УДК 519.234.3

В. Б. Горяинов

# АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ ОЦЕНОК НАИМЕНЬШИХ МОДУЛЕЙ КОЭФФИЦИЕНТОВ ПРОСТРАНСТВЕННОЙ АВТОРЕГРЕССИИ

*Для процесса пространственной авторегрессии порядка* $(1,1)$ *установлена асимптотическая нормальность оценок наименьших модулей, вычислена относительная асимптотическая эффективность этих оценок по отношению к оценкам наименьших квадратов.*

**E-mail: vb-goryainov@mail.ru**

***Ключевые слова***: пространственная авторегрессия, оценки наименьших модулей, асимптотическая нормальность, асимптотическая относительная эффективность.

В работе изучается авторегрессионное поле $X_{ij}$, описываемое уравнением

$$X_{ij} = a_{10}X_{i-1,j}+a_{01}X_{i,j-1}+a_{11}X_{i-1,j-1}+\varepsilon_{ij}, \ \ i,j = 0, \pm 1, \pm 2, \ldots, \quad (1)$$

где $a = (a_{10}, a_{01}, a_{11})$ — вектор авторегрессионных коэффициентов, а $\varepsilon_{ij}$ — независимые одинаково распределенные случайные величины с нулевым математическим ожиданием $\mathsf{E}\varepsilon_{ij} = 0$ и конечной дисперсией $\sigma^2 = \mathsf{D}\varepsilon_{ij}$.

В большинстве работ по оцениванию вектора $a$ предполагается гауссовость поля $\varepsilon_{ij}$, а сами оценки — различными модификациями оценок максимального правдоподобия и наименьших квадратов [1, 2]. Хорошо известно [3], что оценки наименьших квадратов чувствительны к нарушению предположения нормальности $\varepsilon_{ij}$. В линейных регрессионных моделях и авторегрессионных процессах в этом случае обычно используются оценки, более устойчивые к загрязнению выборки резко выделяющимися наблюдениями, например оценки наименьших модулей [4].

В настоящей работе для параметра $a$ поля (1) строятся и изучаются оценки наименьших модулей, а именно установлена состоятельность и асимптотическая нормальность оценок наименьших модулей для $a$. Это, в частности, позволило вычислить эффективность оценок наименьших модулей по отношению к оценкам наименьших квадратов для основных типов вероятностных распределений поля $\varepsilon_{ij}$.

**Постановка задачи и формулировка основных результатов.** Рассмотрим поле (1), где $a = (a_{10}, a_{01}, a_{11})$ — неизвестный вектор параметров, а $\varepsilon_{ij}$ — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$ и плотностью $f(x)$.

Будем предполагать поле (1) стационарным. Как показано в [1], достаточным условием этого является отсутствие корней уравнения

$$1 - a_{10}z_1 - a_{01}z_2 - a_{11}z_1z_2 = 0$$

внутри единичного полидиска

$$|z_1| \leq 1, \quad |z_2| \leq 1.$$

Обозначим $\hat{a} = (\hat{a}_{10}, \hat{a}_{01}, \hat{a}_{11})^{\mathrm{T}}$ оценку наименьших модулей параметра $a$, построенную по наблюдениям $X = \{X_{ij}\}$, $i = 0, \ldots, m$, $j = 0, \ldots, n$. Другими словами, $\hat{a}$ — точка минимума функции

$$\mathcal{L}(a) = \sum_{i=1}^{m}\sum_{j=1}^{n} |X_{ij} - a_{10}X_{i-1,j} - a_{01}X_{i,j-1} - a_{11}X_{i-1,j-1}| \,. \tag{2}$$

