УДК 519.234 DOI: 10.18698/1812-3368-2019-5-4-18

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ЭКСПОНЕНЦИАЛЬНОЙ АВТОРЕГРЕССИОННОЙ МОДЕЛИ

В.Б. Горяинов vb-goryainov@bmstu.ru
В.М. Кайнг waimyokhing@gmail.com

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Цель работы — сравнение оценки наименьших квадратов и оценки наименьших модулей в зависимости от распределения вероятности обновляющего процесса авторегрессионного уравнения. Для достижения этой цели с помощью компьютерного моделирования многократно воспроизведена последовательность наблюдений экспоненциального авторегрессионного процесса, по каждой последовательности вычислены оценка наименьших квадратов и оценка наименьших модулей. Полученные последовательности оценок использованы для вычисления выборочных дисперсий оценки наименьших квадратов и оценки наименьших модулей. Лучшей полагалась оценка с наименьшей выборочной дисперсией. Количественной мерой сравнения оценок выступала выборочная относительная эффективность оценок, определяемая как обратное отношение их выборочных дисперсий. В качестве моделей распределения вероятности обновляющего процесса использованы нормальное распределение, загрязненное нормальное распределение (распределение Тьюки) с различными значениями доли и величины загрязнения, логистическое распределение, распределение Лапласа и распределение Стьюдента с различным числом степеней свободы (в частности, с одной степенью свободы, т. е. распределение Коши). Для каждого распределения вероятности получены асимптотические значения выборочной относительной эффективности при неограниченном увеличении объема выборки наблюдений авторегрессионного процесса. Оказалось, что оценка наименьших модулей лучше оценки наименьших квадратов для распределения Лапласа и загрязненного нормального распределения с достаточно большими уровнями доли и величины загрязнения. В остальных случаях предпочтительнее оценка наименьших квадратов

Ключевые слова

Экспоненциальная авторегрессия, оценка наименьших квадратов, оценка наименьших модулей

Поступила 09.04.2019

Введение. Долгое время основным инструментом для анализа временных рядов были линейные модели, в частности линейные авторегрессионные модели [1]. Линейные модели и сейчас нередко используют на начальной стадии изучения временных рядов, поскольку являются удобными инструментами для прогнозирования и управления. Однако в большинстве случаев временные ряды, встречающиеся в различных областях науки техники, обнаруживают свойства, которые невозможно описать, оставаясь в рамках линейных моделей. Понимая это, исследователи предложили несколько нелинейных моделей временных рядов [2]. Одна из самых удачных нелинейных моделей — модель экспоненциальной авторегрессии [3], которая, в частности, описывает предельные циклы, резонансные скачки, зависимость частоты колебаний от амплитуды, что невозможно в рамках линейной модели. Эта модель применяется во многих областях науки и техники, например, при моделировании нелинейных периодических процессов [4, 5], при моделировании качки корабля [3], в экономике [6], биологии [3], океанологии [7], климатологии [8]. При изучении экспоненциальных авторегрессионных моделей важной задачей является оценивание коэффициентов соответствующего авторегрессионного уравнения. Существуют различные методы оценивания коэффициентов, например, с помощью фильтра Калмана [7].

Настоящая работа посвящена оценкам наименьших модулей и наименьших квадратов. *Цель работы* — сравнение точности этих оценок в зависимости от распределения вероятности обновляющего процесса авторегрессионного уравнения.

Постановка задачи. Экспоненциальная авторегрессионная модель первого порядка описывается рекуррентным уравнением

$$X_t = \left(a_0 + b_0 e^{-c_0 X_{t-1}^2}\right) X_{t-1} + \varepsilon_t, \tag{1}$$

где a_0, b_0, c_0 — действительные числа, являющиеся параметрами модели; ε_t, $t = 1, 2, \ldots$ — обновляющий процесс, представляющий собой последовательность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием $\mathrm{E}\varepsilon_t = 0$ и конечной дисперсией $\mathrm{D}\varepsilon_t = \mathrm{E}\varepsilon_t^2 = \sigma^2 < \infty$.

Рассмотрим задачу оценивания вектора параметров (a_0, b_0, c_0) по наблюдениям $X_1, X_2, \ldots, X_n$ процесса X_t, удовлетворяющего этому уравнению. Наиболее распространенными оценками являются оценки наименьших квадратов $(\tilde{a}, \tilde{b}, \tilde{c})$ и наименьших модулей (a^*, b^*, c^*), определяемые как точки минимума функций

$$g_{\text{н.к}}(a,b,c)=\sum_{t=2}^{n}\left(X_t-\left(a+be^{-cX_{t-1}^2}\right)X_{t-1}\right)^2$$

и

$$g_{\text{н.м}}(a,b,c)=\sum_{t=2}^{n}\left|X_t-\left(a+be^{-cX_{t-1}^2}\right)X_{t-1}\right|.$$

Сравним между собой точности этих оценок в зависимости от распределения вероятности членов ε_t обновляющей последовательности.

Метод сравнения оценок. В статистике случайных процессов существуют различные методы сравнения качества оценок. Так, можно изучать поведение функционала влияния оценок при возмущении наблюдений $X_1, X_2, \ldots, X_n$, когда оценки могут утрачивать свойство состоятельности [9]. Здесь используем другой подход, связанный со сравнением дисперсий оценок [10].

