ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 23 МАТЕМАТИКИ Выпуск 4
2016

С. С. Перелевский, Е. А. Пчелинцев (Томск, НИ ТГУ). **Улучшенная процедура выбора модели для оценивания функции регрессии по дискретным данным.**

В данной работе предлагается робастная адаптивная процедура оценивания неизвестной функции в непрерывной гетероскедастичной регрессионной модели. Доказывается оракульное неравенство для среднеквадратического риска для процедуры выбора модели, построенной на основе улучшенных взвешенных оценок МНК функции гетероскедастичной регрессии [1–3]. Пусть наблюдаемый процесс описывается уравнением

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

где $x_j = j/n$, $S(\cdot) \in W_r^k$ — неизвестная функция, которую требуется оценить (соболевский класс W_r^k определен в [1]), $(\xi_j)_{1\leqslant j\leqslant n}$ — последовательность н. о. р. с. в. с нулевым средним и единичной дисперсией, $(\sigma_j)_{1\leqslant j\leqslant n}$ — неизвестные коэффициенты волатильности, которые могут зависеть от x_j и такие, что $\min_{1\leqslant j\leqslant n}\sigma_j \geqslant \sigma_*$, $\max_{1\leqslant j\leqslant n}\sigma_j \leqslant \sigma^*$, где σ_*, σ^* — некоторые известные постоянные.

Регрессионная модель (1) является важной в стохастических дифференциальных уравнениях для аппроксимации процессов в непрерывным времени, путем использования оценок, имеющих асимптотически минимальные отклонения.

Предлагается построить процедуру выбора модели для оценивания функции S в модели (1) на основе улучшенных оценок [2–3],

$$S^*_\lambda(x) = \sum_{j=1}^{[j_0]} \lambda(j)\theta^*_{j,n}\varphi_j(x) + \sum_{j=[j_0]+1}^{[\omega_\alpha]} \lambda(j)\widehat{\theta}_{j,n}\varphi_j(x), \tag{2}$$

где

$$\theta^*_{j,n} = \begin{cases} \left(1 - \frac{(j_0-1)\sigma_*^2}{n\|\widehat{\theta}_n\|(r+\sqrt{j_0\sigma^{*2}/n})}\right)\widehat{\theta}_{j,n}, & 1\leqslant j\leqslant j_0; \\ \widehat{\theta}_{j,n}, & j > j_0; \end{cases}$$

$$\widehat{\theta}_{j,n} = \frac{1}{n}\sum_{l=1}^{n} y_l\ \varphi_j(x_l),$$

где $\lambda \in \Lambda$ — вектор весовых коэффициентов, элементы которого принимают значение от 0 до 1. Для построения процедуры выбора модели необходимо минимизировать относительно λ функцию

$$J_n(\lambda) = \sum_{j=1}^{n} \lambda^2(j)\theta^{*2}_{j,n} - 2\sum_{j=1}^{n}\lambda(j)(\tilde{\theta}_{j,n} - \theta^*_{j,n}\theta_{j,n}) + \rho\widehat{P_n}(\lambda).$$

Здесь слагаемое

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

— функция штрафа, который мы должны заплатить за использование вместо истинных значений параметров их оценки

$$\widetilde{\theta}_{j,n} = \theta^*_{j,n}\widehat{\theta}_{j,n} - \frac{\sigma^2}{n}.$$

Обозначим $\lambda^* = \arg\min_{\lambda\in\Lambda} J_n(\lambda)$. Как показано в [1]

$$\lambda^*(j) = \mathbf{1}_{1\leqslant j\leqslant j_0} + (1-(j/\omega)^k)\mathbf{1}_{j_0+1\leqslant j\leqslant\omega},$$

где $j_0 = [\omega/\ln n]$, $\omega = \overline{\omega} + (A_k r n)^{1/(2k+1)}$, $A_k = (k+1)(2k+1)/(\pi^{2k}k)$ и $\overline{\omega}$ — некоторая неотрицательная постоянная. Тогда процедура выбора модели определяется равенством

$$S^* = S^*_{\lambda^*}. \tag{3}$$

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

$$R(S^*,S) \leqslant \frac{1+3\rho}{1-3\rho}\min_{\lambda\in\Lambda} R(S^*_\lambda,S) + \frac{1}{n}\Psi_n(\rho), \tag{4}$$

где

$$\Psi_n(\rho) = \frac{cn(2\rho+c(1-\rho))+2\nu\sigma^*}{\rho(1-3\rho)}, \quad c = \frac{(j_0-1)\sigma_*}{n(r+\sqrt{j_0\sigma^{2*}/n})}.$$

СПИСОК ЛИТЕРАТУРЫ

1. *Galthouk L. I., Pergamenshchikov S. M.* Adaptive sequential estimation for ergodic diffusion processes in quadratic metric. — J. Nonparametric Statistics, 2011, v. 23, № 2, p. 255–285.
2. *Pchelintsev E.* Improved estimation in a non-Gaussian parametric regression. — Statistical Inference for Stochastic Processes, 2013, v. 16, is. 1, p. 15–28.
3. *Перелевский С. С., Пчелинцев Е. А.* Адаптивное улучшенное оценивание функции гетероскедастичной регрессии в непрерывном времени. — Обозрение прикл. и промышл. матем., 2015, т. 22, в. 4, с. 491–492.