А. В. Горяинов, В. Б. Горяинов, М-оценки параметров процесса авторегрессии со случайными коэффициентами, *Автомат. и телемех.*, 2018, выпуск 8, 50–65

https://www.mathnet.ru/at15122

Автоматика и телемеханика, № 8, 2018

© 2018 г. А.В. ГОРЯИНОВ, канд. физ.-мат. наук (agoryainov@gmail.com)
(Московский авиационный институт),
В.Б. ГОРЯИНОВ, д-р физ.-мат. наук (vb-goryainov@mail.ru)
(Московский государственный технический университет им. Н.Э. Баумана)

М-ОЦЕНКИ ПАРАМЕТРОВ ПРОЦЕССА АВТОРЕГРЕССИИ СО СЛУЧАЙНЫМИ КОЭФФИЦИЕНТАМИ

Доказана асимптотическая нормальность М-оценок авторегрессионных параметров уравнения авторегрессии со случайными коэффициентами. Для уравнения первого порядка изложен метод вычисления асимптотической относительной эффективности М-оценки с аналитической ρ-функцией по отношению к оценке наименьших квадратов. Метод основан на разложении асимптотической дисперсии М-оценки в сходящийся ряд. Показано, что М-оценка эффективнее оценки наименьших квадратов, если обновляющий процесс имеет загрязненное гауссовское распределение.

Ключевые слова: авторегрессионная модель со случайными коэффициентами, оценка наименьших квадратов, М-оценка, асимптотическая относительная эффективность, распределение Тьюки.

1. Введение

В классической теории временны́х рядов основное место занимают гауссовские линейные модели, в частности модели авторегрессии, скользящего среднего и смешанные модели авторегрессии—скользящего среднего [1, 2]. Эти модели позволили описать поведение целого ряда случайных процессов в самых разнообразных областях науки и техники. Однако существует множество примеров стохастических систем, поведение которых не укладывается в линейную схему. Поэтому в конце XX в. появилось большое число нелинейных моделей, с помощью которых удалось описать эффекты, необъяснимые в рамках линейной теории [3]. Так, с помощью нелинейных моделей были построены лучшие, чем в линейных моделях, прогнозы финансово-экономических показателей [4], численности популяций в модели хищник—жертва [5], числа солнечных пятен [6] и изменения речного стока [7].

Одной из нелинейных моделей временны́х рядов является модель авторегрессии со случайными коэффициентами (RCA-модель), разнообразные приложения которой имеются, например, в [8–11]. Согласно модели RCA(p) временной ряд X_t удовлетворяет уравнению (см., например, [12]) вида

$$X_t = (\varphi_1 + \eta_{t1})X_{t-1} + (\varphi_2 + \eta_{t2})X_{t-2} + \cdots + (\varphi_p + \eta_{tp})X_{t-p} + \varepsilon_t, \tag{1}$$

в котором $\varphi = (\varphi_1, \ldots, \varphi_p)^{\mathrm{T}}$ — вектор неизвестных неслучайных параметров, $\eta_t = (\eta_{t1}, \ldots, \eta_{tp})^{\mathrm{T}}$, а $(\eta_t^{\mathrm{T}}, \varepsilon_t)^{\mathrm{T}}$ — последовательность независимых одинаково распределенных случайных векторов с нулевым математическим ожиданием $\mathsf{E}\eta_t = 0$, $\mathsf{E}\varepsilon_t = 0$. Основная задача при анализе такого уравнения состоит в оценивании параметра φ.

Традиционным методом оценивания φ является метод наименьших квадратов [13], хотя при отклонении распределения ε_t от гауссовского предпочтительнее использовать метод наименьших модулей [14]. Вместо метода максимального правдоподобия из-за сложного вида функции правдоподобия для большинства распределений η_t и ε_t в модели (1) используется метод квазимаксимального правдоподобия [15], в котором функция правдоподобия строится в предположении гауссовости η_t и ε_t. Эти методы при умеренных предположениях о вероятностных свойствах авторегрессионных коэффициентов $\varphi + \eta_t$ и обновляющего процесса ε_t позволяют получить состоятельные и асимптотически нормальные оценки (см. [13–15]).

В частном случае модели (1), когда авторегрессионные коэффициенты не случайные ($\eta_t = 0$), известны оценки, которые при определенных обстоятельствах обладают более высокой эффективностью, чем оценка наименьших квадратов. Например, M-оценка [16] предпочтительнее оценки наименьших квадратов, если распределение вероятности обновляющего процесса ε_t авторегрессионного уравнения отклоняется (в определенном смысле) от гауссовского.

По аналогии с M-оценкой параметров уравнения авторегрессии с неслучайными коэффициентами определим M-оценку $\hat{\varphi}_n$ параметра φ уравнения (1) по наблюдениям $X_{1-p}, \dots, X_n$ как точку минимума функции

$$\mathcal{L}(\varphi) = \sum_{t=1}^{n} \rho(X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p}), \tag{2}$$

где ρ — некоторая функция. Заметим, что при $\rho(x) = x^2$ эта оценка совпадает с оценкой наименьших квадратов, а при $\rho(x) = |x|$ — с оценкой наименьших модулей.

Данная статья посвящена изучению асимптотических свойств указанной M-оценки в модели RCA(p) и сравнительному анализу асимптотической относительной эффективности (АОЭ) этой оценки по отношению к оценке наименьших квадратов для некоторых вероятностных распределений процесса ε_t и случайного коэффициента η_t в модели RCA(1). Важно отметить, что результаты исследований АОЭ оценок в модели авторегрессии с неслучайными коэффициентами не переносятся на модель RCA(1) автоматически. Это связано с тем, что для вычисления асимптотической дисперсии оценки коэффициента φ в модели RCA(1) нужно уметь находить распределение вероятности X_t при известных распределениях ε_s и η_s, $s \leqslant t$, что не представляется возможным из-за сложной зависимости X_t от ε_s и η_s, $s \leqslant t$, вызванной нелинейностью модели (1). В статье предлагается преодолеть это затруднение с помощью представления асимптотической дисперсии M-оценки в виде функции от моментов $\mathsf{E}X_t^n$, которые в свою очередь выражаются через моменты $\mathsf{E}\eta_1^i$ и $\mathsf{E}\varepsilon_1^i$, $i = 1, 2, \dots, n$.

Статья имеет следующую структуру. В разделе 2 изложена постановка задачи. В разделе 3 доказана асимптотическая нормальность M-оценки с выпуклой четной функцией $\rho(x)$. В разделе 4 для модели RCA(1) предложен метод вычисления АОЭ M-оценки с аналитической ρ-функцией по отношению к оценке наименьших квадратов. Метод основан на разложении ρ-функции в сходящийся ряд. В разделе 5 приведен численный пример. В разделе 6 рассмотрено уравнение RCA(1) с обновляющим процессом ε_t, описываемым загрязненным усеченным гауссовским распределением, и проведено исследование зависимости АОЭ от параметров загрязняющего распределения. Показано, что АОЭ М-оценки по отношению к оценке наименьших квадратов растет с увеличением как дисперсии засоряющего распределения, так и доли засорения.