Точность любой оценки $\hat{\theta}$ скалярного параметра θ распределения вероятности произвольной случайной величины естественно измерять ее абсолютным отклонением от оцениваемого параметра $|\hat{\theta}-\theta|$. Поскольку любая оценка является случайной величиной, разность $\hat{\theta}-\theta$ также случайна. В связи с этим величину $|\hat{\theta}-\theta|$ разумно усреднить и измерять точность оценки математическим ожиданием $\mathrm{E}|\hat{\theta}-\theta|$ или, что более удобно, величиной $\mathrm{E}(\hat{\theta}-\theta)^2$. Отметим, что для несмещенных оценок, т. е. таких, что $\mathrm{E}\hat{\theta}=\theta$, величина $\mathrm{E}(\hat{\theta}-\theta)^2$ совпадает с дисперсией $\mathrm{D}\hat{\theta}$ оценки $\hat{\theta}$.

Из двух оценок $\hat{\theta}_1$ и $\hat{\theta}_2$ параметра θ лучшей будем полагать более точную оценку, т. е. ту, отклонение которой от оцениваемого параметра является меньшим, а количественно измерять преимущество одной оценки, например, $\hat{\theta}_1$ по отношению к другой $\hat{\theta}_2$ будем отношением $e(\hat{\theta}_1,\hat{\theta}_2)=\dfrac{\mathrm{E}(\hat{\theta}_2-\theta)^2}{\mathrm{E}(\hat{\theta}_1-\theta)^2}$, которое назовем относительной эффективностью $\hat{\theta}_1$ по отношению к $\hat{\theta}_2$.

Обычно точность $\mathrm{E}(\hat{\theta}-\theta)^2$ любой оценки $\hat{\theta}$ произвольного параметра θ с ростом числа наблюдений n стремится к нулю пропорционально величине $1/n$ [11]. В этом случае относительная эффективность $e(\hat{\theta}_1,\hat{\theta}_2)$ показывает, во сколько раз больше наблюдений необходимо оценке $\hat{\theta}_2$ для достижения точности оценки $\hat{\theta}_1$. Например, $e(\hat{\theta}_1,\hat{\theta}_2)=2$ означает, что для достижения одинаковой точности оценке $\hat{\theta}_2$ требуется в 2 раза больше наблюдений, чем оценке $\hat{\theta}_1$.

Моделирование распределения обновляющего процесса. Относительная эффективность зависит от распределения вероятности случайных величин ε_t. Поэтому целесообразно знать ее для основных наиболее распространенных вероятностных распределений. Аргументы для выбора тестовых распределений следующие.

Нормальное, или гауссово, распределение вероятности, имеющее плотность $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, является наиболее распространенным при моделировании стохастических процессов, что обусловлено центральной предельной теоремой теории вероятностей. Однако именно в силу предельного характера этой теоремы на практике распределение вероятности практически всегда в той или иной степени отклоняется от нормального.

Наиболее распространенная модель такого отклонения — приближенное нормальное распределение вероятности, или распределение Тьюки [12] с плотностью

$$f(x)=(1-\gamma)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\gamma\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{x^2}{2\tau^2}},\quad \tau>1,\ \ 0<\gamma<1,$$

представляющее собой смесь с весами $1-\gamma$ и γ двух нормальных распределений с нулевым математическим ожиданием — стандартного нормального (с единичной дисперсией) и с дисперсией $\tau^2>1$. Замена стандартного нормального распределения приближенным нормальным означает, что в последовательности ε_t, $t=1,2,\ldots$, стандартных нормальных величин случайно с вероятностью γ элементы последовательности заменяются нормальными случайными величинами с бóльшей дисперсией τ^2. Эта модель, также называемая загрязненным или засоренным нормальным распределением, хорошо описывает аномальные выбросы в наблюдениях, связанные, например, со сбоями измерительной аппаратуры или с систематическим иррегулярным внешним воздействием на процесс. Величина γ показывает долю загрязнения, а величина τ — силу загрязнения.

Еще один способ моделирования приближенного нормального распределения заключается в замене его распределением Стьюдента с плотностью

$$f(x)=\frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)\left(1+\frac{x^2}{m}\right)^{\frac{m+1}{2}}},$$

где m — параметр распределения, называемый числом степеней свободы. Распределение Стьюдента с большим числом степеней свободы практически неотличимо от нормального, однако с уменьшением m оно все больше и больше отклоняется от нормального вплоть до того, что при $m=2$ у распределения Стьюдента отсутствует дисперсия, а при $m=1$ (называемого в этом случае распределением Коши) — еще и математическое ожидание.

Кроме нормального распределения, в приложениях естественным образом возникает и распределение Лапласа. Это происходит в случае, когда разумно предполагать, что нормальным является условное распределение величин ε_t, дисперсия которых случайна. Например, точность измерений может определяться степенью запыленности атмосферы. Если полагать известным только среднее значение дисперсии, а в остальном ее положение на числовой оси полагать максимально неопределенным, то методами теории информации можно получить [13], что безусловное распределение ε_t будет двусторонним экспоненциальным распределением, или распределением Лапласа с плотностью $f(x)=\frac{1}{\sqrt{2}\sigma}e^{-\frac{\sqrt{2}|x|}{\sigma}}$.

