МАТЕМАТИКА

УДК 517.16; 519.2

Е.А. Пчелинцев

ПРОЦЕДУРА ДЖЕЙМСА – СТЕЙНА ДЛЯ УСЛОВНО-ГАУССОВСКОЙ РЕГРЕССИИ

В статье рассматривается задача оценивания p-мерного ($p\geq 2$) вектора среднего многомерного условно-гауссовского распределения при квадратической функции потерь. Задача такого типа возникает, например, при оценивании параметров непрерывной регрессионной модели с негауссовским процессом Орнштейна – Уленбека. Предлагается модификация процедуры Джеймса – Стейна вида $\theta^*(Y) = (1-c/\|Y\|)Y$, где Y – наблюдение и $c > 0$ – специальная константа. Для этой оценки найдена явная верхняя граница для квадратического риска и показано, что ее риск строго меньше риска обычной оценки максимального правдоподобия для размерности $p\geq 2$. Эта процедура применяется к проблеме параметрического оценивания непрерывной условно-гауссовской регрессии и к оцениванию вектора среднего многомерного нормального распределения, когда ковариационная матрица неизвестна и зависит от некоторых мешающих параметров.

Ключевые слова: *условно-гауссовская регрессия*, *улучшенное оценивание*, *процедура Джеймса – Стейна*, *негауссовский процесс Орнштейна – Уленбека.*

Введение

В 1961 г. Джеймс и Стейн, рассматривая задачу оценивания вектора среднего θ p-мерного нормального распределения случайного вектора Y с единичной ковариационной матрицей I_p, ввели оценку

$$\hat{\theta}_{JS} = \left(1 - \frac{p-2}{\|Y\|^2}\right) Y , \tag{1}$$

которая для $p\geq 3$ превосходит оценку максимального правдоподобия

$$\hat{\theta}_{ML} = Y \tag{2}$$

при квадратическом риске

$$R(\theta, \hat{\theta}) = \mathbf{E}_\theta \left\|\theta - \hat{\theta}\right\|^2 , \tag{3}$$

т.е. для всех значений параметра θ

$$R(\theta, \hat{\theta}_{JS}) < R(\theta, \hat{\theta}_{ML}) .$$

Этот неожиданный результат вызвал большой интерес математиков статистиков и стимулировал многих авторов к развитию теории улучшенного оценивания. Задача Джеймса – Стейна была изучена для более общих моделей, в том числе с неизвестной ковариационной матрицей [1, 2, 6, 9]. Значительные усилия были направлены на решение задачи улучшенного оценивания в негауссовских моделях.

В [3, 5] были предложены процедуры улучшенного оценивания для регрессии со сферически симметричными распределениями шумов. В [4] изучалась непараметрическая модель регрессии с зависимыми шумами.

Оценки Джеймса – Стейна и другие улучшенные оценки нашли применение в эконометрике и в задачах, связанных с обработкой сигналов.

В статье рассматривается задача оценивания среднего в модели регрессии, в которой шум является только условно-гауссовским. Предположим, что наблюдаемый p-мерный случайный вектор Y удовлетворяет уравнению

$$Y = \theta + \xi, \tag{4}$$

где θ – вектор постоянных параметров из некоторого компакта $\Theta \subset \mathbb{R}^p$, ξ – условно-гауссовский случайный вектор с нулевым средним и условной ковариационной матрицей $D(\mathcal{G})$, т.е. $Law(\xi|\mathcal{G})=\mathcal{N}_p(0, D(\mathcal{G}))$, где $\mathcal{G}$ – некоторая фиксированная σ-алгебра.

Для данной модели исследование оценки Джеймса – Стейна провести не удается и вопрос о нахождении её риска остается открытым. Предлагается рассмотреть следующую модификацию оценки Джеймса – Стейна:

$$\theta^* = \left(1 - \frac{c}{\|Y\|}\right) Y, \tag{5}$$

где c – положительная константа, определяемая ниже. Основное отличие оценки (5) от (1) в том, что вместо квадрата нормы $\|Y\|^2$ используется норма в первой степени $\|Y\|$. Как показано ниже, оценка θ^* позволяет решать задачу улучшенного оценивания регрессии с условно-гауссовскими шумами при некоторых дополнительных предположениях. В теореме 2.1 раздела 2 устанавливается, что оценка (5) превосходит по точности оценку максимального правдоподобия (2) равномерно по θ для любого компакта $\Theta \subset \mathbb{R}^p$ и всех размерностей $p \geq 2$. Для сравнения заметим, что преимущество оценки Джеймса – Стейна для гауссовской регрессии проявляется только для $p \geq 3$. В разделе 3 оценка (5) применяется для решения задачи улучшенного параметрического оценивания в непрерывной регрессионной модели с негауссовским шумом, являющимся смесью белого шума и импульсных возмущений, описываемых составным процессом Пуассона. В разделе 4 оценка (5) применяется для оценивания параметра в дискретной регрессионной модели с гауссовским шумом, зависящим от неизвестных мешающих параметров. В приложении приводятся некоторые технические результаты.