2. Постановка задачи

Рассмотрим временной ряд X_t, $t = 0, \pm 1, \pm 2, \ldots$, удовлетворяющий разностному уравнению (1), в котором $(\eta_t^{\mathrm{T}}, \varepsilon_t)^{\mathrm{T}}$, $t = 0, \pm 1, \pm 2, \ldots$ — последовательность независимых одинаково распределенных случайных векторов с нулевым математическим ожиданием $\mathsf{E}\eta_t = 0$, $\mathsf{E}\varepsilon_t = 0$ и ковариационной матрицей

$$\begin{pmatrix} \Sigma_\eta & \sigma_{\eta\varepsilon} \\ \sigma_{\eta\varepsilon}^{\mathrm{T}} & \sigma_\varepsilon^2 \end{pmatrix},$$

где Σ_η — ковариационная матрица размеров $p \times p$ вектора η_t, $\sigma_\varepsilon^2 = \mathsf{D}\varepsilon_t$ — дисперсия ε_t, $\sigma_{\eta\varepsilon} = (\mathsf{cov}(\eta_{t1}, \varepsilon_t), \ldots, \mathsf{cov}(\eta_{tp}, \varepsilon_t))^{\mathrm{T}}$.

Далее всюду предполагается, что случайный процесс X_t, $t = 0, \pm 1, \pm 2, \ldots$, является эргодическим. Достаточные условия эргодичности определяются (см. [13]) следующим образом. Введем матрицы

$$\Phi = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{p-1} & \varphi_p \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}, \qquad H_t = \begin{pmatrix} \eta_{t1} & \eta_{t2} & \cdots & \eta_{tp} \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

размеров $p \times p$ и векторы $X(t) = (X_t, X_{t-1}, \ldots, X_{t-p+1})^{\mathrm{T}}$, $U_t = (\varepsilon_t, 0, \ldots, 0)^{\mathrm{T}}$ размера $p \times 1$. Если все собственные числа матрицы $\mathsf{E}(H_t \bigotimes H_t) + \Phi \bigotimes \Phi$ размеров $p^2 \times p^2$ по модулю меньше единицы, то стационарное и эргодическое решение уравнения (1) существует и имеет вид

$$X(t) = U_t + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} (\Phi + H_{t-i}) \right) U(t-j).$$

Для авторегрессионного уравнения первого порядка ($p = 1$) указанное достаточное условие стационарности и эргодичности превращается в неравенство $\varphi^2 + \mathsf{D}(\eta_t) < 1$.

Обозначим через $\mathfrak{F}_t$ σ-алгебру событий, порожденную множеством случайных величин $\{\eta_s, \varepsilon_s, \ s \leqslant t\}$. Оценка наименьших квадратов $\tilde{\varphi}_n$ параметра φ по наблюдениям $X_0, X_1 \ldots, X_n$ случайного процесса X_t, описываемого уравнением (1), определяется (см. [17]) как точка минимума функции

$$L_{LS}(\varphi) = \sum_{t=1}^{n} (X_t - \mathsf{E}(X_t | \mathfrak{F}_{t-1}))^2,$$

где $\mathsf{E}(X_t|\mathfrak{F}_{t-1})$ — условное математическое ожидание $\mathsf{E}X_t$ относительно σ-алгебры $\mathfrak{F}_{t-1}$. Поскольку $\mathsf{E}(X_t|\mathfrak{F}_{t-1}) = \varphi^{\mathrm{T}} X(t-1)$, то

$$L_{LS}(\varphi) = \sum_{t=1}^{n} \left(X_t - \varphi^{\mathrm{T}} X(t-1) \right)^2 . \tag{3}$$

В [13] доказано, что оценка наименьших квадратов является состоятельной и асимптотически нормальной в случае, когда $\mathsf{E}(X_t^4) < \infty$ и все собственные числа матрицы $\mathsf{E}(H_t \bigotimes H_t) + \Phi \bigotimes \Phi$ по модулю меньше единицы.

Однако наряду с указанными важными свойствами состоятельности и асимптотической нормальности оценка наименьших квадратов имеет и существенный недостаток, а именно: в ситуациях, когда среди X_t имеются выбросы (резко выделяющиеся наблюдения), функция (3), а следовательно, и ее точка минимума $\tilde{\varphi}_n$ будут претерпевать значительные искажения. Это обстоятельство обусловлено тем, что L_{LS} является квадратичной функцией невязки $X_t - \varphi^{\mathrm{T}} X(t-1)$.

Для большинства статистических моделей качество оценивания их параметров можно повысить, если в (2) функцию $\rho(x) = x^2$ заменить функцией, которая, будучи четной, является либо ограниченной, либо растет на бесконечности медленнее, чем x^2. Наиболее распространенной ρ-функцией является (см. [18]) ρ-функция Хьюбера

$$\rho_H(x) = \begin{cases} x^2, & \text{если } |x| \leqslant k, \\ 2k|x| - k^2, & \text{если } |x| > k. \end{cases} \tag{4}$$

Эта функции зависит от параметра $k > 0$, изменение которого позволяет регулировать устойчивость М-оценки к сильно выделяющимся наблюдениям. В частности, при приближении k к бесконечности точка минимума (2) в большинстве моделей имеет свойства, схожие со свойствами оценки наименьших квадратов, а при приближении k к нулю — со свойствами оценки наименьших модулей.

3. Асимптотическая нормальность М-оценки

Основным результатом статьи является доказательство асимптотической нормальности М-оценок $\hat{\varphi}_n$, приведенное в следующей теореме. Обозначим $B = \mathsf{E}\left[\rho''(\eta_1^{\mathrm{T}} X_0 + \varepsilon_1) X_0 X_0^{\mathrm{T}}\right]$.

Теорема. *Пусть собственные числа матрицы $\mathsf{E}(H_t \bigotimes H_t) + \Phi \bigotimes \Phi$ по модулю меньше единицы, $\mathsf{E}\eta_t = 0$, $\mathsf{E}\varepsilon_t = 0$, $\mathsf{E}|\eta_t|^4 < \infty$, $\mathsf{E}\varepsilon_t^4 < \infty$, $\mathsf{E}|X_0|^4 < \infty$, плотность распределения вероятности случайного вектора $(\eta_t^{\mathrm{T}}, \varepsilon_t)$ — четная функция, функция $\rho(x)$ — четная и выпуклая, а ее вторая производная $\rho''(x)$ ограничена и почти всюду непрерывна.*

Тогда случайная величина $\sqrt{n}(\hat{\varphi}_n - \varphi)$ является асимптотически нормальной с нулевым математическим ожиданием и ковариационной матрицей

$$B^{-1} \mathsf{E}\left[\left(\rho'\left(\eta_1^{\mathrm{T}} X(0) + \varepsilon_1 \right) \right)^2 X(0) X^{\mathrm{T}}(0) \right] (B^{-1})^{\mathrm{T}}. \tag{5}$$

Доказательство теоремы приведено в Приложении.