Нередко нормальное распределение можно перепутать с логистическим распределением с плотностью $f(x)=\frac{e^{-x}}{(1+e^{-x})^2}$ (в частности, их нередко путает критерий согласия К. Пирсона).

В связи с изложенным выше сравним с помощью компьютерного моделирования оценку наименьших квадратов и оценку наименьших модулей в случае, когда распределение вероятности ε_t является нормальным, загрязненным нормальным, логистическим, Лапласа и Стьюдента.

Описание компьютерного эксперимента. Вектор наблюдений X_1, $X_2,\ldots,X_n$ генерировался $N=1000$ раз, его каждая реализация $x_{i1},x_{i2},\ldots$, x_{in}, $i=1,2,\ \ldots,\ N$, получена с помощью рекуррентного соотношения (1) с нулевым начальным условием $X_0=0$, параметры в (1) для определенности полагались $a_0=-0,3$, $b_0=-0,8$, $c_0=1$, n = 50, 100, 200, 500. Случайные величины ε_t получены с использованием датчиков псевдослучайных чисел программы *MATLAB*. Поскольку оценки $(\tilde{a},\tilde{b},\tilde{c})$ и (a^*,b^*,c^*) векторные, они сравнивались покоординатно. Для определенности рассмотрим оценки $\tilde{b}$ и b^*. Для каждой i-й реализации $x_{i1},x_{i2},\ldots,x_{in}$, $i=1,2,\ldots,N$, путем минимизации функций $g_{\text{н.к}}(a,b,c)$ и $g_{\text{н.м}}(a,b,c)$ строились оценки $\tilde{b}_i$ и b_i^*, $i=1,\ldots,N$, и относительная эффективность $e(\tilde{b},b^*)$ оценивалась величиной

$$e_N(\tilde{b}, b^*) = \frac{\sum_{i=1}^{N}(b_i^* - b_0)^2}{\sum_{i=1}^{N}(\tilde{b}_i - b_0)^2}.$$

Функции $g_{\text{н.к}}(a,b,c)$ и $g_{\text{н.м}}(a,b,c)$ минимизировались с помощью алгоритма Левенберга — Марквардта [14], суть которого заключается в комбинации метода Ньютона с методом градиентного спуска.

Результаты моделирования и их обсуждение. Результаты моделирования приведены в таблице.

Оценка $e_N(\tilde{b}, b^*)$ относительной эффективности метода наименьших квадратов по отношению к методу наименьших модулей при различных распределениях ε_t

Распределение ε_t	n				
	50	100	200	500	∞
Нормальное	1,666	1,540	1,480	1,501	1,571
Логистическое	1,208	1,317	1,182	1,293	1,216
Тьюки:					
$\gamma = 0{,}01,\ \tau = 3$	1,575	1,560	1,453	1,443	1,474
$\gamma = 0{,}01,\ \tau = 10$	0,762	0,808	0,786	0,8882	0,804
$\gamma = 0{,}1,\ \tau = 3$	1,005	0,999	0,947	1,049	1,002
$\gamma = 0{,}1,\ \tau = 10$	0,171	0,169	0,161	0,174	0,174
Стьюдента:					
$m = 15$	1,314	1,468	1,382	1,369	1,407
$m = 10$	1,324	1,368	1,359	1,303	1,321
$m = 5$	1,010	1,074	1,097	0,977	1,041
$m = 4$	0,896	0,906	0,880	0,904	0,889
Лапласа	0,678	0,620	0,613	0,559	0,500
Коши	$3{,}527 \cdot 10^{-5}$	$1{,}827 \cdot 10^{-4}$	$5{,}315 \cdot 10^{-5}$	$1{,}187 \cdot 10^{-5}$	0

В столбцах со второго по пятый для определенности приведены оценки относительной эффективности $\tilde{b}$ по отношению к b^* (ситуация с оцениванием a_0 и c_0 аналогичная). Погрешность метода статистических испытаний с увеличением числа испытаний N уменьшается достаточно медленно (пропорционально $1/\sqrt{N}$). Это, а также погрешности численного метода минимизации функций $g_{\text{н.к}}(a,b,c)$ и $g_{\text{н.м}}(a,b,c)$ приводят к большому разбросу значений в ячейках таблицы. Тем не менее метод наименьших квадратов эффективнее метода наименьших модулей для нормального и логистического распределений, распределения Стьюдента с большим числом степеней свободы и для распределения Тьюки с небольшими значениями γ и τ. Метод наименьших модулей предпочтительнее метода наименьших квадратов для распределения Стьюдента с небольшим числом степеней свободы и для распределения Лапласа. Крайний правый столбец таблицы содержит асимптотические значения $e_N(\tilde{b}, b^*)$ при $n, N \to \infty$, полученные в следующем разделе, которые хорошо согласуются с результатами моделирования.