2. Верхняя граница для риска оценки

В данном разделе получена верхняя граница для риска оценки (5) при некоторых условиях на случайную ковариационную матрицу $D(\mathcal{G})$.

Предположим, что

($\mathbf{C_1}$) – минимальное собственное значение матрицы $D(\mathcal{G})$ отделено снизу от нуля, т.е. существует положительная постоянная λ_*, такая, что

$$\lambda_{\min}(D(\mathcal{G})) \geq \lambda_* \quad \text{п.н.};$$

($\mathbf{C_2}$) – математическое ожидание максимального собственного значения матрицы $D(\mathcal{G})$ ограничено сверху на некотором компакте $\Theta \subset \mathbb{R}^p$, т.е.

$$\sup_{\theta\in\Theta}\mathbf{E}_\theta\lambda_{\max}(D(\mathcal{G}))\le a^*,$$

где a^* – некоторая положительная постоянная.

Обозначим разность рисков оценок (5) и (2) как

$$\Delta(\theta):=R(\theta,\theta^*)-R(\theta,\hat\theta_{ML}).$$

Нам потребуется также следующая константа

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

где $\kappa=d/\sqrt{a^*}$, $I(\alpha)=\int_0^\infty\frac{e^{-r^2/2}}{\alpha+r}dr$ и $d=\sup\{\|\theta\|:\ \theta\in\Theta\}$.

Основной результат для условно-гауссовской регрессии содержит следующая теорема.

Теорема 2.1. Пусть шум ξ в (4) имеет p-мерное условно-гауссовское распределение $\mathcal{N}_p(0,\ D(\mathcal{G}))$ и его ковариационная матрица $D(\mathcal{G})$ удовлетворяет условиям $(\mathbf{C_1})$, $(\mathbf{C_2})$ для некоторого компакта $\Theta\subset\mathbb{R}^p$. Тогда оценка (5) с $c=(p-1)\lambda_*\gamma_p$ доминирует оценку максимального правдоподобия (2) для всех $p\ge2$, т.е.

$$\sup_{\theta\in\Theta}\Delta(\theta)\le-[(p-1)\lambda_*\gamma_p]^2.$$

Доказательство. Сначала установим нижнюю границу для математического ожидания случайной величины $\|Y\|^{-1}$.

Лемма 2.2. В условиях теоремы 2.1

$$\inf_{\theta\in\Theta}\mathbf{E}_\theta\frac{1}{\|Y\|}\ge\gamma_p.$$

Доказательство леммы приводится в приложении.

Чтобы получить верхнюю границу для $\Delta(\theta)$, нам потребуется модифицировать доказательство леммы Стейна [9] применительно к модели (4) со случайной ковариационной матрицей.

Рассмотрим риски оценок (2) и (5)

$$R(\hat\theta_{ML},\theta)=\mathbf{E}_\theta\left\|\hat\theta_{ML}-\theta\right\|^2=\mathbf{E}_\theta\left(\mathbf{E}\left(\left\|\hat\theta_{ML}-\theta\right\|^2\mid\mathcal{G}\right)\right)=\mathbf{E}_\theta trD(\mathcal{G});$$

$$R\left(\theta^*,\theta\right)=R\left(\hat\theta_{ML},\theta\right)+\mathbf{E}_\theta\left[E\left((g(Y)-1)^2\|Y\|^2\mid\mathcal{G}\right)\right]+2\sum_{j=1}^d\mathbf{E}_\theta\left[\mathbf{E}\left((g(Y)-1)Y_j(Y_j-\theta_j)\mid\mathcal{G}\right)\right],$$

где $g(Y)=1-c/\|Y\|$.

Обозначив $f(Y)=(g(Y)-1)Y_j$ и используя условную плотность распределения вектора Y относительно σ-алгебры $\mathcal{G}$

$$p_Y(x\mid\mathcal{G})=\frac{1}{(2\pi)^{p/2}\sqrt{\det D(\mathcal{G})}}\exp\left(-\frac{(x-\theta)'D^{-1}(\mathcal{G})(x-\theta)}{2}\right),$$

имеем

$$I_j := \mathbf{E}\left(f(Y)(Y_j-\theta_j)\,|\,\mathcal{G}\right)=\int\limits_{\mathbb{R}^p} f(x)(x_j-\theta_j)p_Y(x\,|\,\mathcal{G})dx\,,\ \ j=\overline{1,p}\,.$$

Делая замену переменной $u=D^{-1/2}(\mathcal{G})(x-\theta)$ и полагая $\tilde{f}(u)=f(D^{1/2}(\mathcal{G})u+\theta)$, находим

$$I_j=\frac{1}{(2\pi)^{p/2}}\sum_{l=1}^{p}<D^{1/2}(\mathcal{G})>_{jl}\int\limits_{\mathbb{R}^p}\tilde{f}(u)u_l\exp\left(-\frac{\|u\|^2}{2}\right)du,\ \ j=\overline{1,p}\,,$$

