Е. А. П ч е л и н ц е в (Томск, ТГУ, Руан, Руанский университет)**. Улучшенные оценки параметров регрессии с импульсными шумами.**

Предлагается процедура оценивания параметров регрессии с непрерывным временем при негауссовских шумах, включающих импульсные воздействия, и с неизвестными корреляционными свойствами. Показывается, что эта процедура превосходит по среднеквадратической точности оценки МНК и является оптимальной в минимаксном смысле.

Пусть наблюдаемый процесс y_t описывается уравнением

$$dy_t = S(t)\,dt + d\xi_t, \quad 0 \leqslant t \leqslant n, \tag{1}$$

где $S(t) = \sum_{j=1}^{p} \theta_j \phi_j(t)$, $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$ — вектор неизвестных параметров из некоторого компакта $\Theta \subset \mathbf{R}^p$, $(\phi_j)_{1\leqslant j\leqslant p}$ — система линейно независимых 1-периодических ортонормальных ограниченных функций из пространства $L_2[0,1]$, $(\xi_t)_{t\geqslant 0}$ — скалярный негауссовский процесс Орнштейна–Уленбека, т. е.

$$d\xi_t = a\xi_t\,dt + du_t, \tag{2}$$

где a ($a \leqslant 0$) — мешающий параметр, $(u_t)_{t\geqslant 0}$ — процесс Леви вида

$$u_t = \varrho_1 w_t + \varrho_2 z_t, \tag{3}$$

$(w_t)_{t\geqslant 0}$ — стандартное броуновское движение, $(z_t)_{t\geqslant 0}$ — составной пуассоновский процесс, т. е.

$$z_t = \sum_{j=1}^{N_t} Y_j, \tag{4}$$

$(N_t)_{t\geqslant 0}$ — пуассоновский процесс с интенсивностью $\lambda > 0$ и $(Y_j)_{j\geqslant 1}$ — последовательность независимых одинаково распределенных стандартных гауссовских случайных величин. Параметры a, ϱ_1, ϱ_2 и λ неизвестны.

Задача: оценить неизвестный параметр θ по наблюдениям $(y_t)_{0\leqslant t\leqslant n}$. Качество оценки $\widehat{\theta}$ измеряется квадратическим риском $R(\theta,\widehat{\theta}) = \mathbf{E}_\theta \|\theta - \widehat{\theta}\|^2$.

Классической оценкой для θ в модели (1) является оценка МНК $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p)'$ с $\widehat{\theta}_j = n^{-1}\int_0^n \phi_j(t)\,dy_t$, $j = 1, 2, \ldots, p$. Отсюда с учетом (1) имеем

$$\widehat{\theta} = \theta + \varepsilon_n \zeta(n), \tag{5}$$

где $\varepsilon_n = n^{-1/2}$ и $\zeta(n)$ — случайный вектор с координатами $\zeta_j(n) = n^{-1/2}\int_0^n \phi_j(t)\,d\xi_t$. Эта параметрическая модель не является гауссовской из-за наличия импульсных воздействий z_t в составе шума ξ_t. Однако вектор $\zeta(n)$, как следует из (2), имеет условно-гауссовское распределение с нулевым средним и условной ковариационной матрицей $V_n(\mathcal{G}) = \operatorname{cov}(\zeta(n), \zeta(n)'|\mathcal{G})$ с элементами $v_{ij}(n) = \mathbf{E}\,(\zeta_i(n)\zeta_j(n)|\mathcal{G})$, где $\mathcal{G} = \sigma\{N_t, t \geqslant 0\}$ — σ-алгебра, порожденная пуассоновским процессом.

Таким образом, исходная задача оценивания параметра θ в (1) сводится к оцениванию параметра θ в условно-гауссовской дискретной регрессионной модели (5).

В качестве оценки неизвестного параметра θ предлагается использовать следующую величину:

$$\theta^* = \left(1 - \frac{\varrho_1^2 (p-1)\gamma_p}{n\|\widehat{\theta}\|}\right)\widehat{\theta}, \tag{6}$$

где

$$\gamma_p = \frac{1}{2^{p/2-1}\Gamma(p/2)d}\left[\sum_{j=0}^{p-2} 2^{(j-1)/2}(-1)^{p-j}(\kappa)^{p-1-j}\Gamma\left(\frac{j+1}{2}\right) - (-\kappa)^p I(\kappa)\right],$$

$$\kappa = \frac{d}{\varepsilon_n \sqrt{p\varrho^*(1+2K^2)}}, \quad I(a) = \int_0^\infty \frac{e^{-x^2/2}}{a+x}\,dx, \quad d = \sup\{\|\theta\|\colon\ \theta \in \Theta\}.$$

Здесь $\varrho^* = \varrho_1^2 + \lambda\varrho_2^2$ и K — положительная константа, которой не превосходят функции $(\phi_j)_{1\leqslant j\leqslant p}$. Обозначим $\Delta(\theta) := R(\theta,\theta^*) - R(\theta,\widehat{\theta})$.

Основной результат содержит следующая теорема (см. [1]).

Теорема. *Пусть регрессионная модель определяется уравнениями* (1)–(4). *Тогда для всех $n \geqslant 1$ и $p \geqslant 2$ оценка* (6) *для θ превосходит по точности оценку $\widehat{\theta}$ МНК*:

$$\sup_{\theta\in\Theta} \Delta(\theta) \leqslant -\left[\frac{\varrho_1^2(p-1)\gamma_p}{n}\right]^2.$$

З а м е ч а н и е. Можно показать, что оценка θ^* является также оптимальной в минимаксном смысле.

СПИСОК ЛИТЕРАТУРЫ

1. *Pchelintsev E.* On improved estimation in a conditionally Gaussian regression, 2011. http://hal.archives-ouvertes.fr/hal-00595298/fr/.