Асимптотическое поведение оценок. Анализ таблицы показывает, что величины $e_N(\tilde{b}, b^*)$ монотонно изменяются с ростом n. Это наводит на мысль о существовании предельного значения величины $e_N(\tilde{b}, b^*)$ при $n, N \to \infty$. Покажем, как можно найти этот предел.

Аппроксимируем функцию $g_{\text{н.к}}(a,b,c)$, обозначив для простоты $g(a,b,c)$, ее разложением по формуле Тейлора в точке (a_0, b_0, c_0) до второго порядка включительно. Обозначим

$$\theta = \left(\sqrt{n}(a-a_0), \sqrt{n}(b-b_0), \sqrt{n}(c-a_c)\right),$$

$$A = \frac{1}{\sqrt{n}}\left(\frac{\partial g(a_0,b_0,c_0)}{\partial a}, \frac{\partial g(a_0,b_0,c_0)}{\partial b}, \frac{\partial g(a_0,b_0,c_0)}{\partial c}\right)^{\text{т}},$$

$$B = \frac{1}{n}\begin{pmatrix} \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial a^2} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial a \partial b} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial a \partial c} \\ \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial a \partial b} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial b^2} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial b \partial c} \\ \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial a \partial c} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial b \partial c} & \dfrac{\partial^2 g(a_0,b_0,c_0)}{\partial c^2} \end{pmatrix}.$$

В этих обозначениях

$$g(a,b,c) = g(a_0,b_0,c_0) + A^{\text{т}}\theta + \frac{1}{2}\theta^{\text{т}} B\theta + \delta(a,b,c),$$

где $\delta(a,b,c)$ — бесконечно малая функция более высокого порядка при $(a,b,c) \to (a_0,b_0,c_0)$, чем $(a-a_0)^2 + (b-b_0)^2 + (c-c_0)^2$.

Непосредственное дифференцирование приводит к

$$A = \frac{1}{\sqrt{n}}\left(-2\sum_{t=2}^{n}\varepsilon_t X_{t-1},\ -2\sum_{t=2}^{n}\varepsilon_t X_{t-1} e^{-c_0 X_{t-1}^2},\ 2b_0\sum_{t=2}^{n}\varepsilon_t X_{t-1}^3 e^{-c_0 X_{t-1}^2}\right)^{\text{т}},$$

$$\frac{\partial^2 g(a_0,b_0,c_0)}{\partial a^2} = 2\sum_{t=2}^{n} X_{t-1}^2,\ \frac{\partial^2 g(a_0,b_0,c_0)}{\partial a \partial b} = 2\sum_{t=2}^{n} X_{t-1}^2 e^{-c_0 X_{t-1}^2},$$

$$\frac{\partial^2 g(a_0,b_0,c_0)}{\partial a\partial c} = -2b_0\sum_{t=2}^{n} X_{t-1}^4 e^{-c_0 X_{t-1}^2}, \quad \frac{\partial^2 g(a_0,b_0,c_0)}{\partial b^2} = 2\sum_{t=2}^{n} X_{t-1}^2 e^{-2c_0 X_{t-1}^2},$$

$$\frac{\partial^2 g(a_0,b_0,c_0)}{\partial b\partial c} = -2b_0\sum_{t=2}^{n} X_{t-1}^4 e^{-2c_0 X_{t-1}^2} + 2\sum_{t=2}^{n} \varepsilon_t X_{t-1}^3 e^{-c_0 X_{t-1}^2},$$

$$\frac{\partial^2 g(a_0,b_0,c_0)}{\partial c^2} = 2b_0^2\sum_{t=2}^{n} X_{t-1}^6 e^{-2c_0 X_{t-1}^2} - 2\sum_{t=2}^{n} \varepsilon_t X_{t-1}^5 b_0 e^{-c_0 X_{t-1}^2}.$$

Будем предполагать модель (1) стационарной и эргодической. Достаточным условием этого является, например, одновременное выполнение условий $|a|<1,\ c>0$ и $\sigma^2<\infty$ [15]. В этом случае существует $\lim\limits_{n\to\infty} B.$

Найдем, например, $\lim\limits_{n\to\infty}\frac{1}{n}\frac{\partial^2 g(a_0,b_0,c_0)}{\partial b\partial c}$. Из стационарности и эргодичности последовательностей ε_t и X_t вытекает стационарность и эргодичность последовательностей $X_{t-1}^4 b_0 e^{-2c_0 X_{t-1}^2}$ и $\varepsilon_t X_{t-1}^3 e^{-c_0 X_{t-1}^2}$ [16]. Поэтому по закону больших чисел для стационарных и эргодических последовательностей [16] с учетом $\mathrm{E}\varepsilon_t = 0$ и независимости ε_t от X_{t-1} получим при $n\to\infty$ по вероятности

$$\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^4 e^{-2c_0 X_{t-1}^2} \to \mathrm{E}\left(X_0^4 e^{-2c_0 X_0}\right),$$

$$\frac{1}{n}\sum_{t=2}^{n} \varepsilon_t X_{t-1}^3 e^{-c_0 X_{t-1}^2} \to \mathrm{E}\varepsilon_t \mathrm{E}\left(X_{t-1}^3 e^{-c_0 X_{t-1}^2}\right) = 0.$$