где $<A>_{ij}$ обозначает (i, j)-й элемент матрицы A. Эти величины можно переписать как

$$I_j=\sum_{l=1}^{p}\sum_{k=1}^{p}\mathbf{E}\left(<D^{1/2}(\mathcal{G})>_{jl}<D^{1/2}(\mathcal{G})>_{kl}\frac{\partial f}{\partial u_k}(u)\big|_{u=Y}\,|\,\mathcal{G}\right),\ \ j=\overline{1,p}\,.$$

Таким образом, квадратический риск для оценки (5) представляется в виде

$$R\left(\theta^*,\theta\right)=R\left(\hat{\theta}_{ML},\theta\right)+\mathbf{E}_\theta\left(g(Y)-1\right)^2\|Y\|^2+$$

$$+2\mathbf{E}_\theta\left(\sum_{j=1}^{p}\sum_{l=1}^{p}\sum_{k=1}^{p}<D^{1/2}(\mathcal{G})>_{jl}<D^{1/2}(\mathcal{G})>_{kl}\frac{\partial}{\partial u_k}\left[\left(g(u)-1\right)u_j\right]\big|_{u=Y}\right).$$

После простых преобразований получаем

$$R\left(\theta^*,\theta\right)=R\left(\hat{\theta}_{ML},\theta\right)+\mathbf{E}_\theta W(Y)\,,$$

где $$W(x)=c^2+2c\frac{x'D(\mathcal{G})x}{\|x\|^3}-2trD(\mathcal{G})\cdot c\frac{1}{\|x\|}.$$

Отсюда разность рисков $\Delta(\theta)$ принимает вид

$$\Delta(\theta)=\mathbf{E}_\theta W(Y).$$

Поскольку $x'Ax\le\lambda_{\max}(A)\|x\|^2$, приходим к неравенству

$$\Delta(\theta)\le c^2-2c\mathbf{E}_\theta\frac{trD(\mathcal{G})-\lambda_{\max}(D(\mathcal{G}))}{\|Y\|}.$$

Следовательно,

$$\Delta(\theta)\le c^2-2c\sum_{i=2}^{d}\mathbf{E}_\theta\frac{\lambda_i(D(\mathcal{G}))}{\|Y\|}.$$

Принимая во внимание условие $(\mathbf{C_1})$ и лемму 2.2, имеем

$$\Delta(\theta)\le c^2-2(p-1)\lambda_*\gamma_p c=:\phi(c_n).$$

Минимизируя функцию $\phi(c)$ по c, приходим к желаемому результату

$$\Delta(\theta)\le-\left[(p-1)\lambda_*\gamma_p\right]^2.$$

Теорема 2.1 доказана.

В отличие от (1), оценку (5) можно использовать для регрессионной модели с зависимыми гауссовскими шумами.

Следствие 2.3. Пусть в (4) шум $\xi \sim \mathcal{N}_p(0, D)$ с положительно определенной неслучайной матрицей $D>0$ и $\lambda_{\min}(D) \geq \lambda_* > 0$. Тогда оценка (5) с $c=(p-1)\lambda_*\gamma_p$ доминирует оценку максимального правдоподобия для всех $p \geq 2$ и любого компакта $\Theta \subset \mathbb{R}^p$, т.е.

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq -[(p-1)\lambda_*\gamma_p]^2 .$$

Замечание. Заметим, что если $D=\sigma^2 I_p$, то

$$\sup_{\theta \in \Theta} \Delta(\theta) \leq -[(p-1)\sigma^2\gamma_p]^2 .$$

Следствие 2.4. Если $\xi \sim \mathcal{N}_p(0, I_p)$ и $\theta=0$ в модели (4), то риск оценки (5) дается формулой

$$R(0,\theta^*) = p - \left[\frac{(p-1)\Gamma((p-1)/2)}{\sqrt{2}\Gamma(p/2)}\right]^2 =: r_p .$$

Применяя формулу Стирлинга для гамма-функции $\Gamma(x)=\sqrt{2\pi}x^{x-1/2}e^{-x}(1+o(1))$, можно проверить, что $r_p \to 0{,}5$ при $p \to \infty$. Поведение риска r_p при малых значениях p представлено на рис. 1. Заметим, что в этом случае риск оценки Джеймса – Стейна является постоянным для всех $p \geq 3$, т.е. $R(\theta, \hat{\theta}_{JS}) = 2$, а риск оценки максимального правдоподобия $\hat{\theta}_{ML}$ равен p и стремится к бесконечности при $p \to \infty$. Таким образом, оценка (5) для регрессии с единичной матрицей шумов превосходит по точности как оценку максимального правдоподобия, так и оценку Джеймса – Стейна.