Нетрудно показать, что условия теоремы выполнены для ρ-функции Хьюбера (4) и (при достаточно малых φ и $\mathsf{E}|\eta_t|^4$) — для основных вероятностных распределений, в частности гауссовского, логистического и Стьюдента с числом степеней свободы, большим четырех.

С л е д с т в и е. *Пусть в уравнении* (1) $p = 1$ *и* $\varphi^2 + \mathsf{E}\eta_t^2 < 1$.

Тогда в условиях теоремы случайная величина $\sqrt{n}(\hat{\varphi}_n - \varphi)$ *является асимптотически нормальной с нулевым математическим ожиданием и дисперсией*

$$\frac{\mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2]}{(\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2])^2}. \tag{6}$$

4. Асимптотическая относительная эффективность

При сравнении между собой двух оценок неизвестного параметра целесообразно считать лучшей ту оценку, которая имеет меньшее рассеяние вокруг оцениваемого параметра. Для асимптотически несмещенных оценок скалярного параметра естественной мерой рассеяния служит асимптотическая дисперсия. Именно поэтому сравнительной количественной характеристикой точности таких оценок служит обратное отношение их асимптотических дисперсий, называемое асимптотической относительной эффективностью (АОЭ) [19, § 5.2].

Пусть X_t описывается уравнением RCA(1)

$$X_t = (\varphi + \eta_t)X_{t-1} + \varepsilon_t, \qquad t = 0, \pm 1, \pm 2, \dots, \tag{7}$$

и выполнено условие $\varphi^2 + \sigma_\eta^2 < 1$, где σ_η^2 — дисперсия η_t. Тогда существует решение уравнения (7), которое представляется в виде сходящегося с вероятностью единица ряда

$$X_t = \varepsilon_t + \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} (\varphi + \eta_{t-j})\, \varepsilon_{t-i}. \tag{8}$$

В этом случае АОЭ М-оценки $\hat{\varphi}_n$ по отношению к оценке наименьших квадратов $\tilde{\varphi}_n$ равна $e = \lim_{n\to\infty} \frac{\mathsf{D}\tilde{\varphi}_n}{\mathsf{D}\hat{\varphi}_n}$. Для простоты предположим, что случайные последовательности η_t и ε_t независимы. Подставляя $\rho(x) = x^2$ в (6) и учитывая независимость X_0, η_1 и ε_1 между собой, получим, что оценка наименьших квадратов асимптотически нормальна с асимптотической дисперсией

$$\mathsf{D}\tilde{\varphi}_n = \frac{\sigma_\eta^2 \mathsf{E}X_0^4 + \sigma_\varepsilon^2 \mathsf{E}X_0^2}{(\mathsf{E}X_0^2)^2}. \tag{9}$$

Асимптотическая дисперсия М-оценки задана выражением (6). Поэтому АОЭ М-оценки по отношению к оценке наименьших квадратов равна

$$e = \frac{(\sigma_\eta^2 \mathsf{E}X_0^4 + \sigma_\varepsilon^2 \mathsf{E}X_0^2)(\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2])^2}{(\mathsf{E}X_0^2)^2 \mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2]}. \tag{10}$$

Если $\eta_1 = 0$ (что соответствует модели авторегрессии с детерминированным коэффициентом), то (10) в силу независимости ε_1 от X_0 превращается в

$$e = \frac{\sigma_\varepsilon^2 (\mathsf{E}[\rho''(\varepsilon_1)])^2}{\mathsf{E}[(\rho'(\varepsilon_1))^2]}.$$

Таким образом, в этом частном случае найти e, зная вид функции ρ и распределение вероятности ε_1, можно либо аналитически, либо при помощи приближенных методов вычисления соответствующих интегралов.

В общем случае для нахождения e в RCA-модели нужно знать распределение X_0, найти которое из-за сложной зависимости X_0 от η_t и ε_t, $t \leqslant 0$, не представляется возможным. Например, из гауссовости η_t и ε_t не следует гауссовость X_t.

Эту трудность можно преодолеть, предположив, что функция ρ аналитическая, а распределения η_t и ε_t таковы, что справедливы разложения:

$$\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1) X_0^2] = \sum_{n=0}^{\infty} \frac{\rho^{(n+2)}(0)}{n!} \mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^n X_0^2], \tag{11}$$

$$\mathsf{E}[\rho'(\eta_1 X_0 + \varepsilon_1) X_0] = \sum_{n=0}^{\infty} \frac{\rho^{(n+1)}(0)}{n!} \mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^n X_0]. \tag{12}$$

В этом случае из независимости η_1, ε_1 и X_0 следует, что

$$\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^n X_0^2] = \sum_{i=0}^{n} C_n^i \mathsf{E}\eta_1^i \mathsf{E}X_0^{i+2} \mathsf{E}\varepsilon_1^{n-i},$$

$$\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^n X_0] = \sum_{i=0}^{n} C_n^i \mathsf{E}\eta_1^i \mathsf{E}X_0^{i+1} \mathsf{E}\varepsilon_1^{n-i}.$$

Поэтому вычисление АОЭ сводится к рекуррентному выражению моментов $\mathsf{E}X_0^n$ через моменты $\mathsf{E}\varepsilon_1^n$ и $\mathsf{E}\eta_1^n$, $n = 1, 2, \ldots$ Из (7) вытекает, что

$$\mathsf{E}X_t^n = \mathsf{E}\left((\varphi + \eta_t) X_{t-1} + \varepsilon_t\right)^n.$$

Так как $\mathsf{E}\eta_t^n$, $\mathsf{E}\varepsilon_t^n$ и $\mathsf{E}X_t^n$ в силу стационарности η_t^n, ε_t^n и X_t^n не зависят от t, то

$$\mathsf{E}X_0^n = \sum_{i=0}^{n} \mathsf{E}(\varphi + \eta_1)^i \mathsf{E}X_0^i \mathsf{E}\varepsilon_1^{n-i}.$$

Отсюда

$$\mathsf{E}X_0^n = \frac{1}{1 - \mathsf{E}(\varphi + \eta_1)^n} \sum_{i=0}^{n-1} C_n^i \mathsf{E}X_0^i \mathsf{E}\varepsilon_1^{n-i} \mathsf{E}(\varphi + \eta_1)^i. \tag{13}$$