**Теорема.** *Пусть функция распределения $F(x)$ и плотность $f(x)$ независимых одинаково распределенных случайных величин $\varepsilon_{ij}$ в (1) удовлетворяет следующим условиям*:

$$F(0) = \frac{1}{2}, \tag{3}$$

$$f(0) > 0, \quad f(x) \text{ непрерывна в нуле}, \tag{4}$$

$$\int_{-\infty}^{\infty} |f'(x)|\, dx < \infty, \tag{5}$$

$$\mathsf{E}\varepsilon_{ij} = 0, \tag{6}$$

$$\mathsf{E}\varepsilon_{ij}^2 < \infty. \tag{7}$$

*Тогда при $m, n \to \infty$ случайный вектор $\sqrt{mn}(\hat{a} - a)$ асимптотически нормален с нулевым математическим ожиданием и ковариационной матрицей $\dfrac{1}{4f^2(0)}R^{-1}$, где*

$$R = \begin{pmatrix} r_{0,0} & r_{1,-1} & r_{0,-1} \\ r_{-1,1} & r_{0,0} & r_{-1,0} \\ r_{0,1} & r_{1,0} & r_{0,0} \end{pmatrix},$$

*а $r_{\alpha-p,\beta-q} = \mathsf{E}(X_{\alpha\beta}X_{pq})$, $(\alpha, \beta) \in \mathcal{I}$, $(p, q) \in \mathcal{I}$.*

**Доказательство.** Обозначим для всех $i = 1, \ldots, m$, $j = 1, \ldots, n$

$$u_{ij} = \frac{1}{\sqrt{mn}}(b_{10}X_{i-1,j} + b_{01}X_{i,j-1} + b_{11}X_{i-1,j-1}),$$

$$v_{ij} = |\varepsilon_{ij} - u_{ij}| - |\varepsilon_{ij}|, \quad z_{ij} = u_{ij}\,\mathrm{sign}(\varepsilon_{ij}), \quad w_{ij} = v_{ij} + z_{ij}.$$

Отметим, что $\hat{a}$ минимизирует (2) тогда и только тогда, когда $\hat{b} =$

$= \sqrt{mn}(\hat{a} - a)$ минимизирует функцию

$$L(b) = \sum_{i=1}^{m}\sum_{j=1}^{n} v_{ij}.$$

Установим на множестве пар индексов $(i, j)$, $i, j \in \mathbb{Z}$, порядок, полагая $(p, q) < (i, j)$, если $q < j$ или $q = j$, но $p < i$. Обозначим $\mathfrak{A}_{ij}$ $\sigma$-алгебру, порожденную $\{\varepsilon_{pq} : (p, q) < (i, j)\}$. Учитывая, что

$$\mathsf{E}[z_{ij}] = \mathsf{E}[\mathsf{E}(z_{ij}|\mathfrak{A}_{ij})] = \mathsf{E}[u_{ij}\mathsf{E}(\operatorname{sign}(\varepsilon_{ij})|\mathfrak{A}_{ij})] = 0,$$

представим $L(b)$ в виде

$$L(b) = L_1 - L_2 + L_3,$$

где

$$L_1 = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}],\ L_2 = \sum_{i=1}^{m}\sum_{j=1}^{n} z_{ij},\ L_3 = \sum_{i=1}^{m}\sum_{j=1}^{n}(w_{ij} - \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]).$$

Обозначим

$$e_{ij}(t) = \mathsf{E}(|\varepsilon_{ij} - t| - \varepsilon_{ij}), \quad i = 1, \dots, m, j = 1, \dots, n.$$

Отметим, что

$$e_{ij}(t) = \begin{cases} \int\limits_{-\infty}^{0} tf(x)\,dx + \int\limits_{0}^{t}(t - 2x)f(x)\,dx - \int\limits_{t}^{\infty} tf(x)\,dx, & t \geq 0, \\ \int\limits_{-\infty}^{t} tf(x)\,dx + \int\limits_{0}^{t}(2x - t)f(x)\,dx - \int\limits_{0}^{\infty} tf(x)\,dx, & t < 0. \end{cases}$$

