УДК 519.234.3 DOI: 10.18698/1812-3368-2017-6-19-30

РОБАСТНОЕ ОЦЕНИВАНИЕ В ПОРОГОВОЙ АВТОРЕГРЕССИИ

В.Б. Горяинов[1] vb-goryainov@bmstu.ru
Е.Р. Горяинова[2] el-goryainova@mail.ru

[1] МГТУ им. Н.Э. Баумана, Москва, Российская Федерация
[2] Национальный исследовательский университет «Высшая школа экономики», Москва, Российская Федерация

Аннотация

Изучены робастные свойства М-оценок параметров самовозбуждающейся пороговой авторегрессионной модели. Функция потерь, определяющая М-оценки, предполагалась выпуклой и дважды дифференцируемой. Порог авторегрессионной модели считался известным и единственным. Доказана асимптотическая нормальность М-оценок параметров авторегрессионного уравнения. Найдена асимптотическая относительная эффективность М-оценок, оценки наименьших квадратов и оценки наименьших модулей по отношению друг к другу. Значения асимптотической относительной эффективности вычислены для нормального распределения, двойного экспоненциального распределения (распределения Лапласа) и загрязненного нормального распределения (распределения Тьюки). Исследована зависимость асимптотической относительной эффективности указанных оценок от параметров (доли загрязнения и величины загрязнения) распределения Тьюки. Для всех трех оценок в пространстве параметров распределения Тьюки построены линии равной эффективности, которые позволили выделить области предпочтения для каждой пары рассмотренных оценок. Показано, что при небольшом отклонении распределения обновляющего процесса от нормального М-оценки являются эффективнее оценок наименьших квадратов и наименьших модулей. Приведены рекомендации по использованию указанных оценок в практических исследованиях

Ключевые слова

Пороговая модель авторегрессии, М-оценки, асимптотическая нормальность, асимптотическая относительная эффективность, распределение Тьюки

Поступила в редакцию 20.02.2017

Введение. В последнее время в теории временных рядов большое внимание уделяется нелинейным моделям, широко используемым в различных областях науки и техники [1].

Одним из наиболее распространенных классов нелинейных моделей является класс пороговых моделей, описываемых кусочно-линейными уравнениями в зависимости от величины пороговой переменной [2]. Возникновение таких моделей, в частности, было обусловлено желанием описать «эффект асимметрии» в поведении финансовых индексов. Этот эффект возрастания волатильности после падения цен не удавалось описать ни в рамках линейных, ни в рамках условно гетерокседастических моделей.

Когда каждый линейный режим описывается авторегрессионным уравнением, модель называется пороговой авторегрессионной моделью (*threshold autoregression*, или *TAR*) [3]. Если пороговой переменной является одно из прошлых наблюдений этого же временного ряда, пороговая модель называется самовозбуждающейся (*self excited threshold autoregression*, или *SETAR*). Примеры, приведенные в работе [2], показывают, что модели *TAR* обеспечивают лучшее описание реальных временных рядов, чем линейные модели. В частности, *SETAR*-модели способны описывать предельные циклы, резонансные скачки и некоторые другие асимметричные особенности, наблюдаемые в самых разнообразных, в том числе экономических и финансовых, временных рядах.

Одна из наиболее важных задач в теории нелинейных временных рядов — оценивание параметров описывающей этот ряд модели. Настоящая работа посвящена М-оценкам авторегрессионных коэффициентов *SETAR*-модели с одним порогом, доказана асимптотическая нормальность этих оценок и исследована их асимптотическая относительная эффективность по отношению к оценкам наименьших квадратов и наименьших модулей.

Постановка задачи. Рассмотрим простейший вариант самовозбуждающейся пороговой модели, а именно случайный процесс X_t, описываемый уравнением

$$X_t = a_1 X_{t-1}^{+} + a_2 X_{t-1}^{-} + \varepsilon_t, \quad t = 1,2,\ldots, \tag{1}$$

где $x^{+} = \max(x,0)$, а $x^{-} = \min(x,0)$. В (1) авторегрессионные коэффициенты представляют собой неслучайные действительные числа, а ε_t, $t \geq 1$ — последовательность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием $\mathrm{E}\varepsilon_t = 0$. Эквивалентным образом модель (1) может быть представлена в виде $X_t = h(X_{t-1}) + \varepsilon_t$, $t = 1,2,\ldots$, где $h(X_{t-1}) = \left(a_1 I(X_{t-1} > 0) + a_2 I(X_{t-1} \leq 0)\right) X_{t-1}$; $I(A)$ — индикаторная функция множества A.