К сожалению, вычислить АОЭ при помощи (11)–(12) нельзя в важном частном случае, когда η_1 и ε_1 являются гауссовскими случайными величинами с нулевым средним. Действительно, тогда для любого четного $n = 2m$ имеем:

$$\mathsf{E}\eta_1^{2m} = (m-1)!!\sigma_\eta^{2m}, \qquad \mathsf{E}\varepsilon_1^{2m} = (m-1)!!\sigma_\varepsilon^{2m}, \qquad \mathsf{E}\eta_1^{2m-1} = \mathsf{E}\varepsilon_1^{2m-1} = 0,$$

где $(m-1)!! = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (m-1)$. Поэтому $\mathsf{E}\eta_1^{2m} \to \infty$ при $m \to \infty$ и $1 - \mathsf{E}(\varphi + \eta_1)^{2m} < 0$ при всех достаточно больших m, а сумма в правой части (13) всегда положительна. Таким образом, $\mathsf{E}X_0^{2m} < 0$ для достаточно больших m, что невозможно.

В большинстве приложений использование гауссовских случайных величин обосновывается центральной предельной теоремой, в которой речь идет о бесконечной сумме случайных величин. На практике сумма всегда конечна, поэтому предположение о гауссовости выполняется лишь приближенно. На фоне этого несоответствия между реальным явлением и его математической моделью гауссовскую модель для большинства приложений практически без всякого ущерба можно заменить усеченной гауссовской моделью, в которой η_t и ε_t имеют распределения с плотностями $f_\eta(x) = f(x, k_\omega, \omega^2)$ и $f_\varepsilon(x) = f(x, k_\sigma, \sigma^2)$ соответственно, где

$$f(x,l,d) = \begin{cases} \dfrac{1}{2\Phi_0(l)\sqrt{2\pi d}} e^{-\frac{x^2}{2d}}, & \text{если } |x| \leqslant l\sqrt{d}, \\ 0, & \text{если } |x| > l\sqrt{d}, \end{cases} \tag{14}$$

$\Phi_0(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$ — функция Лапласа, а ω, k_ω, σ и k_σ — некоторые положительные постоянные. Выбирая k_ω и k_σ достаточно большими, можно для большинства приложений достаточно хорошо аппроксимировать гауссовские распределения случайных величин η_t, ε_t их усеченными версиями. Поэтому вычисление АОЭ, если ε_t и η_t имеют усеченные гауссовские распределения, является важным для изучения применения М-оценок на практике.

Покажем, что если распределения η_t и ε_t имеют вид (14), то АОЭ М-оценки уже можно вычислить при помощи разложений (11)–(12). В этом случае $\mathsf{E}\eta_1^n = \mathsf{E}\varepsilon_1^n = 0$ для нечетных n. Отсюда и из (8) вытекает, что $\mathsf{E}X_0^n = 0$ для таких n. Кроме того,

$$\mathsf{E}\varepsilon_1^{2n} \leqslant \int_{-\sigma k_\sigma}^{\sigma k_\sigma} (\sigma k_\sigma)^{2n} \frac{1}{2\Phi_0(k_\sigma)\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}\, dx = (\sigma k_\sigma)^{2n}. \tag{15}$$

Аналогичные формулы справедливы и для $\mathsf{E}\eta_1^n$, в частности, $\mathsf{E}\eta_1^{2n} \leqslant (\omega k_\omega)^{2n}$. Далее

$$\mathsf{E}(\varphi + \eta_1)^n = \sum_{i=0}^{n} C_n^i \varphi^i \mathsf{E}\eta_1^{n-i}, \tag{16}$$

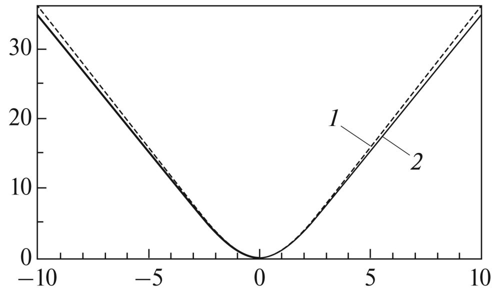

Рис. 1. ρ-функции М-оценок: *1* — функция Хьюбера $\rho_H(x)$; *2* — функция $\rho_A(x)$.

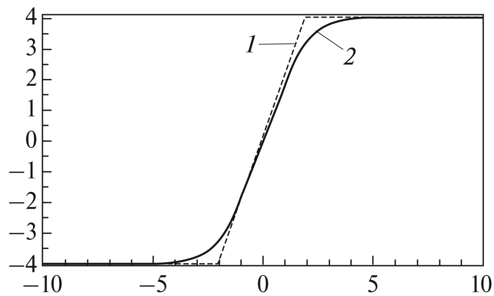

Рис. 2. Производные ρ-функций М-оценок: *1* — производная функции Хьюбера $\rho_H(x)$; *2* — производная функции $\rho_A(x)$.

откуда

$$(17)\qquad |\mathsf{E}(\varphi+\eta_1)^n| \leqslant \sum_{i=0}^{n} C_n^i \left|\mathsf{E}\eta_1^i\right| |\varphi|^{n-i} \leqslant (|\varphi| + \omega k_\omega)^n.$$

Из неравенств (15) и (17), представления (8) и независимости случайных величин $\{\eta_t, \varepsilon_t, t = 0, \pm 1, \pm 2, \dots\}$ следует, что при $|\varphi| + \omega k_\omega < 1$

$$(18)\qquad \mathsf{E}X_0^{2n} \leqslant \frac{(\sigma k_\sigma)^{2n}}{(1 - |\varphi| - \omega k_\omega)^{2n}}.$$

Отсюда вытекает сходимость (11)–(12), если η_t и ε_t имеют усеченные гауссовские распределения, а ρ — аналитическая функция.

На практике при построении М-оценок параметров различных статистических моделей наиболее распространенной ρ-функцией является ρ-функция $\rho_H(x)$ Хьюбера (4). Популярность ρ-функции Хьюбера обусловлена ее выпуклостью и линейным ростом на бесконечности. Выпуклость обеспечивает единственность минимума целевой функции (2), а линейный (не квадратичный) рост на бесконечности — преимущество М-оценки над оценкой наименьших квадратов при отклонении вероятностного распределения наблюдений от гауссовского. Например, в авторегрессионной модели с неслучайными коэффициентами (модель (1) с $\eta_t = 0$ для любых t) и негауссовским распределением обновляющего процесса ε_t М-оценка обычно эффективнее оценки наименьших квадратов [16].