Так как $e'_{ij}(t) = 2F(t) - 1$, то с учетом $e_{ij}(0) = 0$, $e'_{ij}(0) = 0$, $e''_{ij}(0) = 2f(0)$ получим

$$e_{ij}(t) = f(\tau t)t^2, \quad \tau = \tau_{ij}(t) \in (0, 1), \quad i = 1, \dots, m, \quad j = 1, \dots, n. \tag{8}$$

Учитывая (8), независимость $\varepsilon_{ij}$ от $\mathfrak{A}_{ij}$ и измеримость $u_{ij}$ от $\mathfrak{A}_{ij}$, представим $L_1$ в виде $L_1 = L_{11} + L_{12}$, где

$$L_{11} = f(0)\sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}[u_{ij}^2], \quad L_{12} = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}[(f(\tau u_{ij}) - f(0))u_{ij}^2].$$

Из закона больших чисел [5, гл.VI, § 3] для стационарных процессов следует, что при $m, n \to \infty$ по вероятности

$$L_{11} = f(0)\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{(p,q)\in\mathcal{I}}\sum_{(\alpha,\beta)\in\mathcal{I}} \mathsf{E}[X_{i-p,j-q}X_{i-\alpha,j-\beta}b_{pq}b_{\alpha\beta}] \to$$

$$\to f(0) \sum_{(p,q)\in\mathcal{I}} \sum_{(\alpha,\beta)\in\mathcal{I}} \mathsf{E}(X_{\alpha-p,\beta-q}X_{00})b_{pq}b_{\alpha\beta} =$$

$$= f(0) \sum_{(p,q)\in\mathcal{I}} \sum_{(\alpha,\beta)\in\mathcal{I}} R_{\alpha-p,\beta-q}b_{pq}b_{\alpha\beta}.$$

Так как $f(x)$ непрерывна в нуле, семейство

$$\zeta_{mn} = |f(\tau u_{11}) - f(0)|u_{11}^2, \quad i = 1,\dots,m, j = 1,\dots,n,$$

равномерно интегрируемо и $u_{ij} \to 0$ при $m, n \to \infty$, то [5]

$$\mathsf{E}|L_{12}| \le \sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}[|f(\tau u_{ij}) - f(0)|u_{ij}^2] =$$

$$= \mathsf{E}[|f(\tau u_{11}) - f(0)|(b_{10}X_{01} + b_{01}X_{10} + b_{11}X_{00})^2] \to 0 \text{ при } m, n \to \infty,$$

т.е. $L_{12} = o_p(1)$ при $m, n \to \infty$.

Обозначим

$$\eta_{pq} = \frac{1}{\sqrt{mn}} \sum_{i=1}^{m}\sum_{j=1}^{n} b_{pq}X_{i-p,j-q}\,\mathrm{sign}(\varepsilon_{ij}), \quad (p,q)\in\mathcal{I}, \quad \eta = (\eta_{10}, \eta_{01}, \eta_{11}).$$

Тогда $L_2 = \eta^{\mathrm{T}} b$ и согласно центральной предельной теореме для мартингалов [6] вектор $\eta$ асимптотически нормален с нулевым средним и ковариационной матрицей $R$.

Далее

$$\mathsf{E}L_3 = \sum_{i=1}^{m}\sum_{j=1}^{n} (\mathsf{E}w_{ij} - \mathsf{E}(\mathsf{E}[w_{ij}|\mathfrak{A}_{ij}])) = 0.$$

Так как $\mathfrak{A}_{pq} \subseteq \mathfrak{A}_{ij}$ для всех $(p,q) < (i,j)$, то

$$\mathsf{E}\Big[\big(w_{ij} - \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]\big)|\mathfrak{A}_{pq}\Big] =$$

$$= \mathsf{E}[w_{ij}|\mathfrak{A}_{pq}] - \mathsf{E}(\mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]|\mathfrak{A}_{pq}) = \mathsf{E}[w_{ij}|\mathfrak{A}_{pq}] - \mathsf{E}[w_{ij}|\mathfrak{A}_{pq}] = 0.$$