Известно, что если плотность распределения вероятности $f(x)$ случайных величин ε_t четная, то распределение вероятности X_t симметрично относительно нуля тогда и только тогда, когда $a_1 = a_2$. Поэтому модель (1) иногда называют асимметричной авторегрессионной моделью.

Далее всюду будет предполагаться, что $f(x) > 0$ для любых $x \in \mathbb{R}$ и существует $\mathrm{E}\,|\,\varepsilon_t\,|^{2+\delta} < \infty$ для некоторого $\delta > 0$. В этом случае процесс X_t является стационарным и эргодическим тогда и только тогда, когда одновременно $a_1 < 1$, $a_2 < 1$ и $a_1 a_2 < 1$ [4].

Цель настоящей работы — исследование свойств М-оценок коэффициентов $a = (a_1, a_2)^{\mathrm{т}}$ по наблюдениям $X_1, \ldots, X_n$, а также сравнение эффективности

М-оценок по отношению к оценкам наименьших квадратов и наименьших модулей.

Оценка наименьших квадратов. Оценку наименьших квадратов $a^* = (a_1^*, a_2^*)^{\text{т}}$ параметра $a = (a_1, a_2)^{\text{т}}$ по наблюдениям $X_1, X_2, \ldots, X_n$ определяют как точку минимума функции

$$L_{LSQ}(a) = \sum_{t=1}^{n} (X_t - a_1 X_{t-1}^{+} - a_2 X_{t-1}^{-})^2.$$

С учетом того, что $X_t^+ X_t^- = 0,$ для любого t, имеем

$$a_1^* = \frac{\sum_{t=1}^{n} X_t X_{t-1}^{+}}{\sum_{t=1}^{n} (X_{t-1}^{+})^2}; \qquad a_2^* = \frac{\sum_{t=1}^{n} X_t X_{t-1}^{-}}{\sum_{t=1}^{n} (X_{t-1}^{-})^2}.$$

Эти оценки состоятельны и асимптотически нормальны с ковариационной матрицей $\sigma^2 K^{-1},$ где

$$K = \begin{pmatrix} \mathrm{E}(X_1^+)^2 & 0 \\ 0 & \mathrm{E}(X_1^-)^2 \end{pmatrix}.$$

Состоятельность означает, что при $n \to \infty$ последовательность a^* сходится по вероятности к $a,$ асимптотическая нормальность — что для любых действительных x и y

$$\lim_{n\to\infty} \left\{ \sqrt{n\mathrm{E}X_1^{+2}}\,(a_{1n}^* - a_1) < x\sigma, \sqrt{n\mathrm{E}X_1^{-2}}\,(a_{2n}^* - a_2) < y\sigma \right\} = \Phi(x)\Phi(y),$$

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ — функция распределения вероятности стандартной нормальной случайной величины; $\sigma^2 = \mathrm{E}\varepsilon_t^2$ — общая дисперсия всех $\varepsilon_t,$ $t = 1,\ 2, \ldots$

Оценка наименьших модулей. Оценку наименьших модулей $\tilde{a} = (\tilde{a}_1,\ \tilde{a}_2)^{\text{т}}$ параметра a по наблюдениям $X_1, X_2, \ldots, X_n$ определяют как точку минимума функции

$$L_{LAD}(a) = \sum_{t=1}^{n} | X_t - a_1 X_{t-1}^{+} - a_2 X_{t-1}^{-} |.$$

Оценки наименьших модулей состоятельны и асимптотически нормальны с ковариационной матрицей $\dfrac{1}{4 f^2(0)} K^{-1}$ [5].