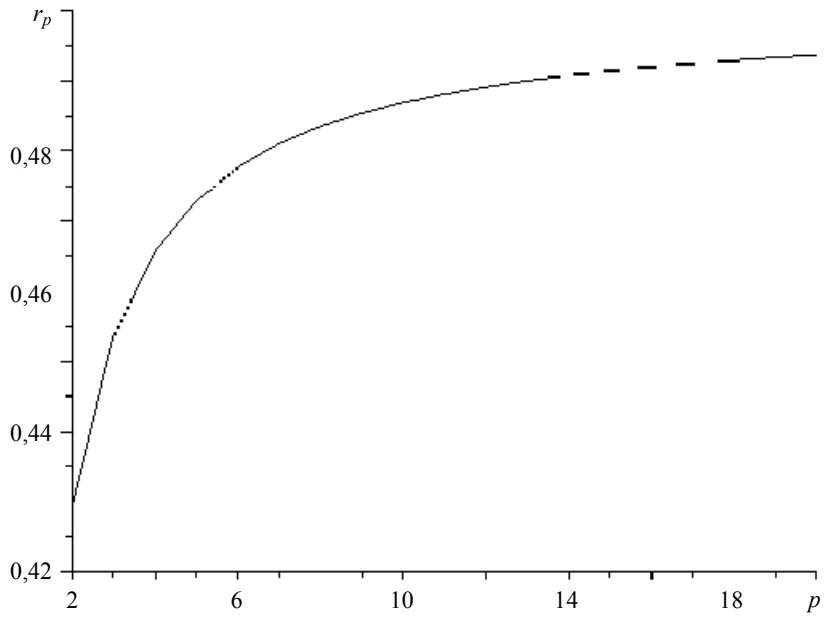

Рис. 1. Риск оценки (5) при $\theta = 0$

3. Улучшенное оценивание в негауссовской регрессионной модели

В данном разделе применим предложенную оценку (5) к негауссовской непрерывной регрессионной модели. Пусть наблюдаемый процесс (y_t) описывается уравнением

$$dy_t = \sum_{j=1}^{p} \theta_j \phi_j(t)dt + d\xi_t, \quad 0 \le t \le n, \tag{6}$$

где $\theta=(\theta_1,\ldots, \theta_p)'$ – вектор неизвестных параметров из некоторого компакта $\Theta \subset \mathbb{R}^p$, n – натуральное число. Предположим, что $(\phi_j)_{1\le j\le p}$ – система линейно независимых 1-периодических ортонормальных функций из пространства $\mathcal{L}_2[0,1]$; шум $(\xi_t)_{t\ge 0}$ – скалярный негауссовский процесс Орнштейна – Уленбека, т.е.

$$d\xi_t = a\xi_t dt + du_t, \tag{7}$$

где $a\le 0$ – мешающий параметр, а (u_t) – процесс Леви вида

$$u_t = \varrho_1 w_t + \varrho_2 z_t, \tag{8}$$

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

$$z_t = \sum_{j=1}^{N_t} Y_j, \tag{9}$$

$(N_t)_{t\ge 0}$ – однородный пуассоновский процесс с интенсивностью $\lambda > 0$, а $(Y_j)_{j\ge 1}$ – последовательность н.о.р. гауссовских случайных величин с параметрами (0, 1). Параметры a, ϱ_1, ϱ_2 и λ – неизвестные постоянные.

Заметим, что регрессионная модель (6) – (9) позволяет рассматривать также модели с зависимыми шумами, включающими импульсные возмущения.

Задача – оценить неизвестный вектор параметров $\theta=(\theta_1,\ldots, \theta_p)'$ в модели (6) по наблюдениям $(y_t)_{0\le t\le n}$.

Качество оценки $\tilde{\theta}$ будем измерять квадратическим риском

$$R(\theta,\tilde{\theta}) = \mathbf{E}_\theta \left\|\theta-\tilde{\theta}\right\|^2.$$

Пусть $\mathcal{G} = \sigma\{N_t, t\ge 0\}$ обозначает σ-алгебру, порожденную пуассоновским процессом. Заметим, что модель (6) является условно-гауссовской относительно определенной σ-алгебры $\mathcal{G}$. Поэтому можно использовать оценку (5), чтобы получить улучшенную оценку параметра θ. С этой целью представим исходную регрессионную модель с непрерывным временем (6) в виде модели типа (4).

Классической оценкой для неизвестного вектора θ в модели (6) по наблюдениям $(y_t)_{0\le t\le n}$ является оценка по методу наименьших квадратов (МНК) $\hat{\theta}=(\hat{\theta}_1,\ldots,\hat{\theta}_p)'$ с

$$\hat{\theta}_j = \frac{1}{n}\int_0^n \phi_j(t)\, dy_t, \quad j=\overline{1,p}.$$

Отсюда, принимая во внимание (6), имеем

$$\hat{\theta} = \theta + \varepsilon_n \zeta(n), \tag{10}$$

где $\varepsilon_n = n^{-1/2}$ и $\zeta(n)$ – случайный вектор с координатами

$$\zeta_j(n) = \frac{1}{\sqrt{n}} \int_0^n \phi_j(t) d\xi_t .$$

Заметим, что вектор $\zeta(n)$ имеет условно-гауссовское распределение с нулевым средним и случайной ковариационной матрицей $V_n(\mathcal{G}) = \mathrm{cov}(\zeta(n), \zeta(n)' \mid \mathcal{G})$ с элементами $v_{ij}(n) = \mathbf{E}(\zeta_i(n)\zeta_j(n) \mid \mathcal{G})$.