К сожалению, функция $\rho_H(x)$ не является аналитической, что не позволяет использовать ее в разложениях (11)–(12). Тем не менее разумно предположить, что если вместо $\rho_H(x)$ взять аналитическую четную выпуклую растущую линейно на бесконечности ρ-функцию, то М-оценка в RCA-модели (7) будет эффективнее оценки наименьших квадратов для многих вероятностных распределений ε_t, в частности, для тех распределений, для которых М-оценка эффективнее оценки наименьших квадратов в авторегрессионной модели с неслучайным коэффициентом.

Исследуем свойства М-оценки параметра φ в (7), построенной при помощи аналитической ρ-функции

$$\rho_A(x) = 4kx\Phi_0\left(\frac{x\sqrt{\pi}}{k\sqrt{2}}\right) + \frac{4k^2}{\pi}e^{-\frac{\pi x^2}{4k^2}} - \frac{4k^2}{\pi}, \tag{19}$$

которая как и ρ-функция Хьюбера $\rho_H(x)$ является выпуклой и растущей линейно на бесконечности. Графики функций $\rho_H(x)$ и $\rho_A(x)$ для $k = 2$ приведены на рис. 1, а графики их производных — на рис. 2.

Применяя разложения (11) и (12) к функции $\rho_A(x)$, имеем:

$$\mathsf{E}\left[\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2\right] = 2\sum_{n=0}^{\infty}\frac{(-\pi)^n \mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^{2n}X_0^2]}{n!(4k^2)^n}, \tag{20}$$

$$\mathsf{E}\left[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2\right] = \tag{21}$$

$$= 4\sum_{n=1}^{\infty}\left(\sum_{i=0}^{n}\frac{1}{(i-1)!(n-i)!(2i-1)(2(n-i)+1)}\right)\frac{(-\pi)^{n-1}\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^{2n}X_0^2]}{(4k^2)^{n-1}},$$

где

$$\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^{2n}X_0^2] = \sum_{i=0}^{2n} C_{2n}^i \mathsf{E}\eta_1^i \mathsf{E}X_0^{i+2}\mathsf{E}\varepsilon_1^{2n-i}.$$

Легко видеть, что

$$\sum_{i=0}^{n}\frac{1}{i!(n-i)!(2i+1)(2(n-i)+1)} \leqslant \frac{2^n}{n!},$$

$$\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^{2n}X_0^2] \leqslant \frac{(\sigma k_\sigma)^{2n+2}(1-|\varphi|)^{2n}}{(1-|\varphi|-\omega k_\omega)^{2n+2}}.$$

Поэтому правые части (20) и (21) при рассматриваемом в статье ограничении $|\varphi| + \omega k_\omega < 1$ сходятся, что позволяет вычислить АОЭ М-оценки, построенной при помощи ρ-функции $\rho_A(x)$, сколь угодно точно.

5. Численный пример

В линейных регрессионных моделях и в моделях авторегрессии с неслучайными коэффициентами М-оценки обычно являются более точными, чем оценки наименьших квадратов в случае негауссовских ошибок наблюдений и негауссовского обновляющего процесса [16]. В частности, метод наименьших квадратов достаточно чувствителен даже к небольшим отклонениям распределения наблюдений от гауссовского.

Одной из распространенных моделей отклонения распределения от гауссовского является распределение Тьюки, или “загрязненное гауссовское распределение” (см. [18, с. 49]), плотность которого имеет вид

$$f(x) = (1-\gamma)\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{x^2}{2\tau^2}} + \gamma\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}, \qquad 0 \leqslant \gamma \leqslant 1, \qquad \sigma > \tau. \tag{22}$$

Последовательность случайных величин, имеющих распределение Тьюки, имитирует типичное на практике засорение последовательности центрированных гауссовских величин с дисперсией τ^2 небольшой (0,01–0,15) долей γ центрированных гауссовских величин с дисперсией $\sigma^2 > \tau^2$.

К сожалению, если ε_1 будет иметь распределение (22), то ряды (20) и (21) при достаточно больших значениях величины $\sigma_\varepsilon^2 = (1-\gamma)\tau^2 + \gamma\sigma^2$ будут расходиться даже в тех случаях, когда η_1 имеет усеченное гауссовское распределение. Это следует из того, что при $n \to \infty$ величина $\mathsf{E}[(\eta_1 X_0 + \varepsilon_1)^{2n} X_0^2]$ имеет порядок $n!\sigma_\varepsilon^{2n}$.

Поэтому предположим, что ε_1 имеет распределение вида

$$f_\varepsilon(x) = (1-\gamma)g_1(x) + \gamma g_2(x), \tag{23}$$

где $g_1(x) = f(x, k_\tau, \tau^2)$ и $g_2(x) = f(x, k_\sigma, \sigma^2)$ соответственно, $f(x, l, d)$ определена в (14), а k_τ, τ, k_σ, σ — некоторые положительные постоянные. Назовем такое распределение усеченным распределением Тьюки. Для достаточно больших k_τ и k_σ в большинстве приложений распределением (23) можно безболезненно заменить распределение (22).

Пример 1. Предположим, что случайные величины η_t имеют усеченное гауссовское распределение (14) с плотностью $f_\eta(x) = f(x, k_\omega, \omega^2)$, а члены ε_t обновляющей последовательности имеют усеченное распределение Тьюки (23). Далее пусть, например, $\varphi = 0{,}2$, $\omega = 0{,}01$, $k_\omega = k_\tau = k_\sigma = 3$, $\tau = 1$, $\sigma = 3$ и $\gamma = 0{,}1$. Вычислим АОЭ М-оценки по отношению к оценке наименьших квадратов, которая по определению равна величине, определяемой выражением (10). Используя разложения (20)–(21), получим, что $e \approx 1{,}336$. Таким образом, в этом случае М-оценка превосходит оценку наименьших квадратов и оценке наименьших квадратов для достижения точности М-оценки необходимо в 1,336 больше наблюдений.

Этот результат справедлив, если число наблюдений n авторегрессионного процесса X_t достаточно велико, поскольку он получен при ковариационной матрице (5), имеющей асимптотический характер. Поэтому чтобы понять, насколько надежно пользоваться асимптотическими результатами, относительная эффективность для различных n оценивалась при помощи компьютерного моделирования. Для каждого n компьютерный эксперимент проводился $N = 50 \cdot 10^3$ раз, реализации $X_1, \ldots, X_n$ длины n процесса X_t моделировались при помощи датчика псевдослучайных чисел по рекуррентной формуле (1) с начальным условием $X_0 = 0$. Дисперсии оценок $\hat{\varphi}_n$ и $\tilde{\varphi}_n$ параметра φ оценивались выборочными дисперсиями по N реализациям, относительная эффективность М-оценки по отношению к оценке наименьших квадратов аппроксимировалась обратным отношением выборочных дисперсий. По результатам моделирования для $n = 50, 100, 200, 300, 400$ относительная эффективность оказалась равной соответственно 1,182, 1,253, 1,316, 1,328 и 1,336. Таким образом, результаты моделирования достаточно хорошо согласуются с асимптотическим результатом.