Функция $L_{LAD}(a)$ кусочно-линейна и выпукла. Ее минимум можно найти стандартными методами минимизации кусочно-линейных выпуклых функций, например методами линейного программирования [6] или методом Нелдера — Мида [7], который так же известен как метод деформируемого многогранника, или симплекс-метод. Однако наиболее распространенным способом минимизации функции $L_{LAD}(a)$ является итерационный взвешенный метод наименьших квадратов [8], при котором эта функция представляется в виде

$$L_{LAD}(a)=\sum_{t=1}^{n} w_t(a_1, a_2)(X_t-a_1X_{t-1}^{+}-a_2X_{t-1}^{-})^2 \tag{2}$$

с весами $w_t(a_1,a_2)=1/\,|\,X_t-a_1X_{t-1}^{+}-a_2X_{t-1}^{-}\,|$. Точка минимума $\tilde{a}$ функции $L_{LAD}(a)$ является пределом последовательности $a^{(k)}=(a_1^{(k)},a_2^{(k)})^{\text{т}}$, k-й член которой находится с помощью минимизации функции

$$\tilde{L}_{LAD}(a)=\sum_{t=1}^{n} w_t(a_1^{(k-1)},a_2^{(k-1)})(X_t-a_1X_{t-1}^{+}-a_2X_{t-1}^{-})^2.$$

В качестве начального приближения $a^{(0)}=(a_1^{(0)},\, a_2^{(0)})^{\text{т}}$ можно использовать, например, оценку наименьших квадратов.

Сравнение оценок наименьших квадратов и наименьших модулей. Точность состоятельной и асимптотически нормальной оценки определяется ее асимптотической дисперсией — чем меньше асимптотическая дисперсия, тем точнее оценка. Из двух оценок лучше та, которая имеет меньшую дисперсию. Сравнительной характеристикой точности двух асимптотически нормальных оценок является асимптотическая относительная эффективность, определяемая как обратное отношение асимптотических дисперсий оценок. Асимптотическая относительная эффективность одной оценки по отношению к другой показывает, во сколько раз меньше наблюдений необходимо первой оценке для достижения точности второй оценки. Таким образом, асимптотическая относительная эффективность оценки наименьших квадратов (ОНК) по отношению к оценке наименьших модулей (ОНМ) равна

$$e(\text{ОНК, ОНМ})=\frac{1}{4\sigma^2 f^2(0)}. \tag{3}$$

Обычно полагают, что флуктуационные возмущения ε_t модели (1) являются суммой большого числа примерно одинаковых малых случайных величин, имеющих самое разнообразное происхождение и, следовательно, на основании центральной предельной теоремы подчинены нормальному закону. Если ε_t нормальные случайные величины с нулевым математическим ожиданием, то $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\dfrac{x^2}{2\sigma^2}\right)$, поэтому $e(\text{ОНК, ОНМ})=\pi/2$, т. е. оценка наименьших квадратов лучше оценки наименьших модулей в $\pi/2$ раз — примерно в 1,5 раза. Другими словами, для достижения одинаковой с оценкой наименьших квадратов точности оценке наименьших модулей необходимо в $\pi/2$ раз больше наблюдений.

Однако нормальное распределение является не единственно возможным естественным распределением возмущений ε_t. Во многих случаях разумно полагать, что распределение случайных величин ε_t является нормальным, но со случайной дисперсией η (точнее нормальным является условное распределение ε_t при условии, что случайная дисперсия η приняла какое-то конкретное значение y). Если о дисперсии η дополнительно не делается никаких предположений, кроме существования средней дисперсии $\mathrm{E}\eta$, то, полагая положение дисперсии η на числовой оси максимально неопределенным, методами теории информации можно получить [9], что безусловное распределение ε_t будет двусторонним экспоненциальным распределением, или распределением Лапласа с плотностью

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

где $\sigma^2=\mathrm{D}\varepsilon_t=(\mathrm{E}\eta)^2/2.$ В этом случае $e(\text{ОНК, ОНМ})=1/2,$ т. е. оценка наименьших модулей в 2 раза эффективнее оценки наименьших квадратов.