Поэтому

$$\mathsf{E}\Big[\big(w_{pq} - \mathsf{E}[w_{pq}|\mathfrak{A}_{pq}]\big)\big(w_{ij} - \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]\big)\Big] =$$

$$= \mathsf{E}\Big[\big(w_{pq} - \mathsf{E}[w_{pq}|\mathfrak{A}_{pq}]\big)\mathsf{E}\big[\big(w_{ij} - \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]\big)|\mathfrak{A}_{pq}\big]\Big] = 0.$$

Следовательно,

$$\mathsf{E}L^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}\big(w_{ij} - \mathsf{E}[w_{ij}|\mathfrak{A}_{ij}]\big)^2 \le \sum_{i=1}^{m}\sum_{j=1}^{n} \mathsf{E}w_{ij}^2.$$

Отметим, что

$$|x-t|-|x|+t\,\mathrm{sign}(x)=\begin{cases}2(t-x), & t>x>0,\\ 2(x-t), & t<x<0,\\ 0, & \text{иначе}\ .\end{cases}$$

Поэтому

$$||x-t|-|x|+t\,\mathrm{sign}(x)|\le 2|t|I\{|t|\ge|x|\}.$$

Отсюда, из (7) и из того, что при $m,n\to\infty$ $\mathsf{P}\{|u_{11}|>\sqrt{mn}|\varepsilon_{11}|\}\to 0$, получим в силу абсолютной непрерывности интеграла Лебега

$$\mathsf{E}L_3^2\le\sum_{i=1}^{m}\sum_{j=1}^{n}\mathsf{E}(4u_{ij}^2I\{|u_{ij}|>|\varepsilon_{ij}|\})=$$

$$=4\mathsf{E}[u_{11}^2I\{|u_{11}|>\sqrt{mn}|\varepsilon_{11}|\}]\to 0,\quad m,n\to\infty.$$

Таким образом, $L_3=o_p(1)$ при $m,n\to\infty$.

Итак,

$$L=f(0)b^{\mathrm{T}}Rb+\eta^{\mathrm{T}}b+r(b),$$

где

$$r(b)=o_p(1)\text{ при } m,n\to\infty.$$

Обозначим

$$b^*=-\frac{1}{2f(0)}R^{-1}\eta.$$

Отметим, что $b^*$ асимптотически нормален с математическим ожиданием 0 и ковариационной матрицей $\dfrac{1}{4f^2(0)}R^{-1}$. Покажем, что $\hat b-b^*=o_p(1)$, откуда будет следовать утверждение теоремы. Другими словами, зафиксируем $\varepsilon>0$ и $\varepsilon_1>0$ и покажем, что существуют $m_0$ и $n_0$ такие, что для любых $m>m_0$, $n>n_0$ справедливо

$$\mathsf{P}\{|\hat b-b^*|>\varepsilon\}<\varepsilon_1.$$

Обозначим $U$ шар радиуса $\varepsilon$ с центром в $b^*$, а $S$ его границу. Так как $b^*=o_p(1)$, то существует компакт $K\in\mathbb{R}^3$ и существуют $m_1$, $n_1$ такие, что $\mathsf{P}\{U\notin K\}<\varepsilon_1$ для любых $m>m_1$, $n>n_1$. Так как $R$ положительно определена, то

$$\gamma=\inf_{|z|=\varepsilon}z^TRz>0.$$

Известно [7, 8], что сходящаяся по вероятности на выпуклом множестве последовательность выпуклых функций сходится равномерно на любом компактном подмножестве, т. е. $r(b)\to 0$ равномерно на $K$.