6. Исследование АОЭ при нарушении предположения о гауссовости обновляющего процесса

Поскольку ρ-функция $\rho_A(x)$ (как и ρ-функция $\rho_H(x)$) зависит от параметра k, то М-оценки на основе $\rho_A(x)$ образуют целое семейство оценок. Параметр k является фиксированным и не определяется величинами $X_0, X_1 \ldots, X_n$ случайного процесса (7), так как выбирается еще до начала эксперимента по наблюдению за процессом X_t. Оценим зависимость АОЭ от параметра k, что позволит получить рекомендации по его выбору.

Если ε_1 имеет распределение вида (23), η_1 имеет усеченное гауссовское распределение, а функция ρ имеет вид (19), то согласно (10) АОЭ является функцией от φ, ω, k_ω, k_τ, τ, k_σ, σ, γ и k. Зафиксировав для определенности значения $\varphi = 0{,}2$, $\omega = 0{,}01$, $k_\omega = k_\tau = k_\sigma = 3$, $\tau = 1$ и $\gamma = 0{,}02$, рассмотрим АОЭ как функцию параметра k ρ-функции $\rho_A(x)$ при различных значениях дисперсии σ^2 засоряющего распределения. На рис. 3 приведены графики зависимости $e = e(k)$ для $\sigma = 1, 2, 3, 4$, которые показывают, что лучшим выбором значения параметра k для ρ-функции $\rho_A(x)$ будет число близкое к двум.

Теперь, положив параметр ρ-функции $k = 2$, исследуем зависимость АОЭ от параметров σ и γ отклонения распределения обновляющего процесса ε_t от

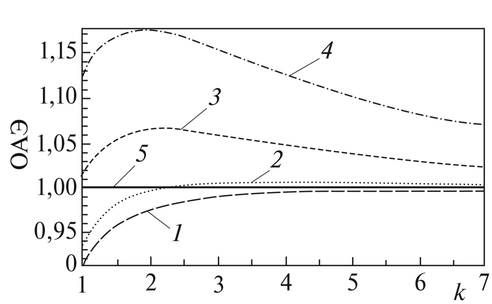

Рис. 3. Зависимость АОЭ от k для разных σ: *1* — $\sigma = 1$; *2* — $\sigma = 2$; *3* — $\sigma = 3$; *4* — $\sigma = 4$; *5* — АОЭ $= 1$.

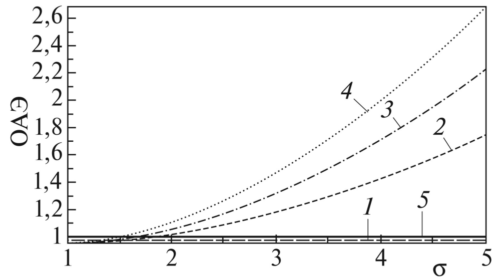

Рис. 4. Зависимость АОЭ от σ для разных значений γ: *1* — $\gamma = 0$; *2* — $\gamma = 0{,}05$; *3* — $\gamma = 0{,}1$; *4* — $\gamma = 0{,}2$; *5* — АОЭ $= 1$.

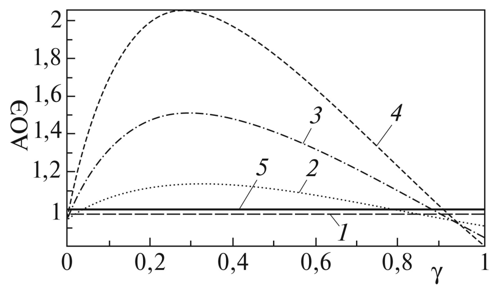

Рис. 5. Зависимость АОЭ от γ для разных σ: *1* — $\sigma = 1$; *2* — $\sigma = 2$; *3* — $\sigma = 3$; *4* — $\sigma = 4$; *5* — АОЭ $= 1$.

усеченного гауссовского распределения в модели (23). Параметры φ, ω, k_ω, k_τ, k_σ и τ оставим на прежнем уровне.

Сначала исследуем зависимость АОЭ от σ, построив графики зависимости $e = e(\sigma)$ при различных уровнях засорения γ. На рис. 4 приведены графики $e = e(\sigma)$ для уровней загрязнения $\gamma = 0, 0{,}05, 0{,}1, 0{,}2$. Случай $\gamma = 0$ сводит распределение (23) обновляющего процесса к усеченному гауссовскому распределению, и в этой ситуации оценка наименьших квадратов является более эффективной, чем М-оценка. Для значений $\gamma > 0$ М-оценка превосходит оценку наименьших квадратов тем сильнее, чем больше дисперсия σ^2 и доля γ загрязнения.

Теперь исследуем зависимость АОЭ от γ при фиксированных σ. На рис. 5 приведена зависимость $e = e(\gamma)$ для значений $\sigma = 1, 2, 3, 4$. Видно, что с ростом доли загрязнения γ АОЭ М-оценки увеличивается, становясь с увеличением σ больше единицы. Парадоксальность уменьшения АОЭ при дальнейшем росте γ только кажущаяся. Дело в том, что при $\gamma > 0{,}5$ роль загрязнения начинают играть гауссовские величины с дисперсией τ^2, а не σ^2 и случай $\gamma = 1$, как и случай $\gamma = 0$, соответствует отсутствию загрязнений. Отметим, что наибольшее значение АОЭ соответствует не $\gamma = 0{,}5$, а меньшему значению γ, так как при $\gamma > 0{,}5$ дисперсия загрязнений меньше дисперсии основной доли случайных величин.

Исследование зависимости АОЭ от φ и от дисперсии σ_η^2 при ограничении $|\varphi| + \omega k_\omega < 1$ показало, что всегда $e < 1$.

7. Заключение

В статье доказана асимптотическая нормальность M-оценок параметров авторегрессионного уравнения со случайными коэффициентами. Разработан метод вычисления асимптотической относительной эффективности M-оценок, построенных при помощи аналитических ρ-функций, по отношению к оценке наименьших квадратов. Для аналитического варианта ρ-функции Хьюбера исследована зависимость асимптотической относительной эффективности от параметров усеченной версии загрязненного гауссовского распределения обновляющего шума. Показано, что уже при небольшом отклонении распределения от усеченного гауссовского M-оценки становятся эффективнее оценки наименьших квадратов.