Согласно центральной предельной теореме, предположение о нормальности выполняется лишь приближенно. Общепринятая модель приближенного нормального распределения — распределение Тьюки, являющееся сочетанием нормальных распределений, при котором с небольшой вероятностью дисперсия нормального распределения значительно увеличивается, иногда в несколько раз. Плотность распределения вероятности распределения Тьюки имеет вид

$$f(x)=(1-\gamma)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}+\gamma\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{x^2}{2\tau^2}},\qquad \tau>1,\quad 0<\gamma<1. \tag{4}$$

Другими словами, если последовательность случайных величин ε_t имеет распределение Тьюки, то среди ε_t с вероятностью $1-\gamma$ встречаются стандартные (с нулевым математическим ожиданием и единичной дисперсией) нормальные величины, а с вероятностью γ — нормальные величины с дисперсией τ^2. Чем больше γ и τ, тем сильнее распределение Тьюки отклоняется от нормального.

В этом случае $\sigma^2=1+\tau^2\gamma-\gamma,$ $f(0)=(1-\gamma+\gamma/\tau)/\sqrt{2\pi},$ а величина $e(\text{ОНК, ОНМ}),$ согласно (3), будет равна

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

С ростом γ и τ величина $e(\gamma,\tau)$ становится меньше единицы, т. е. оценка наименьших модулей становится лучше оценки наименьших квадратов. Кроме того, так как $e(\gamma,\tau)\to 0$ при $\tau\to\infty$ и любом фиксированном γ, то оценки наименьших модулей могут быть эффективнее оценки наименьших квадратов в любое число раз. При небольших τ оценка наименьших квадратов всегда (при любых γ) эффективнее оценки наименьших модулей. Действительно, минимум $e(\gamma,\tau)$ как функция γ, достигаемый в точке $\gamma=\dfrac{\tau+2}{3(\tau+1)}$, равен $\min\limits_{0\le\gamma\le1} e(\gamma,\tau)=\dfrac{27\pi\tau^2(\tau+1)^2}{8(\tau^2+\tau+1)^3}$ и становится меньше единицы только при $\tau>\tau_0$, $\tau_0\approx 2,3$.

Следовательно, в наиболее распространенном на практике случае небольшого отклонения распределения ε_t от нормального эффективность обеих оценок невелика — оценка наименьших квадратов уже неэффективна, а оценка наименьших модулей еще неэффективна.

М-оценки. Получить оценки, достаточно эффективные во всем диапазоне значений τ и γ можно, рассмотрев оценки, которые являются точкой минимума функции

$$\mathcal{L}(a_1,a_2)=\sum_{t=1}^{n}\rho(X_t-a_1X_{t-1}^{+}-a_2X_{t-1}^{-}), \tag{5}$$

где функция $\rho(x)$ ведет себя в окрестности начала координат как x^2, а вне этой окрестности как $|x|$. Наиболее распространенным является семейство ρ-функций Хьюбера [10]:

$$\rho_k(x)=\begin{cases} x^2, & \text{если } |x|\le k;\\ 2k|x|-k^2, & \text{если } |x|>k,\end{cases} \tag{6}$$

где k — параметр, изменяющийся от нуля (метод наименьших модулей) до бесконечности (метод наименьших квадратов). Оценки $\hat a=(\hat a_1,\hat a_2)^{\mathrm{т}}$, полученные минимизацией (5) с произвольной ρ-функцией, называют М-оценками (получившими такое название от оценок максимального правдоподобия, которые являются частным случаем М-оценок с ρ-функцией, $\rho(x)=-\ln f(x)$).

Имеет место следующая теорема.