Поэтому при $n\to\infty$ по вероятности имеем

$$\frac{1}{n}\frac{\partial^2 g(a_0,b_0,c_0)}{\partial b}\partial c \to -2b_0\mathrm{E}\left(X_0^4 e^{-2c_0 X_0}\right).$$

Аналогично вычисляя остальные элементы матрицы $\lim\limits_{n\to\infty} B,$ получаем, что по вероятности $\lim\limits_{n\to\infty} B = 2K,$ где

$$K = \begin{pmatrix} \mathrm{E}X_0^2 & \mathrm{E}\left(X_0^2 e^{-c_0 X_0^2}\right) & -\mathrm{E}\left(X_0^4 b_0 e^{-c_0 X_0^2}\right) \\ \mathrm{E}\left(X_0^2 e^{-c_0 X_0^2}\right) & \mathrm{E}\left[X_0^2 e^{-2c_0 X_0^2}\right] & -\mathrm{E}\left(X_0^4 b_0 e^{-2c_0 X_0^2}\right) \\ -\mathrm{E}\left(X_0^4 b_0 e^{-c_0 X_0^2}\right) & -\mathrm{E}\left(X_0^4 b_0 e^{-2c_0 X_0^2}\right) & \mathrm{E}\left(X_0^6 b_0^2 e^{-2c_0 X_0^2}\right) \end{pmatrix}.$$

Таким образом, $g(a,b,c) = g(a_0,b_0,c_0) + A^{\mathrm{T}}\theta + \theta^{\mathrm{T}}K\theta + \tilde{\delta}(a,b,c),$ где при $n\to\infty$ $\tilde{\delta}(a,b,c)\to 0$ по вероятности.

Рассуждая так же, как и в работе [17], получаем, что асимптотическое распределение точки $(\tilde{a}, \tilde{b}, \tilde{c})$ минимума функции $g(a, b, c)$ совпадает с асимптотическим распределением точки минимума квадратичной формы $A^{\text{т}}\theta + \theta^{\text{т}}K\theta$, минимум очевидно равен $-2K^{-1}A$.

Найдем асимптотическое распределение случайного вектора $-2K^{-1}A$, для чего докажем, что вектор A является асимптотически нормальным, т. е. слабо (по распределению) сходится при $n \to \infty$ к нормальному случайному вектору. Обозначим через $\mathcal{A}_t$ σ-алгебру событий, порожденную множеством случайных величин $\{X_s,\ s \le t\}$. По определению ε_t не зависит от $\mathcal{A}_t$ и X_{t-1} измерима относительно $\mathcal{A}_{t-1}$. Поэтому из свойств условных математических ожиданий [18] следует, что $\mathrm{E}[\varepsilon_t X_{t-1} \mid \mathcal{A}_{t-1}] = X_{t-1}\mathrm{E}[\varepsilon_t \mid \mathcal{A}_{t-1}] = X_{t-1}\mathrm{E}\varepsilon_t = 0$.

Кроме того, по предположению $0 < \mathrm{E}\varepsilon_t^2 < \infty$, откуда вытекает $0 < \mathrm{E}X_{t-1}^2 < \infty$. Таким образом, по центральной предельной теореме для мартингалов [19] случайная последовательность

$$\frac{1}{\sqrt{n}}\frac{\partial g(a_0, b_0, c_0)}{\partial a} = -\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t X_{t-1}$$

является асимптотически нормальной. Поскольку случайные величины ε_t и X_{t-1} независимы и $\mathrm{E}\varepsilon_t = 0$,

$$\mathrm{E}\left(\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t X_{t-1}\right) = \frac{2}{\sqrt{n}}\sum_{t=2}^{n}\mathrm{E}[\varepsilon_t X_{t-1}] = \frac{2}{\sqrt{n}}\sum_{t=2}^{n}\mathrm{E}\varepsilon_t\mathrm{E}X_{t-1} = 0.$$

Далее, так как при $s < t$ случайные величины ε_t и $\varepsilon_s X_{s-1} X_{t-1}$ независимы и $\mathrm{E}\varepsilon_t = 0$, для всех $s < t$

$$\mathrm{E}\left[\varepsilon_s X_{s-1}\varepsilon_t X_{t-1}\right] = \mathrm{E}\varepsilon_t \mathrm{E}\left[\varepsilon_s X_{s-1} X_{t-1}\right] = 0.$$

Кроме того, из независимости ε_t и X_{t-1} вытекает $\mathrm{E}(\varepsilon_t X_{t-1})^2 = \sigma^2\mathrm{E}X_{t-1}^2 = \sigma^2\mathrm{E}X_0^2$. Следовательно, при $n \to \infty$

$$\mathrm{D}\left(\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\varepsilon_t X_{t-1}\right) = \frac{4}{n}\sum_{t=2}^{n}\mathrm{E}\,(\varepsilon_t X_{t-1})^2 \to 4\sigma^2\mathrm{E}X_0^2.$$

Таким образом, случайная последовательность $\dfrac{1}{\sqrt{n}}\dfrac{\partial g(a_0, b_0, c_0)}{\partial a}$ асимптотически нормальна с нулевым математическим ожиданием и дисперсией $4\sigma^2\mathrm{E}X_0^2$.