ПРИЛОЖЕНИЕ

Доказательство теоремы. Доказательство разобьем на три шага. Сначала приблизим целевую функцию $\mathcal{L}(\varphi)$ квадратичной формой $\widetilde{\mathcal{L}}(\varphi)$, раскладывая $\mathcal{L}(\varphi)$ по формуле Тейлора в окрестности φ. Затем установим асимптотическую нормальность точки минимума $\tilde{\varphi}_n$ функции $\tilde{\mathcal{L}}(\varphi)$. И докажем сходимость к нулю по вероятности последовательности $\sqrt{n}(\hat{\varphi}_n - \tilde{\varphi}_n)$.

Обозначим $\zeta_t = \eta_t^{\mathrm{T}} X(t-1) + \varepsilon_t$. Заметим, что $\hat{\varphi}_n$ является точкой минимума $\mathcal{L}(\varphi)$ тогда и только тогда, когда $\hat{b}_n = \sqrt{n}(\hat{\varphi}_n - \varphi)$ — точка минимума функции

$$L_n(b) = \sum_{t=1}^{n} \rho\left(\zeta_t - \frac{b^{\mathrm{T}} X(t-1)}{\sqrt{n}}\right) - \sum_{t=1}^{n} \rho\left(\zeta_t\right).$$

Раскладывая выражение $L_n(b)$ в окрестности нуля по степеням b, получим, что

$$L_n(b) = -A_n^{\mathrm{T}} b + \frac{1}{2} b^{\mathrm{T}} B_n b + \alpha_n(b),$$

где

$$A_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \rho'(\zeta_t) X(t-1), \qquad B_n = \frac{1}{n} \sum_{t=1}^{n} \rho''(\zeta_t) X(t-1) X^{\mathrm{T}}(t-1),$$

$$\alpha_n(b) = \frac{1}{2n} \sum_{t=1}^{n} \left(\rho''\left(\zeta_t - \frac{\tau b^{\mathrm{T}} X(t-1)}{\sqrt{n}}\right) - \rho''(\zeta_t)\right) b^{\mathrm{T}} X(t-1) X^{\mathrm{T}}(t-1) b,$$
$$0 < \tau < 1.$$

Случайные последовательности

$$\zeta_{1t} = \rho''(\zeta_t) X(t-1) X^{\mathrm{T}}(t-1), \qquad t = 0, \pm 1, \pm 2, \ldots,$$

и

$$\zeta_{2t} = \left(\rho''\left(\zeta_t - \frac{\tau b^{\mathrm{T}} X(t-1)}{\sqrt{n}}\right) - \rho''(\zeta_t)\right) b^{\mathrm{T}} X(t-1) X^{\mathrm{T}}(t-1) b,$$
$$t = 0, \pm 1, \pm 2, \dots,$$

являются стационарными и эргодическими, как преобразования последовательности $(\eta_t^{\mathrm{T}}, \varepsilon_t)$ одинаково распределенных и независимых случайных векторов (см. [20, с. 170, 182]). Из стационарности ζ_{2t} следует, что

$$|\alpha_n(b)| \leqslant \frac{1}{2} \mathsf{E}\left[\left|\rho''\left(\zeta_1 - \frac{\tau b^{\mathrm{T}} X_0}{\sqrt{n}}\right) - \rho''(\zeta_1)\right| |X_0|^2 |b|^2\right].$$

Поскольку функция ρ'' ограничена и непрерывна почти всюду, то согласно теореме Лебега о мажорируемой сходимости $\mathsf{E}|\alpha_n(b)| \to 0$ при $n \to \infty$. Отсюда в силу неравенства Чебышева следует, что $\alpha_n(b) = o_p(1)$ при $n \to \infty$ (здесь и в дальнейшем $o_p(1)$ означает последовательность случайных величин, сходящихся к нулю по вероятности). Из эргодичности ζ_{1t} следует ([20, с. 181]), что существует предел по вероятности

$$B = \lim_{n\to\infty} B_n = \mathsf{E}[\rho''(\zeta_1) X_0 X_0^{\mathrm{T}}].$$

Поэтому

$$L_n(b) = -A_n^{\mathrm{T}} b + \frac{1}{2} b^{\mathrm{T}} B b + \gamma_n(b),$$

где $\gamma_n(b) = o_p(1)$ при $n \to \infty$.

Обозначим через $\tilde{b}_n = B^{-1} A_n$ точку минимума функции $\tilde{L}_n(b) = -A_n^{\mathrm{T}} b + \frac{1}{2} b^{\mathrm{T}} B b$. Покажем, что последовательность $\tilde{b}_n$ асимптотически нормальна и $\hat{b}_n - \tilde{b}_n = o_p(1)$, что повлечет асимптотическую нормальность последовательности $\hat{b}_n$.

Так как плотность распределения вероятности случайного вектора $(\eta_t^{\mathrm{T}}, \varepsilon_t)$ — четная функция, то условная плотность распределения вероятности случайной величины ζ_t при условии $\mathfrak{F}_{t-1}$ — тоже четная. Поскольку ρ', как производная четной функции, является нечетной, то $\mathsf{E}[\rho'(\zeta_t) X(t-1) | \mathfrak{F}_{t-1}] = 0$. Поэтому по центральной предельной теореме для мартингалов (см. [21]) последовательность A_n является асимптотически нормальной с нулевым математическим ожиданием и ковариационной матрицей $\mathsf{E}[(\rho'(\zeta_1))^2 X(0) X^{\mathrm{T}}(0)]$. Следовательно, последовательность $\tilde{b}_n$ является асимптотически нормальной с нулевым математическим ожиданием и ковариационной матрицей (5). Существование этой матрицы вытекает из ограниченности $\rho''(x)$ и условия $\mathsf{E}X_0|^4 < \infty$.

Теперь докажем, что $\hat{b}_n - \tilde{b}_n = o_p(1)$ при $n \to \infty$, т.е. для любого $\varepsilon > 0$ справедливо $\mathsf{P}\{|\hat{b}_n - \tilde{b}_n| > \varepsilon\} \to 0$ при $n \to \infty$. Другими словами, нужно доказать, что для любого $\varepsilon_1 > 0$ существует $N(\varepsilon_1)$, такое что для любого $n > N(\varepsilon_1)$ выполнено неравенство

$$\mathsf{P}\left\{|\hat{b}_n - \tilde{b}_n| > \varepsilon\right\} < \varepsilon_1. \tag{П.1}$$

Так как последовательность $\{\tilde{b}_n\}$ асимптотически нормальна с ковариационной матрицей (5), то она ограничена по вероятности. Поэтому существует компакт $K \subset \mathbb{R}^p$ такой, что все шары $U_\varepsilon(\tilde{b}_n) = \{b \in \mathbb{R}^p : |b - \tilde{b}_n| \leqslant \varepsilon\}$, $n \in \mathbb{N}$, радиуса ε с центром в $\tilde{b}_n$ лежат в K с вероятностью $1 - \varepsilon_1/2$. Таким образом, по формуле полной вероятности

$$\mathsf{P}\Big\{|\hat{b}_n - \tilde{b}_n| > \varepsilon\Big\} < \mathsf{P}\Big\{|\hat{b}_n - \tilde{b}_n| > \varepsilon \Big| U_\varepsilon(\tilde{b}_n) \subset K\Big\} + \frac{\varepsilon_1}{2}.$$