Теорема. *Пусть процесс* X_t, *описываемый уравнением (1), является стационарным,* ρ *— выпуклая функция,* ρ'' *непрерывна почти всюду и ограничена, а плотность* $f(x)$ *независимых одинаково распределенных случайных величин* ε_t *в (1) удовлетворяет условиям* $f(x)>0$, $x\in\mathbb{R}$, $\mathrm{E}\varepsilon_t=0$, $\mathrm{E}(\rho'(\varepsilon_t)^2)>0$, $\mathrm{E}\rho'(\varepsilon_t)=0$, $0<\mathrm{E}\rho''(\varepsilon_t)<\infty$ *и* $\mathrm{E}\varepsilon_t^{2+\delta}<\infty$ *для некоторого* $\delta>0$. *Тогда при* $n\to\infty$ *случайный вектор* $\sqrt{n}(\hat a-a)$ *является асимптотически нормальным с нулевым математическим ожиданием и ковариационной матрицей*

$$K^{-1}\frac{\mathrm{E}[\rho'(\varepsilon_1)^2]}{(\mathrm{E}[\rho''(\varepsilon_1)])^2}.$$

◄ Отметим, что вектор $\hat a$ минимизирует (5) тогда и только тогда, когда вектор $\hat\alpha=\sqrt{n}(\hat a-a)$ является точкой минимума функции

$$L(\alpha)=\sum_{t=1}^{n}\Big(\rho\Big(\varepsilon_t-\alpha^{\mathrm{т}}\tilde X_t n^{-1/2}\Big)-\rho(\varepsilon_t)\Big),$$

где $\tilde{X}_t = (X_{t-1}^{+}, X_{t-1}^{-})^{\mathrm{т}}$. Раскладывая $L(\alpha)$ в окрестности нуля по формуле Тейлора, получаем

$$L(\alpha) = -L_1^{\mathrm{т}}\alpha + 2^{-1}\alpha^{\mathrm{т}} L_2 \alpha + \beta(\alpha),$$

где

$$L_1 = n^{-1/2}\sum_{t=1}^{n}\rho'(\varepsilon_t)\tilde{X}_t; \quad L_2 = n^{-1}\sum_{t=1}^{n}\rho''(\varepsilon_t)\tilde{X}_t\tilde{X}_t^{\mathrm{т}},$$

$$\beta(\alpha) = n^{-1}\sum_{t=1}^{n}\left(\rho''\left(\varepsilon_t - \theta\alpha^{\mathrm{т}}\tilde{X}_t n^{-1/2}\right) - \rho''(\varepsilon_t)\right)\alpha^{\mathrm{т}}\tilde{X}_t\tilde{X}_t^{\mathrm{т}}\alpha, \ \ \theta \in (0,1).$$

Рассуждая так же, как и в работе [10], получаем, что из стационарности X_t и ε_t, $\mathrm{E}\rho''(\varepsilon_t) < \infty$, непрерывности и ограниченности $\rho''(x)$ по теореме о мажорируемой сходимости следует, что $\mathrm{E}\,|\,\beta(\alpha)\,| \to 0$ при $n \to \infty$, отсюда $\beta(\alpha) \to 0$ по вероятности при $n \to \infty$.

Обозначим через $\tilde{\alpha}$ минимум функции $\tilde{L}(\alpha) = -L_1^{\mathrm{т}}\alpha + 2^{-1}\alpha^{\mathrm{т}} W\alpha$, где $W = \mathrm{E}(\rho''(\varepsilon_1))K$. Очевидно, что $\tilde{\alpha} = W^{-1}L_1$. Последовательность $\rho'(\varepsilon_t)\tilde{X}_t$, $t = 1, 2, \ldots$, образует мартингал-разность относительно последовательности σ-алгебр $\mathfrak{A}_t$, порожденной множеством $\{\varepsilon_s,\ s \le t\}$, поскольку $\mathrm{E}[\rho'(\varepsilon_t)\tilde{X}_t\,|\,\mathfrak{A}_{t-1}] = 0$ в силу $\mathrm{E}\rho'(\varepsilon_t) = 0$ и независимости $\mathrm{E}\rho'(\varepsilon_t)$ от $\tilde{X}_t$. Согласно центральной предельной теореме, для мартингалов [12] случайный вектор L_1 является асимптотически нормальным с нулевым математическим ожиданием и ковариационной матрицей $R\mathrm{E}[\rho'(\varepsilon_{11})^2]$. Следовательно, случайная величина $\tilde{\alpha} = W^{-1}L_1$ асимптотически нормальна с нулевым математическим ожиданием и ковариационной матрицей $R^{-1}\mathrm{E}[\rho'(\varepsilon_{11})^2]$. Поэтому для доказательства теоремы достаточно показать, что $\hat{\alpha} - \tilde{\alpha} \to 0$ по вероятности, т. е. $\mathrm{P}\{|\,\hat{\alpha} - \tilde{\alpha}\,| \le \delta\} \to 1$ для любого $\delta > 0$.