Поэтому существуют $m_2$, $n_2$ такие, что для любых $m > m_2$, $n > n_2$

$$\sup_{b \in K} |r(b)| < \frac{\gamma f(0)}{2}.$$

Рассмотрим луч $b(t) = b^* + zt$, $t > 0$, произвольного направления $z$, $|z| = \varepsilon$, выходящий из точки $b^*$ и пересекающий $S$ в точке $b(1)$. Так как $L(b)$ выпукла, то при $t \geq 1$

$$L(b(1)) \leq \frac{1}{t} L(b(t)) + \frac{t-1}{t} L(b^*).$$

Отметим, что

$$L(b(1)) = L(b^*) + f(0) z^T R z + r(b(1)) - r(b^*).$$

Поэтому для всех $t \geq 1$ для любых $m > m_2$, $n > n_2$

$$L(b(t)) \geq L(b^*) + t \left( f(0) z^T R z - 2 \frac{\gamma f(0)}{2} \right) \geq L(b^*).$$

Следовательно, для любых $m > m_0 = \max\{m_1, m_2\}$, $n > n_0 = \max\{n_1, n_2\}$

$$\mathsf{P}\{ \inf_{\substack{t \geq 1 \\ |z| = \varepsilon}} L(b(t)) \geq L(b^*) \} > 1 - \varepsilon_1,$$

т. е. для любых $m > m_0$, $n > n_0$

$$\mathsf{P}\{ |\hat{b} - b^*| < \varepsilon \} \geq 1 - \varepsilon_1.$$

Теорема доказана.

Отметим, что условия теоремы справедливы для основных типов вероятностных распределений $\varepsilon_{ij}$ — нормального, логистического, двойного экспоненциального (Лапласа), равномерного, Стьюдента с числом степеней свободы, бóльшим, чем 2.

**Сравнение оценок наименьших модулей и наименьших квадратов при нарушении предположения нормальности инновационного процесса $\boldsymbol{\varepsilon}_{ij}$.** Оценка наименьших квадратов $\tilde{a}$ определяется как точка минимума функции

$$\Lambda(a) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left( X_{ij} - a_{10} X_{i-1,j} - a_{01} X_{i,j-1} - a_{11} X_{i-1,j-1} \right)^2.$$

В [1] показано, что вектор $\sqrt{mn}(\tilde{a} - a)$ асимптотически нормален с нулевым математическим ожиданием и ковариационной матрицей $\sigma^2 R^{-1}$, где $\sigma^2$ — дисперсия $\varepsilon_{ij}$.

Согласно асимптотической теории оптимальности [9, гл. 5, 6] из двух состоятельных асимптотически нормальных оценок лучшей будет оценка с меньшей асимптотической дисперсией, а отношение этих дисперсий называется асимптотической относительной эффективностью и служит количественной мерой сравнения оценок.

В многомерном случае при сравнении оценок векторного параметра описанный выше подход применим, если матрицы ковариаций оценок пропорциональны друг другу, при этом асимптотическая относительная эффективность определяется как коэффициент пропорциональности соответствующих ковариационных матриц. Следовательно, асимптотическая относительная эффективность $e(f)$ оценок наименьших модулей $\hat{a}$ по отношению к оценкам наименьших квадратов $\tilde{a}$ равна

$$e(f) = 4f^2(0)\sigma^2.$$

Таким образом, для того чтобы получить то же самое предельное распределение для оценок наименьших модулей требуется в $e(f)$ раз больше наблюдений, чем для оценок наименьших квадратов.

Если распределение $\varepsilon_{ij}$ нормальное, т.е.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}},$$

то $e(f) = \dfrac{2}{\pi}$, а значит, оценки наименьших квадратов приблизительно в полтора раза эффективнее оценок наименьших модулей.

Если $\varepsilon_{ij}$ имеет двустороннее показательное распределение

$$f(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}|x|}{\sigma}}, \quad 0 \le \gamma \le 1,$$

то $e(f) = 2$, т. е. оценки наименьших модулей в два раза эффективнее оценок наименьших квадратов.