Таким образом, исходная задача оценивания параметра θ в (6) сводится к задаче оценивания неизвестного параметра θ в условно-гауссовской регрессии (10).

Теорема 3.1. Пусть регрессионная модель определяется уравнениями (6) – (9), $\varrho_1 > 0$. Тогда для всех $n \geq 1$ и $p \geq 2$ оценка для θ

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

доминирует для любого компакта $\Theta \subset \mathbb{R}^p$ оценку МНК $\hat{\theta}$:

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

Для доказательства теоремы достаточно проверить условия ($\mathbf{C_1}$), ($\mathbf{C_2}$) для случайной матрицы $V_n(\mathcal{G})$ и применить теорему 2.1. Доказательство теоремы 3.1 приводится в приложении.

4. Улучшенное оценивание в авторегрессии

В данном разделе оценка (5) применяется для решения задачи улучшенного оценивания неизвестного среднего многомерного нормального распределения, когда ковариационная матрица неизвестна и зависит от некоторых мешающих параметров.

I. Пусть в модели (4) шум $\xi = (\xi_1, ..., \xi_p)'$ описывается гауссовским процессом авторегрессии

$$\xi_k = a\xi_{k-1} + \varepsilon_k, \quad k = \overline{1, p}, \tag{11}$$

где $|a| < 1$, $\mathbf{E}\xi_0 = 0$ и $\varepsilon_1, ..., \varepsilon_p$ – независимые стандартные гауссовские случайные величины. Предположим, что параметр a в уравнении (11) неизвестен и принадлежит сегменту $[-a_0, a_0]$, где a_0 – известное число из интервала (0,1).

Нетрудно проверить, что ковариационная матрица шума ξ имеет вид

$$D(a) = \frac{1}{1-a^2} \begin{pmatrix} 1 & a & \dots & a^{p-1} \\ a & 1 & \dots & a^{p-2} \\ \vdots & \vdots & \vdots & \vdots \\ a^{p-1} & a^{p-2} & \dots & 1 \end{pmatrix}. \tag{12}$$

Далее рассмотрим задачу оценивания вектора параметров θ в модели (4) с шумами (11).

Предложение 4.1. Пусть шум ξ в модели (4) определяется уравнением (11) с $a \in [-a_0, a_0]$. Тогда для всех $p>1/(1-a_0)^2$ оценка для θ, определяемая равенством

$$\theta^* = \left(1 - \left(p - \frac{1}{(1-a_0)^2}\right)\frac{\gamma_p}{\|Y\|}\right)Y,$$

доминирует для любого компакта $\Theta \subset \mathbb{R}^p$ оценку максимального правдоподобия, т.е.

$$\sup_{\theta\in\Theta} \Delta(\theta) \le -\left(p - \frac{1}{(1-a_0)^2}\right)^2 \gamma_p^2.$$

Доказательство. Воспользуемся схемой доказательства теоремы 2.1. Имеем, что $\operatorname{tr} D(a) = p/(1-a^2)$. Оценим сверху максимальное собственное значение матрицы $D(a)$, учитывая, что

$$\lambda_{\max}(D(a)) = \sup_{\|z\|=1} z'D(a)z.$$

Из (12) следует, что

$$z'D(a)z = \sum_{i=1}^{p}\sum_{j=1}^{p} < D(a) >_{ij} z_i z_j = \frac{1}{1-a^2}\left(1 + 2\sum_{i=1}^{p-1}\sum_{l=1}^{p-i} a^l z_i z_{l+i}\right) =$$

$$= \frac{1}{1-a^2}\left(1 + 2\sum_{l=1}^{p-1} a^l \sum_{i=1}^{p-l} z_i z_{i+l}\right).$$

Применяя неравенство Коши – Буняковского, получаем

$$\lambda_{\max}(D(a)) \le \frac{1}{1-a_0^2}\left(1 + 2\sum_{l=1}^{\infty} a_0{}^l\right) = \frac{1}{(1-a_0)^2}.$$

Таким образом,

$$\operatorname{tr} D(a) - \lambda_{\max}(D(a)) \ge p - \frac{1}{(1-a_0)^2}.$$

Отсюда, в силу теоремы 2.1 приходим к утверждению предложения 4.1.

II. Рассмотрим модель (4) с шумом, зависящим от нескольких неизвестных мешающих параметров. Пусть в модели (4) шум $\xi=(\xi_1,...,\xi_p)'$ описывается стационарным многомерным гауссовским процессом авторегрессии q-го порядка $AR(q)$

$$\xi_k = a_1\xi_{k-1} + a_2\xi_{k-2} + ... + a_q\xi_{k-q} + \varepsilon_k, \quad k = \overline{1,p}, \tag{13}$$

где $\varepsilon_1,...,\varepsilon_p$ – независимые стандартные гауссовские случайные величины. Неизвестный вектор $a=(a_1,...,a_q)'$ принадлежит области устойчивости процесса

$$\mathcal{A} = \{a \in \mathbb{R}^q : \max_{1\le i\le q} \operatorname{Re}\lambda_i(a) < 0\},$$

где $(\lambda_i(a))_{1\le i\le q}$ – собственные значения матрицы

$$A = A(a) = \begin{pmatrix} a_1 & \dots & a_q \\ & I_{q-1} & 0 \end{pmatrix};$$

I_q – единичная матрица порядка q.