Доказательство проведем по схеме, приведенной в работах [13, 14]. Зафиксируем $\delta > 0$. Из асимптотической нормальности последовательности $\tilde{\alpha}$ следует ее ограниченность по распределению. Таким образом, существует компакт B из $\mathbb{R}^2$ с вероятностью, сколь угодно близкой к единице, содержащий сразу для всех n шары с центром $\tilde{\alpha}$ и радиусом δ.

Согласно закону больших чисел, для стационарных последовательностей $L_2 \to W$ при $n \to \infty$. В связи с этим $L(\alpha) - \tilde{L}(\alpha) \to 0$ по вероятности при $n \to \infty$. Следовательно, выпуклая функция $L(\alpha) + L_1^{\mathrm{т}}$ сходится к выпуклой функции $2^{-1}\alpha^{\mathrm{т}} W\alpha$ при $n \to \infty$. Таким образом, $\tilde{\beta}(\alpha) \to 0$ равномерно на любом компакте, в частности, $\Delta = \sup\limits_{\alpha \in B}|\,\tilde{\beta}(\alpha)\,| \to 0$ по вероятности при $n \to \infty$ [15].

Обозначим через e произвольный вектор единичной длины, $\alpha^* = \tilde{\alpha} + \delta e$, $\alpha = \tilde{\alpha} + te$, $t > \delta$. Из выпуклости $L(\alpha)$ следует, что

$$L(\alpha^*) \le \left(1 - \frac{\delta}{t}\right)L(\tilde{\alpha}) + \frac{\delta}{t}L(\alpha),$$

поэтому $L(\alpha)\ge L(\tilde\alpha)+\dfrac{t}{\delta}\Big(L(\alpha^*)-(L\tilde\alpha)\Big)\ge L(\tilde\alpha)+\dfrac{t}{\delta}\Big(\tilde L(\alpha^*)-(\tilde L\tilde\alpha)\Big)+\beta(\alpha^*)-\beta(\tilde\alpha)$. Поскольку матрица W положительно определена, $\tilde L(\alpha^*)-(\tilde L\tilde\alpha)=2^{-1}\delta^2e^{\mathrm{т}}We>0$. Следовательно,

$$\inf_{\alpha\notin B}L(\alpha)=L(\tilde\alpha)+\inf_{t>\delta}\frac{t}{\delta}\Big(L(\alpha^*)-L(\tilde\alpha)\Big)\ge L(\tilde\alpha)+2^{-1}\delta^2e^{\mathrm{т}}We-2\Delta.$$

Поэтому с вероятностью, стремящейся к единице, минимум $L(\alpha)$ лежит внутри компакта B, т. е. $\mathrm{P}\{|\hat\alpha-\tilde\alpha|\le\delta\}\to1$ для любого $\delta>0$ при $n\to\infty$. Теорема доказана. ►

Эффективность М-оценки. Для семейства ρ-функций Хьюбера (6)

$$\mathrm{E}[\rho'(\varepsilon_1)^2]=4\left(k^2+2(1-\gamma)(1-k^2)\Phi_0(k)+2\gamma(\tau^2-k^2)\Phi_0\left(\frac{k}{\tau}\right)+\right.$$
$$\left.+\frac{k\sqrt2(\gamma-1)}{\sqrt\pi}e^{-\frac{k^2}{2}}-\frac{k\tau\gamma\sqrt2}{\sqrt\pi}e^{-\frac{k^2}{2\tau^2}}\right);$$

$$\mathrm{E}[\rho''(\varepsilon_1)]=4\left((1-\gamma)\Phi_0(k)+\gamma\Phi_0\left(\frac{k}{\tau}\right)\right),$$

где $\Phi_0(x)=\int\limits_0^x\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$ — функция Лапласа.