Аналогично доказывается, что случайные последовательности $\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial b}$ и $\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial c}$ являются асимптотически нормальными с нулевыми математическими ожиданиями и дисперсиями соответственно $4\sigma^2\mathrm{E}\left(X_0^2 e^{-2c_0X_0^2}\right)$ и $4b_0^2\sigma^2\mathrm{E}\left(X_0^6 e^{-2c_0X_0^2}\right)$.

Далее, рассуждая так же, как при вычислении коэффициентов матрицы K, находим

$$\lim_{n\to\infty}\mathrm{E}\left(\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial a}\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial b}\right)=4\sigma^2\mathrm{E}\left(X_0^2 e^{-c_0X_0^2}\right),$$

$$\lim_{n\to\infty}\mathrm{E}\left(\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial a}\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial c}\right)=-4\sigma^2\mathrm{E}\left(X_0^4 b_0 e^{-c_0X_0^2}\right),$$

$$\lim_{n\to\infty}\mathrm{E}\left(\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial b}\frac{1}{\sqrt{n}}\frac{\partial g(a_0,b_0,c_0)}{\partial c}\right)=-4\sigma^2\mathrm{E}\left(X_0^4 b_0 e^{-2c_0X_0^2}\right).$$

Поэтому случайный вектор A является асимптотически нормальным с нулевым математическим ожиданием и ковариационной матрицей $4\sigma^2 K$. Следовательно [20], асимптотическая ковариационная матрица вектора $-2K^{-1}A$, а значит, и оценки наименьших квадратов $(\tilde{a},\tilde{b},\tilde{c})$ равны σ^2K^{-1}.

Для нахождения асимптотической дисперсии оценки наименьших модулей (a^*,b^*,c^*) аналогичные рассуждения провести нельзя вследствие недифференцируемости функции $g_{\text{н.м}}(a,b,c)$. Для исправления ситуации свернем функцию $|x|$ с последовательностью $f_m(x)$ бесконечно дифференцируемых функций, сходящихся при $m\to\infty$ к δ-функции Дирака $\delta_0(x)$ в нуле. Пусть $f_m(x)=me^{-m^2x^2}\big/\sqrt{\pi}$. Рассмотрим последовательность функций $h_m(x)=\int\limits_{-\infty}^{\infty}|x|\,f_m(x-y)\,dy$, $m=1,2,\ldots$ Несложно доказать, что $h_m(x)\to|x|$ при $m\to\infty$ равномерно по $x\in\mathbb{R}$, поэтому последовательность функций

$$g_m(a,b,c)=\sum_{t=2}^{n}h_m\left(X_t-\left(a+be^{-cX_{t-1}^2}\right)X_{t-1}\right)$$

при $m\to\infty$ равномерно по $(a,b,c)\in\mathbb{R}^3$ сходится к функции $g_{\text{н.м}}(a,b,c)$, а, следовательно, асимптотические дисперсии оценки наименьших модулей (a^*,b^*,c^*) и точки минимума $(\hat{a},\hat{b},\hat{c})$ функции $g_m(a,b,c)$ совпадают.

Предположим, что плотность распределения вероятности $f(x)$ случайных величин ε_t является четной (в частности, все распределения из таблицы таковы). Отсюда и из нечетности производной $h'_m(x)$ функции $h_m(x)$ вытекает, что $\mathrm{E}\, h_{m'}(\varepsilon_t)=0.$ Рассуждая так же, как и при выводе асимптотической дисперсии оценки наименьших квадратов и учитывая $\mathrm{E}h'_m(\varepsilon_t)=0,$ получаем, что асимптотическая дисперсия случайного вектора $(\hat{a},\hat{b},\hat{c})$ равна $\dfrac{\mathrm{E}[h'_m(\varepsilon_1)^2]}{(\mathrm{E}[h''_m(\varepsilon_1)])^2}A^{-1}.$ Поскольку $h_m(x)\to|x|$ при $m\to\infty,$ то $h'_m(x)\to\operatorname{sign}(x)=\dfrac{x}{|x|}$ и $h''_m(x)\to 2\delta_0(x)$ при $m\to\infty.$ Поэтому $\dfrac{\mathrm{E}\,[h'_m(\varepsilon_1)^2]}{(\mathrm{E}\,[h''_m(\varepsilon_1)])^2}\to\dfrac{1}{4f^2(0)}$ при $m\to\infty.$ Следовательно, асимптотическая относительная эффективность оценки наименьших модулей по отношению к оценке наименьших квадратов равна $\sigma^{-2}f^{-2}(0).$

Выводы. Математическое моделирование показало, что оценки наименьших квадратов и наименьших модулей параметров модели экспоненциальной авторегрессии являются состоятельными, т. е. с увеличением числа наблюдений все точнее и точнее оценивают неизвестные параметры. Точность обеих оценок существенно зависит от распределения вероятности обновляющего процесса, входящего в авторегрессионное уравнение. Если распределение вероятности — логистическое, нормальное, загрязненное нормальное с небольшим уровнем загрязнения или распределением Стьюдента с достаточно большим числом степеней свободы, то более эффективна оценка наименьших квадратов и в этом случае ее следует предпочесть оценке наименьших модулей. Однако если вероятностное распределение обновляющего процесса является загрязненным нормальным распределением с достаточно большим уровнем загрязнения, распределением Лапласа или распределением Стьюдента с малым числом степеней свободы, то следует использовать оценку наименьших модулей.

