j+S^2(x_j)+\int_0^1 S^2(t)dt$. Для вычисления весовых коэффициентов $(\lambda_\alpha)_{\alpha\in A}$ положим для $n\geq 3$

$$k = 100 + \sqrt{\ln n},\ \varepsilon = \frac{1}{\ln n},\ m = [\ln^2 n],$$

$$\rho = \frac{1}{3 + \ln^2 n},\ \omega_\alpha = 100 + (A_\beta t n)^{\frac{1}{2\beta+1}}.$$

В таблице приведены результаты поведения эмпирических среднеквадратических рисков по N = 1 000 реализациям процедур выбора модели (4).

Нормированные эмпирические квадратические риски

n	1 001	2 001	10 001	20 001	40 001	50 001
$n^{\frac{2k}{2k+1}} \frac{R(S^*, S)}{\gamma_k(S)}$	4,0856	2,0009	1,0117	1,0020	1,0006	1,0002

Из таблицы видно, что с ростом числа наблюдений n нормированные эмпирические средне-квадратические риски стремятся к единице, что численно подтверждает Следствие 1.

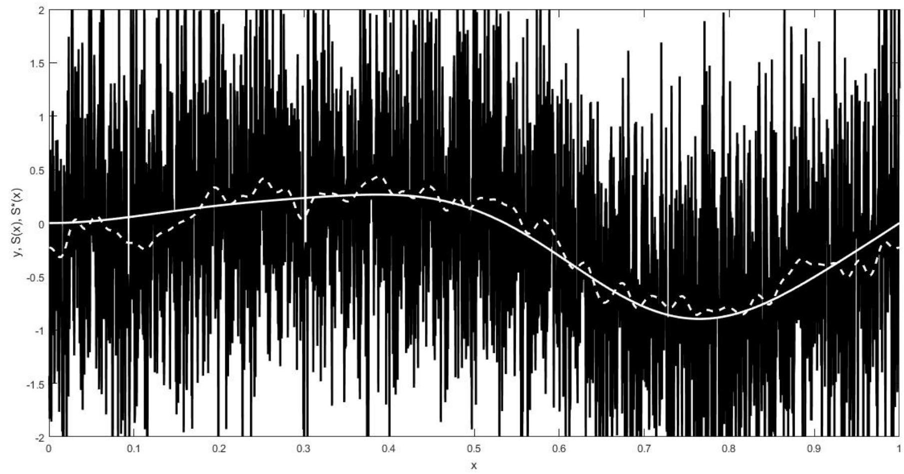

Рис. 1. Графики наблюдений $(y_1,..., y_n)$, функции регрессии S и ее оценки при n = 1 001

Fig. 1. Graphs of the observations $(y_1,..., y_n)$, regression function S and its estimate for n = 1 001

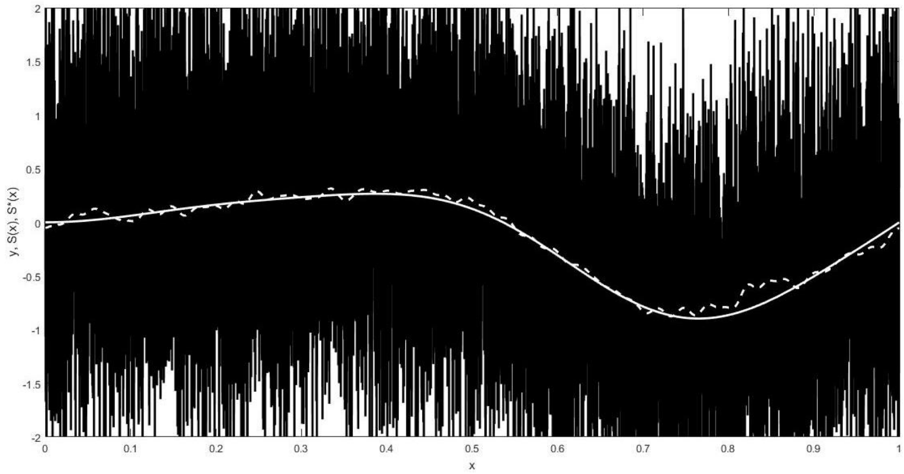

Рис. 2. Графики наблюдений $(y_1,..., y_n)$, функции регрессии S и ее оценки при n = 10 001

Fig. 2. Graphs of the observations $(y_1,..., y_n)$, regression function S and its estimate for n = 10 001

На рис. 1, 2 представлены графики наблюдений $(y_1,..., y_n)$, истинной функции регрессии S (сплошная белая линия) и ее оценки (4) (штрихованная белая линия).