Как можно проверить, ковариационная матрица вектора ξ в данном случае имеет вид

$$D = \begin{pmatrix} < F >_{11} & < AF >_{11} & \dots & < A^{p-1}F >_{11} \\ < AF >_{11} & < F >_{11} & \dots & < A^{p-2}F >_{11} \\ \vdots & \vdots & \vdots & \vdots \\ < A^{p-1}F >_{11} & < A^{p-2}F >_{11} & \dots & < F >_{11} \end{pmatrix}; \qquad (14)$$

F – матрица, удовлетворяющая уравнению

$$F - AFA' = B ,$$

$B = \left\| \delta_{1i}\delta_{1j} \right\|$ – матрица размера $q \times q$, δ_{ij} – символ Кронекера.

Обозначим $p_0 = 1 + \sup\limits_{a \in K}(2c_* \|F\|)/((1-\rho) < F >_{11})$, где K – некоторый компакт из области устойчивости $\mathcal{A}$, $c_* > 0$ и $0 < \rho < 1$ – постоянные.

Предложение 4.2. Пусть вектор шумов ξ в (4) определяется (13). Тогда для всех $p > p_0$ оценка для θ, определяемая формулой

$$\theta^* = \left(1 - \left((p-1) < F >_{11} - \frac{2c_*\rho\|F\|}{1-\rho}\right)\frac{\gamma_p}{\|Y\|}\right)Y ,$$

доминирует для любого компакта $\Theta \subset \mathbb{R}^p$ оценку максимального правдоподобия, т.е.

$$\sup_{\theta \in \Theta} \Delta(\theta) \le -\left((p-1) < F >_{11} - \frac{2c_*\rho\|F\|}{1-\rho}\right)^2 \gamma_p^2 .$$

Доказательство. Используем схему доказательства теоремы 2.1. Имеем, что $\operatorname{tr} D = p < F >_{11}$. Оценим сверху максимальное собственное значение матрицы D, воспользовавшись тем, что $\lambda_{\max}(D) = \sup\limits_{\|z\|=1} z'Dz$. Учитывая (14), находим

$$z'Dz = \left|\sum_{i=1}^{p}\sum_{j=1}^{p} < D >_{ij} z_i z_j\right| = \left| < F >_{11} + 2\sum_{i=1}^{p-1}\sum_{l=1}^{p-i} < A^l F >_{11} z_i z_{l+i}\right| \le$$

$$\le \ < F >_{11} + 2\left|\sum_{l=1}^{p-1} < A^l F >_{11} \sum_{i=1}^{p-l} z_i z_{l+i}\right| .$$

Применяя неравенство Коши – Буняковского и матричные неравенства $\|AB\| \le \|A\|\|B\|$ и $\left\|A^l\right\| \le c_*\rho^l$, получаем

$$\sup_{\|z\|=1} z'Rz \le < F >_{11} + \frac{2c_*\rho\|F\|}{1-\rho} .$$

Таким образом,

$$\operatorname{tr} D - \lambda_{\max}(D) \ge (p-1) < F >_{11} - \frac{2c_*\rho\|F\|}{1-\rho} .$$

Принимая во внимание теорему 2.1, приходим к утверждению предложения 4.2.

5. Приложение

5.1. Доказательство леммы 2.2

Доказательство. Исходя из (4), имеем

$$J = \mathbf{E}_\theta \frac{1}{\|Y\|} = \mathbf{E}_\theta \frac{1}{\|\theta + \xi\|} \geq \mathbf{E}_\theta \frac{1}{d + \|\xi\|}.$$

Перейдем к повторному условному математическому ожиданию. Так как при фиксированной σ-алгебре $\mathcal{G}$ случайный вектор ξ распределен нормально с нулевым средним, то

$$J \geq \mathbf{E}_\theta \frac{1}{(2\pi)^{p/2}\sqrt{\det D(\mathcal{G})}} \int\limits_{\mathbb{R}^p} \frac{\exp\{-x'D(\mathcal{G})^{-1}x/2\}}{d + \|x\|} dx.$$

Обозначая $u = D(\mathcal{G})^{-1/2}x$ и применяя оценку $u'D(\mathcal{G})u \leq \lambda_{\max}(D(\mathcal{G}))\|u\|^2$, находим

$$J \geq \frac{1}{(2\pi)^{p/2}} \mathbf{E}_\theta \int\limits_{\mathbb{R}^p} \frac{\exp\{-\|u\|^2/2\}}{d + \lambda_{\max}^{1/2}(D(\mathcal{G}))\|u\|} du.$$