ЛИТЕРАТУРА

[1] Brockwell P.J., Davis R.A. Introduction to time series and forecasting. *Springer Texts in Statistics*. Cham, Springer, 2016. DOI: https://doi.org/10.1007/978-3-319-29854-2

[2] De Gooijer J.G. Elements of nonlinear time series analysis and forecasting. *Springer Texts in Statistics*. Cham, Springer, 2017. DOI: https://doi.org/10.1007/978-3-319-43252-6

[3] Ozaki T. Time series modeling of neuroscience data. CRC Press, 2012.

[4] Merzougui M. Estimation in periodic restricted EXPAR(1) models. *Comm. Statist. Simulation Comput.*, 2018, vol. 47, iss. 10, pp. 2819–2828.
DOI: https://doi.org/10.1080/03610918.2017.1361975

[5] Merzougui M., Dridi H., Chadli A. Test for periodicity in restrictive EXPAR models. *Comm. Statist. Theory Methods*, 2016, vol. 45, iss. 9, pp. 2770–2783.
DOI: https://doi.org/10.1080/03610926.2014.887110

[6] Olugbode M., El-Masry A., Pointon J. Exchange rate and interest rate exposure of UK industries using first-order autoregressive exponential GARCH-in-mean (EGARCH-M) approach. *The Manchester School*, 2014, vol. 82, iss. 4, pp. 409–464.
DOI: https://doi.org/10.1111/manc.12029

[7] Ghosh H., Gurung B., Gupta P. Fitting EXPAR models through the extended Kalman filter. *Sankhya B*, 2015, vol. 77, iss. 1, pp. 27–44.
DOI: https://doi.org/10.1007/s13571-014-0085-8

[8] Gurung B. An exponential autoregressive (EXPAR) model for the forecasting of all India annual rainfall. *Mausam*, 2015, vol. 66, no. 4, pp. 847–849.

[9] Горяинов В.Б., Горяинова Е.Р. Влияние аномальных наблюдений на оценку наименьших квадратов параметра авторегрессионного уравнения со случайным коэффициентом. *Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки*, 2016, № 2, с. 16–24. DOI: 10.18698/1812-3368-2016-2-16-24

[10] Goryainova E.R., Botvinkin E.A. Experimental and analytic comparison of the accuracy of different estimates of parameters in a linear regression model. *Autom. Remote Control*, 2017, vol. 78, iss. 10, pp. 1819–1836.
DOI: https://doi.org/10.1134/S000511791710006X

[11] Lehmann E.L., Casella G. Theory of point estimation. *Springer Texts in Statistics*. New York, NY, Springer, 1998. DOI: https://doi.org/10.1007/b98854

[12] Huber P., Ronchetti E.M. Robust statistics. Wiley, 2009.

[13] Мудров В.И., Кушко В.Л. Метод наименьших модулей. М., Знание, 1971.

[14] Rhinehart R.R. Nonlinear regression modeling for engineering applications: modeling, model validation, and enabling design of experiments. Wiley, 2016.

[15] Chan K.S., Tong H. On the use of deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probab.*, 1985, vol. 17, iss. 3, pp. 666–678. DOI: https://doi.org/10.2307/1427125

[16] White H. Asymptotic theory for econometricians. Academic Press, 2001.

[17] Горяинов В.Б., Горяинова Е.Р. Робастное оценивание в пороговой авторегрессии. *Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки*, 2017, № 6, с. 19–30. DOI: 10.18698/1812-3368-2017-6-19-30

[18] Ширяев А.Н. Вероятность. М., МЦМНО, 2011.

[19] Billingsley P. Convergence of probability measures. Wiley, 1999.

[20] Magnus J.R., Neudecker H. Matrix differential calculus with applications in statistics and econometrics. Wiley, 2007.

Горяинов Владимир Борисович — д-р физ.-мат. наук, профессор кафедры «Математическое моделирование» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Кайнг Вэй Мьо — аспирант кафедры «Математическое моделирование» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Просьба ссылаться на эту статью следующим образом:

Горяинов В.Б., Кайнг В.М. Оценивание параметров экспоненциальной авторегрессионной модели. *Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки*, 2019, № 5, с. 4–18. DOI: 10.18698/1812-3368-2019-5-4-18

EXPONENTIAL AUTOREGRESSIVE MODEL PARAMETERS ESTIMATION

V.B. Goryainov vb-goryainov@bmstu.ru
W.M. Khing waimyokhing@gmail.com

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

The purpose of the research was to compare the least squares estimatate and the least absolute deviation estimate depending on the probability distribution of the renewal process of the autoregressive equation. To achieve this goal, the sequence of observations of the exponential autoregressive process was repeatedly reproduced using computer simulation, and the least squares estimate and the least absolute deviation estimate were calculated for each sequence. The resulting estimation sequences were used to calculate the sample variances of the least squares estimate and the least absolute deviation estimate. The best estimate was the one with the lowest sample variance. The quantitative measure for the estimates comparison was the sample relative efficiency of estimates, defined as the inverse ratio of their sample variances. Normal distribution, contaminated normal distribution, i.e., Tukey distribution, with different values of the proportion and intensity of contamination, logistic distribution, Laplace distribution and Student distribution with different degrees of freedom, in particular, with one degree of freedom, that is, Cau-