Так как $L_n(b)$ выпукла, то выпуклой будет и последовательность функций $L_n(b) + A_n^{\mathrm{T}} b$, которая при $n \to \infty$ сходится по вероятности к выпуклой функции $\frac{1}{2} b^{\mathrm{T}} B b$. Поэтому (см. [22]) для любого компакта из $\mathbb{R}^p$, в частности, для K, при $n \to \infty$

$$m_n(K) = \sup_{b \in K} |\gamma_n(b)| = o_p(1). \tag{П.2}$$

Заметим, что $\tilde{L}_n(\tilde{b}_n + b) - \tilde{L}_n(\tilde{b}_n) = \frac{1}{2} b^{\mathrm{T}} B b$ для любого $b \in \mathbb{R}^p$. Поэтому

$$L_n(\tilde{b}_n + b) - L_n(\tilde{b}_n) > \frac{1}{2} b^{\mathrm{T}} B b - 2m_n(K).$$

Пусть $b = ce$, где $c \geqslant \varepsilon_1$, а e — произвольный p-мерный единичный вектор, т.е. $(\tilde{b}_n + b) \notin U_\varepsilon(\tilde{b}_n)$. Тогда $L_n(\tilde{b}_n + b) - L_n(\tilde{b}_n) > \frac{1}{2} c^2 e^{\mathrm{T}} B e - 2m_n(K)$. Из (П.2) вытекает, что можно выбрать $N(\varepsilon_1)$ так, чтобы для любого $n > N(\varepsilon_1)$ с вероятностью $1 - \varepsilon_1/2$ выполнялось неравенство $\frac{1}{2} c^2 e^{\mathrm{T}} B e - 2m_n(K) > 0$. Поэтому для любых $n > N(\varepsilon_1)$, $c \geqslant \varepsilon_1$ и $e \in \mathbb{R}^p$, $|e| = 1$ с вероятностью $1 - \varepsilon_1/2$ выполняется неравенство $L_n(\tilde{b}_n + ce) > L_n(\tilde{b}_n)$, т.е. минимум $L_n(b)$ с вероятностью $1 - \varepsilon_1/2$ лежит внутри $U_\varepsilon(\tilde{b}_n)$. Следовательно,

$$\mathsf{P}\left\{|\hat{b}_n - \tilde{b}_n| > \varepsilon | U_\varepsilon(\tilde{b}_n) \in K\right\} < \frac{\varepsilon_1}{2},$$

откуда справедливо (П.1). Теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Андерсон Т.* Статистический анализ временных рядов. М.: Мир, 1976.
2. *Хеннан Э.* Многомерные временные ряды. М.: Мир, 1974.
3. *Tong H.* Nonlinear time series. A dynamical system approach. Oxford: Oxford University Press, 1990.
4. *Tsay R.S.* Analysis of financial time series. Hoboken: Wiley, 2010.
5. *Tong H.* Some Comments on the Canadian Lynx Data // J. Roy. Statist. Soc. Ser. A. 1977. V. 140. P. 432–436.
6. *Subba Rao T.* On the Theory of Bilinear Time Series Models // J. Roy. Statist. Soc. Ser. B. 1981. V. 43. No. 2. P. 244–255.
7. *Tong H., Lim K.S.* Threshold Autoregression, Limit Cycles and Cyclical Data // J. Roy. Statist. Soc. Ser. B. 1980. V. 42. No. 3. P. 245–292.
8. *Singpurwalla N.D., Soyer R.* Assessing Software Reliability Growth Using a Random Coefficient Autoregressive Process and its Ramifications // IEEE Trans. Software Engrg. 1985. V. SE–11. No. 12. P. 1456–1464.

9. *Ghirmai T.* A Random-Coefficient Third-Order Autoregressive Process // Digit. Signal Process. 2015. V. 38. P. 25–46.
10. *Lee H.T., Yoder J.K., Mittelhammer R.C., et al.* A Random Coefficient Autoregressive Markov Regime Switching Model for Dynamic Futures Hedging // J. Futures Market. 2006. V. 26. No. 2. P. 103–129.
11. *Tang D., Yu J., Chen X., Makis V.* An Optimal Condition-based Maintenance Policy for a Degrading System Subject to the Competing Risks of Soft and Hard Failure // Computers & Industrial Engineering. 2015. V. 83. No. 1. P. 100–110.
12. *Nicholls D.F., Quinn B.G.* Random coefficient autoregressive models: an introduction. N.Y.: Springer, 1982.
13. *Hwang S.Y., Basawa I.V.* Parameter Estimation for Generalized Random Coefficient Autoregressive Processes // J. Statist. Plann. Inference. 1998. V. 68. No. 2. P. 323–337.
14. *Горяинов А.В., Горяинова Е.Р.* Сравнение эффективности оценок методов наименьших модулей и наименьших квадратов в авторегрессионной модели со случайным коэффициентом // АиТ. 2016. № 9. С. 84–95.

    Goryainov A.V., Goryainova E.R. Comparison of Efficiency of Estimates by the Methods of Least Absolute Deviations and Least Squares in the Autoregression Model with Random Coefficient // Autom. Remote Control. 2016. V. 77. No. 9. P. 1579–1588.
15. *Aue A., Horváth L., Steinebach J.* Estimation in Random Coefficient Autoregressive Models // J. Time Ser. Anal. 2006. V. 27. N 1. P. 61–76.
16. *Maronna R.A., Martin D., Yohai V.* Robust Statistics: Theory and Methods. Chichester: Wiley, 2006.
17. *Tjøstheim D.* Estimation in Nonlinear Time Series Models // Stochast. Process. Appl. 1986. V. 21. No. 2. P. 251–273.
18. *Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В.А.* Робастность в статистике. Подход на основе функций влияния. М.: Мир, 1989.
19. *Леман Э.* Теория точечного оценивания. М.: Наука, 1991.
20. *Stout W.F.* Almost sure convergence. N.Y.: Acad. Press, 1974.
21. *Ибрагимов И.А.* Центральная предельная теорема для одного класса зависимых случайных величин // Теор. вероят. и ее применен. 1963. Т. 8. Вып. 1. С. 89–94.
22. *Andersen P.K., Gill R.D.* Cox's Regression Model for Counting Processes: a Large Sample Study //Ann. Statist. 1982. V. 10. No. 4. P. 1100–1120.

Статья представлена к публикации членом редколлегии А.И. Кибзуном.

Поступила в редакцию 30.03.2017