Далее, переходя к сферическим координатам, имеем

$$J \geq \frac{2}{2^{p/2}\Gamma(p/2)} \mathbf{E}_\theta \int\limits_0^\infty \frac{r^{p-1}e^{-r^2/2}}{d + \lambda_{\max}^{1/2}(D(\mathcal{G}))r} dr.$$

Отсюда, при использовании неравенства Йенсена, Коши – Буняковского и условия $(\mathbf{C_2})$ следует, что

$$J \geq \frac{\kappa}{2^{p/2-1}\Gamma(p/2)d} \int\limits_0^\infty \frac{r^{p-1}e^{-r^2/2}}{\kappa + r} dr.$$

Вычислим интеграл, стоящий в правой части неравенства. Разделив r^{p-1} на $\kappa + r$, получим

$$\int\limits_0^\infty \frac{r^{p-1}e^{-r^2/2}}{\kappa + r} dr = \sum_{j=0}^{p-2} 2^{\frac{j-1}{2}} (-\kappa)^{p-2-j} \Gamma\left(\frac{j+1}{2}\right) + (-\kappa)^{p-1} I(\kappa),$$

где функция $I(\alpha) = \int\limits_0^\infty \frac{e^{-r^2/2}}{\alpha + r} dr$. Лемма 2.2 доказана.

5.2. Проверка условий $(\mathbf{C_1})$ и $(\mathbf{C_2})$ для матрицы $V_n(\mathcal{G})$

Элементы матрицы $V_n(\mathcal{G})$ могут быть переписаны как [8]

$$v_{ij}(n) = \frac{\varrho_1^2}{n} \int\limits_0^n \phi_i(t)\phi_j(t)dt + \frac{\varrho_1^2}{2n} \int\limits_0^n \left(\phi_i(t)\varepsilon_{\phi_j}(t) + \phi_j(t)\varepsilon_{\phi_i}(t)\right) dt + \frac{\varrho_2^2}{n} \sum_{l \geq 1} \phi_i(T_l)\phi_j(T_l)\chi_{(T_l \leq n)}$$

$$+ \frac{\varrho_2^2}{n} \sum_{l \geq 1} \int\limits_0^n \left(\phi_i(t)L_{\phi_j}(t - T_l, T_l) + \phi_j(t)L_{\phi_i}(t - T_l, T_l)\right) \chi_{(T_l \leq t)} dt, \qquad (15)$$

где $$\varepsilon_f(t) = a\int_0^t e^{a(t-v)} f(v)(1 + e^{2av})\, dv,$$

$$L_f(x,z) = ae^{ax}\left(f(z) + a\int_0^x e^{av} f(v+z)\,dv\right)$$

и $(T_l)_{l\geq 1}$ – моменты скачков пуассоновского процесса $(N_t)_{t\geq 0}$, т.е.

$$T_l = \inf\{t \geq 0 : N_t = l\}.$$

Лемма 5.1. Пусть $(\xi_t)_{t\geq 0}$ определяется уравнением (7) с $a \leq 0$. Тогда минимальное собственное значение матрицы $V_n(\mathcal{G})$ с элементами, определенными (15), удовлетворяет следующему неравенству:

$$\inf_{n\geq 1} \lambda_{\min}(V_n(\mathcal{G})) \geq \varrho_1^2 .$$

Доказательство. Согласно (15), матрицу $V_n(\mathcal{G})$ можем представить в виде

$$V_n(\mathcal{G}) = \varrho_1^2 I_p + F_n + B_n(\mathcal{G}),$$

где F_n – неслучайная матрица с элементами

$$f_{ij}(n) = \frac{\varrho_1^2}{2n}\int_0^n \left(\phi_i(t)\varepsilon_{\phi_j}(t) + \phi_j(t)\varepsilon_{\phi_i}(t)\right)dt ,$$

а $B_n(\mathcal{G})$ – случайная матрица с элементами

$$b_{ij}(n) = \frac{\varrho_2^2}{n}\sum_{l\geq 1}\left[\phi_i(T_l)\phi_j(T_l)\chi_{(T_l\leq n)} + \int_0^n \left(\phi_i(t)L_{\phi_j}(t-T_l,T_l) + \phi_j(t)L_{\phi_i}(t-T_l,T_l)\right)\chi_{(T_l\leq t)}dt\right].$$

Отсюда находим

$$z'V_n(\mathcal{G})z = \varrho_1^2 z'z + z'F_n z + z'B_n(\mathcal{G})z \geq \varrho_1^2 \|z\|^2$$

для всех $n \geq 1$. Поэтому

$$\lambda_{\min}(V_n(\mathcal{G})) \geq \varrho_1^2 .$$

Лемма 5.1 доказана.

Обозначим $\varrho^* = \varrho_1^2 + \lambda\varrho_2^2$ и $M = 1+2K^2$, где $\max\limits_{0\leq t\leq n}\left|\phi_j(t)\right| \leq K$, $j = \overline{1,p}$.