Это позволяет получить аналитическую зависимость асимптотической относительной эффективности М-оценки относительно оценки наименьших квадратов и оценки наименьших модулей, которые имеют вид

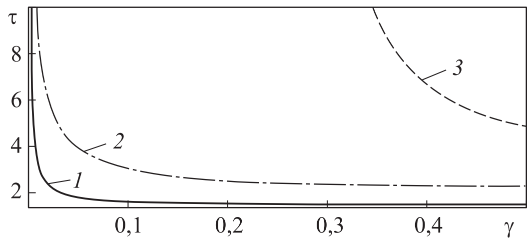
$$e(\text{М, ОНК})=\frac{\sigma^2(\mathrm{E}[\rho''(\varepsilon_1)])^2}{\mathrm{E}[\rho'(\varepsilon_1)^2]}\ \text{ и }\ e(\text{М, ОНМ})=\frac{(\mathrm{E}[\rho''(\varepsilon_1)])^2}{4f^2(0)\mathrm{E}[\rho'(\varepsilon_1)^2]}$$

соответственно.

Выясним, когда М-оценки эффективнее оценок наименьших квадратов и наименьших модулей. На практике, как правило, $\varepsilon\in(0{,}0{,}3)$, $\tau\in(1{,}10)$. Анализ показывает, что для таких значений ε и τ величины $e(\text{М, ОНК})$ и $e(\text{М, ОНМ})$ принимают наибольшие значения при $k\in(1{,}3)$.

Линии уровня $e(\text{М, ОНК})=1$, $e(\text{ОНМ, ОНК})=1$, $e(\text{М, ОНМ})=1$ асимптотических относительных эффективностей $e(\text{М, ОНК})$, $e(\text{ОНМ, ОНК})$ и $e(\text{М, ОНМ})$ при $k=1{,}5$ как функции $\tau(\gamma)$ приведены на рисунке. При этом множество $\{(\gamma,\tau):e(\text{М, ОНК})>1\}$ находится над кривой *1*, множество $\{(\gamma,\tau):e(\text{ОНМ, ОНК})>1\}$ — над кривой *2*, множество $\{(\gamma,\tau):e(\text{М, ОНМ})>1\}$ — под кривой *3*.

Для типичных на практике значений $\gamma\in(0{,}1{,}0{,}3)$, $\tau>2$, М-оценки эффективнее и оценки наименьших квадратов, и оценки наименьших модулей. В частности, если (γ,τ) лежит между кривыми *1* и *2*, то оценка наименьших квадратов все еще лучше оценки наименьших модулей, но уже хуже М-оценки.



Линии уровня $e(\text{М}, \text{ОНК}) = 1$ (*1*), $e(\text{ОНМ}, \text{ОНК}) = 1$ (*2*) и $e(\text{М}, \text{ОНМ}) = 1$ (*3*) асимптотических относительных эффективностей

Следовательно, при оценивании параметров модели пороговой авторегрессии М-оценку следует предпочесть как оценке наименьших квадратов, так и оценке наименьших модулей.

Заключение. Моделирование приближенного нормального распределения распределением Тьюки называется инновационной моделью погрешности наблюдений. Более серьезной причиной снижения точности оценки наименьших квадратов является проведение наблюдений авторегрессионного процесса с погрешностями. Наиболее распространенными моделями погрешностей наблюдений являются аддитивная и замещающая модели. В этом случае оценки, как правило, становятся смещенными, и количественной мерой этого смещения является функционал влияния оценки. Далее планируется для приведенных здесь оценок вычислить функционалы влияния, как это было выполнено в работе [16] для оценки наименьших квадратов параметра авторегрессионной модели со случайным коэффициентом.

ЛИТЕРАТУРА

1. *Franses P.H., Dijk D.V., Opschoor A.* Time series models for business and economic forecasting. Cambridge: Cambridge University Press, 2014. 300 p.

2. *Tong H.* Nonlinear time series: A dynamical approach. New York: Oxford University Press, 1990. 564 p.