Заключение

Отметим, что данная статья является продолжением исследований, представленных авторами в работе [6], и завершает решение задачи улучшенного адаптивного оценивания в непараметрической гетероскедастичной регрессионной модели (1) в смысле среднеквадратической точности. Установлено, что предложенная улучшенная процедура выбора модели является асимптотически эффективной. Проведенный численный анализ подтверждает работоспособность процедуры и достаточно быструю сходимость нормированных рисков к точной нижней границе, определяемой константой Пинскера. Результаты работы рекомендуется применять в задаче калибровки с целью повышения качества статистической идентификации.

ЛИТЕРАТУРА

1. Goldfeld S.M., Quandt R.E. Nonlinear Methods in Econometrics. London : North-Holland, 1972.
2. Cai T., Wang L. Adaptive variance function estimation in heteroscedactic nonparametric regression // Annals of Statistics. 2008. V. 36, No. 5. P. 2025–2054.
3. Galtchouk L., Pergamenshchikov S. Adaptive sequential estimation for ergodic diffusion processes in quadratic metric // Journal of Nonparametric Statistics. 2011. V. 23, No. 2. P. 255–285.
4. Pchelintsev E.A., Perelevskiy S.S., Makarova I.A. Improved nonparametric estimation of the drift in diffusion processes // Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki. 2018. Vol. 160, No 2. P. 364–372.
5. Belomestny D., Reiss M. Spectral calibration of exponential Levy models // Finance and Stochastics. 2006. Vol. 10, No 4. P. 449–474.
6. Пчелинцев Е.А., Перелевский С.С. Адаптивное оценивание в гетероскедастичной непараметрической регрессии // Вестник Томского государственного университета. Математика и механика. 2019. № 57. С. 38–52.
7. Ibragimov I.A., Hasminskii R.Z. Statistical Estimation: Asymptotic Theory. New York : Springer, 1981.
8. Efroimovich S. Sequential design and estimation in heteroscedastic nonparametric regression // Sequential Analysis. 2007. Vol. 26. P. 3–25.
9. Efroimovich S., Pinsker M. Sharp – optimal and adaptive estimation for heteroscedastic nonparametric regression // Statistica Sinica. 1996. Vol. 6. P. 925–942.
10. Galtchouk L., Pergamenshchikov S. Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression // Journal of the Korean Statistical Society. 2009. V. 38, No. 4. P. 305–322.
11. Pinsker M.S. Optimal filtration of square integrable signals in Gaussian white noise // Problems Transimis. information. 1981. N 17. P. 120–133.
12. Galtchouk L., Pergamenshchikov S. Sharp non-asymptotic oracle inequalities for nonparametric heteroscedastic regression models // Journal of Nonparametric Statistics. 2009. V. 21, No. 1. P. 1–16.
13. Pchelintsev E.A., Pchelintsev V.A., Pergamenshchikov S.M. Improved robust model selection methods for a Lévy nonparametric regression in continuous time // Journal of Nonparametric Statistics. 2019. P. 1–17. DOI: 10.1080/10485252.2019.1609672.
14. Brua J.-Y. Asymptotically efficient estimators for nonparametric heteroscedastic regression models // Statistical Methodologie. 2009. Vol. 6, No. 1. P. 47–60.

Поступила в редакцию 10 мая 2019 г.

Pchelintsev E.A., Perelevskiy S.S. (2019) ADAPTIVE EFFICIENT ESTIMATION FOR A FUNCTION IN HETEROSCEDASTIC REGRESSION. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitelnaja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 49. pp. 73–81

DOI: 10.17223/19988605/49/9

We consider the problem of adaptive asymptotically efficient estimation in a nonparametric heteroscedastic regression model

$$y_j = S(x_j) + \sigma_j \xi_j\,,\ 1 \le j \le n\,, \tag{1}$$

where the design points $x_j = j/n$, S is an unknown function to be estimated, $(\xi_j)_{1\le j\le n}$ is a sequence of centered independent random variables with unit variance and $(\sigma_j)_{1\le j\le n}$ are unknown scale functionals depending on the design points and the regression function S.

To estimate the unknown function S in model (1), it is proposed to use a model selection procedure based on improved weighted least squares estimates, defined in [6].