Лемма 5.2. Пусть $(\xi_t)_{t\geq 0}$ определяется уравнением (7) с $a \leq 0$. Тогда максимальное собственное значение матрицы $V_n(\mathcal{G})$ с элементами, определенными (15), удовлетворяет для любого компакта $\Theta \subset \mathbb{R}^p$ следующему неравенству:

$$\sup_{n\geq 1}\sup_{\theta\in\Theta} \mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \leq Mp\varrho^* .$$

Доказательство. Имеем

$$\mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \leq \mathbf{E}_\theta tr(V_n(\mathcal{G})) = \sum_{j=1}^p \mathbf{E}_\theta \zeta_j^2(n) = \frac{\varrho^*}{n}\sum_{j=1}^p \tau_j(n),$$

где $\tau_j(n) = \int_0^n \phi_j^2(t)dt + a\int_0^n \phi_j(t)\int_0^t e^{a(t-u)}\phi_j(u)\left(1+e^{2au}\right)dudt$. Поскольку $(\phi_j)_{1\leq j\leq p}$ – 1-периодические ортонормальные функции, то $\int_0^n \phi_j^2(t)dt = n$ и, ввиду ограничен-

ности этих функций, для любого $a \le 0$ находим

$$a\int_0^n \phi_j(t)\int_0^t e^{a(t-u)}\phi_j(u)\left(1+e^{2au}\right)dudt \le 2K^2|a|\int_0^n\int_0^t e^{a(t-u)}dudt \le 2K^2 n\,.$$

Таким образом, для всех $n \ge 1$ и $\theta \in \Theta$

$$\mathbf{E}_\theta \lambda_{\max}(V_n(\mathcal{G})) \le Mp\varrho^*\,.$$

Лемма 5.2 доказана.

В леммах 5.1 и 5.2 показано, что матрица $V_n(\mathcal{G})$ является положительно определенной и для любого компакта $\Theta \subset \mathbb{R}^p$ удовлетворяет условиям $(\mathbf{C_1})$, $(\mathbf{C_2})$ с $\lambda_* = \varrho_1^2$ и $a^* = Mp\varrho^*$. В силу теоремы 2.1 получаем утверждение теоремы 3.1.

ЛИТЕРАТУРА

1. *Berger J.O.*, *Haff L.R.* A class of minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix // Statist. Decisions. 1983. No. 1. P. 105–129.
2. *Efron B.*, *Morris C.* Families of minimax estimators of the mean of a multivariate normal distribution // Ann. Statist. 1976. No. 4. P. 11–21.
3. *Fourdrinier D.* Statistique inférentielle // D. Fourdrinier. Dunod. 2002. P. 336.
4. *Fourdrinier D.*, *Pergamenshchikov S.* Improved selection model method for the regression with dependent noise // Ann. Inst. Statist. Math. 2007. V. 59 (3). P. 435–464.
5. *Fourdrinier D.*, *Strawderman W.E.*, *William E.* A unified and generalized set of shrinkage bounds on minimax Stein estimates // J. Multivariate Anal. 2008. V. 99. P. 2221–2233.
6. *Gleser L.J.* Minimax estimators of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix // Ann. Statist. 1986. V. 14. No. 1625–1633.
7. *James W.*, *Stein C.* Estimation with quadratic loss // Proceedings of the Fourth Berkeley Symposium on Mathematics Statistics and Probability. V. 1. Berkeley: University of California Press, 1961. P. 361–380.
8. *Konev V.*, *Pergamenchtchikov S.* Efficient robust nonparametric estimation in a semimartingale regression model. URL: http://hal.archives-ouvertes.fr/hal-00526915/fr/ (2010).
9. *Stein C.* Estimation of the mean of a multivariate normal distribution // Ann. Statist. 1981. V. 9(6). P. 1135–1151.

Статья поступила 19.07.2011 г.

Pchelintsev E.A. THE JAMES–STEIN PROCEDURE FOR A CONDITIONALLY GAUSSIAN REGRESSION. The paper considers the problem of estimating a p-dimensional ($p \ge 2$) mean vector of a multivariate conditionally normal distribution under quadratic loss. The problem of this type arises when estimating the parameters in a continuous time regression model with a non-Gaussian Ornstein–Uhlenbeck process. We propose a modification of the James–Stein procedure of the form $\theta^*(Y) = (1 - c/\|Y\|)\,Y$, where Y is an observation and $c > 0$ is a special constant. This estimate allows one to derive an explicit upper bound for the quadratic risk and has a significantly smaller risk than the usual maximum likelihood estimator for the dimensions $p \ge 2$. This procedure is applied to the problem of parametric estimation in a continuous time conditionally Gaussian regression model and to that of estimating the mean vector of a multivariate normal distribution when the covariance matrix is unknown and depends on some nuisance parameters.

Keywords: conditionally Gaussian regression model; improved estimation; James – Stein procedure; non-Gaussian Ornstein – Uhlenbeck process.

PCHELINZEV Evgenii Anatol'evich (Tomsk State University,
Laboratoire de Mathématiques Raphaël Salem, Université de Rouen (France))
E-mail: evgen-pch@yandex.ru