Пусть теперь $\varepsilon_{ij}$ имеют распределение Тьюки с плотностью

$$f(x) = (1-\gamma)\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} + \gamma \frac{1}{\sqrt{2\pi}\tau} e^{-\frac{x^2}{2\tau^2}}.$$

Распределение Тьюки моделирует засорение нормального инновационного поля $\varepsilon_{ij}$ некоторой долей $\gamma$ грубых ошибок, представляющих собой также нормальные случайные величины, но с большей дисперсией $\tau^2$. В этом случае

$$\sigma^2 = (1-\gamma) + \gamma\tau^2,$$

$$4f^2(0) = \frac{2}{\pi}\left(1 - \gamma + \frac{\gamma}{\tau}\right).$$

Поэтому

$$e(f) = \frac{2}{\pi}(1 - \gamma + \gamma/\tau)(1 - \gamma + \gamma\tau^2).$$

С ростом $\tau$ значение $e(f)$ неограниченно возрастает. Таким образом, асимптотическая относительная эффективность оценок наименьших модулей по отношению к оценкам наименьших квадратов может быть сколь угодно велика.

Например, $e(f) > 1$ для $\tau = 3$ и $\gamma = 0{,}1$, т.е. если среди нормальных случайных величин $\varepsilon_{ij}$ каждая десятая в три раза "грубее" остальных, то оценки наименьших модулей по эффективности уже превосходят оценки наименьших квадратов.

**Заключение.** Оценки наименьших модулей параметров пространственной авторегрессионной модели являются состоятельными и асимптотически нормальными. Они значительно (примерно в полтора раза) уступают в эффективности оценкам наименьших квадратов, если наблюдения авторегрессионного поля являются нормальными и измеряются безошибочно. Однако оценки наименьших модулей демонстрируют лучшую эффективность по отношению к оценкам наименьших квадратов (которая может быть сколь угодно большой), если плотность распределения вероятности инновационного поля имеет пик в нуле или сравнительно медленно убывает на бесконечности (имеет "тяжелые хвосты").

## СПИСОК ЛИТЕРАТУРЫ


1. T j o s t h e i m D. Statistical spatial series modelling // Advances in Applied Probability. – 1978. – V. 10. No. 1. – P. 130–154.
2. D a v y d o v Y., P a u l a u s k a s V. On estimation of parameters for spatial autoregressive model // Statistical Inference for Stochastic Processes. – 2008. – V. 11. No. 3. – P. 237–247.
3. M a r o n n a R. A., M a r t i n D., Y o h a i V. Robust statistics. Theory and methods. – Wiley, 2006. – P. 403.
4. B l o o m f i e l d P. Least absolute deviations: Theory, applications, and algorithms. – Birkhauser, 1983. – P. 364.
5. Ш и р я е в А. Н. Вероятность. – М.: Наука. Физ.-мат. лит., 1980. – 576 с.
6. M c L e i s h D. L. Dependent central limit theorems and invariance principles // The Annals of Probability. – 1974. – V. 2. – No 4. – P. 620–628.
7. A n d e r s e n P. K., G i l l R. Cox's regression model for counting processes: A large sample study // Annals of Statistics. – 1982. – V. 10. – P. 1100–1120.
8. P o l l a r d D. Asymptotics for least absolute deviation regression estimators // Econometric Theory. – 1991. – V. 7, No. 2. – P. 186–199.
9. Л е м а н Э. Теория точечного оценивания: Пер. с англ. – М.: Наука. Гл. ред. физ.-мат. лит., 1991. – С. 448.



Статья поступила в редакцию 15.07.2010



Владимир Борисович Горяинов — канд. физ.-мат. наук, доцент кафедры "Математическое моделирование" МГТУ им. Н.Э. Баумана.

V.B. Goryainov — Ph. D. (Phys.-Math.), assoc. professor of "Mathematical Simulation" department of the Bauman Moscow State Technical University.