Keywords

Exponential autoregression, least squares estimate, least absolute deviation estimate

chy distribution, were used as models of probability distribution of the renewal process. For each probability distribution, asymptotic values of the sample relative efficiency were obtained with an unlimited increase in the sample size of the observations of the autoregressive process. Findings of research show that the least absolute deviation estimate is better than the least squares estimate for Laplace distribution and the contaminated normal distribution with sufficiently large levels of the proportion and intensity of contamination. In other cases, the least squares estimate is preferable

Received 09.04.2019

REFERENCES

[1] Brockwell P.J., Davis R.A. Introduction to time series and forecasting *Springer Texts in Statistics*. Cham, Springer, 2016.
DOI: https://doi.org/10.1007/978-3-319-29854-2

[2] De Gooijer J.G. Elements of nonlinear time series analysis and forecasting. *Springer Texts in Statistics*. Cham, Springer, 2017.
DOI: https://doi.org/10.1007/978-3-319-43252-6

[3] Ozaki T. Time series modeling of neuroscience data. CRC Press, 2012.

[4] Merzougui M. Estimation in periodic restricted EXPAR(1) models. *Comm. Statist. Simulation Comput*., 2018, vol. 47, iss. 10, pp. 2819–2828.
DOI: https://doi.org/10.1080/03610918.2017.1361975

[5] Merzougui M., Dridi H., Chadli A. Test for periodicity in restrictive EXPAR models. *Comm. Statist. Theory Methods,* 2016, vol. 45, iss. 9, pp. 2770–2783.
DOI: https://doi.org/10.1080/03610926.2014.887110

[6] Olugbode M., El-Masry A., Pointon J. Exchange rate and interest rate exposure of UK industries using first-order autoregressive exponential GARCH-in-mean (EGARCH-M) approach. *The Manchester School*, 2014, vol. 82, iss. 4, pp. 409–464.
DOI: https://doi.org/10.1111/manc.12029

[7] Ghosh H., Gurung B., Gupta P. Fitting EXPAR models through the extended Kalman filter. *Sankhya B*, 2015, vol. 77, iss. 1, pp. 27–44.
DOI: https://doi.org/10.1007/s13571-014-0085-8

[8] Gurung B. An exponential autoregressive (EXPAR) model for the forecasting of all India annual rainfall. *Mausam*, 2015, vol. 66, no. 4, pp. 847–849.

[9] Goryainov V.B., Goryainova E.R. The influence of anomalous observations on the least squares estimate of the parameter of the autoregressive equation with random coefficient. *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, 2016, no. 2, pp. 16–24 (in Russ.). DOI: 10.18698/1812-3368-2016-2-16-24

[10] Goryainova E.R., Botvinkin E.A. Experimental and analytic comparison of the accuracy of different estimates of parameters in a linear regression model. *Autom. Remote Control*, 2017, vol. 78, iss. 10, pp. 1819–1836.
DOI: https://doi.org/10.1134/S000511791710006X

[11] Lehmann E.L., Casella G. Theory of point estimation. *Springer Texts in Statistics*. New York, NY, Springer, 1998. DOI: https://doi.org/10.1007/b98854

[12] Huber P., Ronchetti E.M. Robust statistics. Wiley, 2009.

[13] Mudrov V.I., Kushko V.L. Metod naimen'shikh moduley [Least absolute deviations method]. Moscow, Znanie Publ., 1971.

[14] Rhinehart R.R. Nonlinear regression modeling for engineering applications: modeling, model validation, and enabling design of experiments. Wiley, 2016.

[15] Chan K.S., Tong H. On the use of deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probab*., 1985, vol. 17, iss. 3, pp. 666–678. DOI: https://doi.org/10.2307/1427125

[16] White H. Asymptotic theory for econometricians. Academic Press, 2001.

[17] Goryainov V.B., Goryainova E.R. Robust estimation in threshold autoregression. *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, 2017, no. 6, pp. 19–30 (in Russ.). DOI: 10.18698/1812-3368-2017-6-19-30

[18] Shiryaev A.N. Veroyatnost' [Probability]. Moscow, MTsMNO Publ., 2011.

[19] Billingsley P. Convergence of probability measures. Wiley, 1999.

[20] Magnus J.R., Neudecker H. Matrix differential calculus with applications in statistics and econometrics. Wiley, 2007.

Goryainov V.B. — Dr. Sc. (Phys.-Math.), Professor, Department of Mathematical Simulation, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Khing W.M. — Post-Graduate Student, Department of Mathematical Simulation, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Please cite this article in English as:

Goryainov V.B., Khing W.M. Exponential autoregressive model parameters estimation. *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, 2019, no. 5, pp. 4–18 (in Russ.). DOI: 10.18698/1812-3368-2019-5-4-18