Denote via $\mathbf{P}_n$ a family of distributions p in $\mathbf{R}^n$ of the vectors $(\xi_1,...,\xi_n)$ in the model (1) such that $\max_{1\le k\le n}\mathbf{E}(\xi_k^4)\le l_n$, where the sequence l_n is such that $\lim_{n\to\infty} l_n=\infty$, $\lim_{n\to\infty} l_n/n^{\delta}=0$, for any $\delta>0$. For estimate S^* we define the following robust mean square risk

$$\Re_n(S^*,S)=\sup_{p\in\mathbf{P}_n}\mathbf{E}_{S,p}\left\|S^*-S\right\|_n^2,$$

where $\mathbf{E}_{S,p}$ is the expectation with respect to the distribution $\mathbf{P}_{S,p}$ of the observations $(y_1,...,y_n)$ with the fixed function S and the fixed distribution $p\in\mathbf{P}_n$ of random variables $(\xi_j)_{1\le j\le n}$ in the model (1). Under some additional conditions on the volatility coefficients in the model (1), the following asymptotic equality is proved:

$$\lim_{n\to\infty} n^{\frac{2k}{2k+1}}\sup_{S\in W_r^k}\frac{\Re(S_{\lambda^*}^*,S)}{\gamma_k(S)}=1,$$

which means that the proposed model selection procedure $S_{\lambda^*}^*$ for estimating the function S from the Sobolev ball W_r^k is asymptotically efficient. The results of the numerical analysis confirm the efficiency of the procedure and sufficiently fast convergence of the normalized risks to the sharp lower limit defined by the Pinsker constant $\gamma_k(S)$.

Keywords: nonparametric heteroscedastic regression; improved estimation; mean square risk; oracle inequalities, asymptotic efficiency.

PCHELINTSEV Evgeny Anatolevich (Candidate of Physical and Mathematical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation).
E–mail: evgen–pch@yandex.ru

PERELEVSKY Svyatoslav Sergeevich (Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation).
E–mail: slavaperelevskiy@mail.ru

REFERENCES

1. Goldfeld, S.M. & Quandt, R.E. (1972) *Nonlinear Methods in Econometrics*. London: North–Holland.
2. Cai, T. & Wang, L. (2008) Adaptive variance function estimation in heteroscedactic nonparametric regression. *Annals of Statistics*. 36(5). pp. 2025–2054.
3. Galtchouk, L. & Pergamenshchikov, S**.** (2011) Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. *Journal of Nonparametric Statistics*. 23(2). pp. 255–285. DOI: 10.1080/10485252.2010.544307
4. Pchelintsev, E.A., Perelevskiy, S.S. & Makarova, I.A. (2018) Improved nonparametric estimation of the drift in diffusion processes. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko–Matematicheskie Nauki – Proceedings of Kazan University. Physics and Mathematics Series.* 160(2). pp. 364–372.
5. Belomestny, D. & Reiss, M. (2006) Spectral calibration of exponential Levy models. *Finance and Stochastics*. 10(4). pp. 449–474. DOI: 10.1007/s00780-006-0021-5
6. Pchelintsev, E.A. & Perelevskiy, S.S. (2019) Adaptive estimation in a heteroscedastic nonparametric regression. *Vestnik Tomskogo gosudarstvennogo universiteta. Matematika i mekhanika – Tomsk State University Journal of Mathematics and Mechanics*. no. 57. pp. 38−52. DOI: 10.17223/19988621/57/3
7. Ibragimov, I.A. & Hasminskii, R.Z. (1981) *Statistical Estimation: Asymptotic Theory*. New York: Springer.
8. Efroimovich, S. (2007) Sequential design and estimation in heteroscedastic nonparametric regression. *Sequential Analysis*. 26. pp. 3–25. DOI: 10.1080/07474940601109670
9. Efroimovich, S. & Pinsker, M. (1996) Sharp – optimal and adaptive estimation for heteroscedastic nonparametric regression. *Statistica Sinica*. 6. pp. 925–942.
10. Galtchouk, L. & Pergamenshchikov, S. (2009) Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression. *Journal of the Korean Statistical Society*. 38(4). pp. 305–322. DOI: 10.1016/j.jkss.2008.12.001
11. Pinsker, M.S. (1981) Optimal filtration of square integrable signals in Gaussian white noise. *Problems of Information Transimission.* 17. pp. 120–133.
12. Galtchouk, L. & Pergamenshchikov, S. (2009) Sharp non–asymptotic oracle inequalities for nonparametric heteroscedastic regression models. *Journal of Nonparametric Statistics*. 21(1). pp. 1–16. DOI: 10.1080/10485250802504096
13. Pchelintsev, E.A., Pchelintsev, V.A. & Pergamenshchikov, S.M. (2019) Improved robust model selection methods for a Lévy nonparametric regression in continuous time. *Journal of Nonparametric Statistics*. 31(3). pp. 1–17. DOI: 10.1080/10485252.2019.1609672.
14. Brua, J.–Y. (2009) Asymptotically efficient estimators for nonparametric heteroscedastic regression models. *Statistical Methodologie*. 6(1). pp. 47–60. DOI: 10.1016/j.stamet.2008.02.009