3. *Tong H.* Threshold models in time series analysis — 30 years on // Statistics and its Interface. 2011. Vol. 4. No. 2. P. 107–118. DOI: 10.4310/SII.2011.v4.n2.a1

4. *Petruccelli J.D., Woolford S.W.* A threshold AR(1) model // J. Appl. Probab. 1984. Vol. 21. Iss. 2. P. 270–286. DOI: 10.1017/S0021900200024670

5. *Wang L., Wang J.* The limiting behavior of least absolute deviation estimators for threshold autoregressive models // J. Multivariate Anal. 2004. Vol. 89. Iss. 2. P. 243–260.
DOI: 10.1016/j.jmva.2004.02.006

6. *Pan P.-Q.* Linear programming computation. Heidelberg: Springer, 2014. 747 p.

7. *Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.* Numerical recipes: The art of scientific computing. New York: CAP, 2007. 1235 p.

8. *Bissantz N., Dümbgen L., Munk A., Stratmann B.* Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces // SIAM J. Optim. 2009. Vol. 19. Iss. 4. P. 1828–1845. DOI: 10.1137/050639132
URL: http://epubs.siam.org/doi/abs/10.1137/050639132

9. *Мудров В.И., Кушко В.Л.* Метод наименьших модулей. М.: Знание, 1971. 64 с.

10. *Huber P., Ronchetti E.M.* Robust statistics. Hoboken: Wiley, 2009. 360 p.

11. *Горяинов В.Б.* М-оценки пространственной авторегрессии // Автоматика и телемеханика. 2012. № 8. С. 119–129.

12. *Häusler E., Luschgy H.* Stable convergence and stable limit theorems. Heidelberg: Springer, 2015. 228 p.

13. *Горяинов В.Б.* Оценки наименьших модулей коэффициентов пространственной авторегрессии // Известия РАН. Теория и системы управления. 2011. № 4. С. 58–65.

14. *Горяинов А.В., Горяинова Е.Р.* Сравнение эффективности оценок методов наименьших модулей и наименьших квадратов в авторегрессионной модели со случайным коэффициентом // Автоматика и телемеханика. 2016. № 9. С. 84–95.

15. *Andersen P.K., Gill R.D.* Cox's regression model for counting processes: A large sample study // Ann. Statist. 1982. Vol. 10. No. 4. P. 1100–1120. DOI: 10.1214/aos/1176345976
URL: https://projecteuclid.org/euclid.aos/1176345976

16. *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

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).

Горяинова Елена Рудольфовна — канд. физ.-мат. наук, доцент департамента математики на факультете экономических наук Национального исследовательского университета «Высшая школа экономики» (Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20).

Просьба ссылаться на эту статью следующим образом:

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

ROBUST ESTIMATION IN THRESHOLD AUTOREGRESSION

V.B. Goryainov[1] vb-goryainov@bmstu.ru
E.R. Goryainova[2] el-goryainova@mail.ru

[1] **Bauman Moscow State Technical University, Moscow, Russian Federation**
[2] **National Research University Higher School of Economics, Moscow, Russian Federation**

Abstract

In this paper we study robust properties of M-estimates of the parameters of self-excited threshold autoregression model. The loss function that determines M-estimates was supposed to be convex and twice differentiable. The threshold of the autoregressive model was considered to be known and unique.

Keywords

Threshold autoregression model, M-estimates, asymptotic normality, asymptotic relative efficiency, Tukey distribution

We proved the asymptotic normality of M-estimates of autoregressive equation parameters. Moreover, we found the asymptotic relative efficiency of M-estimates, the least square and least absolute deviation estimates with respect to each other. First, we calculated the asymptotic relative efficiency values for the normal distribution, as well as the values of the double exponential distribution (Laplace distribution) and contaminated normal distribution (Tukey distribution). Then, we described the dependence of the asymptotic relative efficiency of these estimates on Tukey distribution parameters (the proportion and level of contamination). Next, for all three estimates in the space of Tukey distribution parameters, we built lines of equal efficiency, which made it possible to single out the preference areas for each pair of estimates considered. Findings of the research show that M-estimates are more efficient than the least squares and least absolute deviation estimates if the distribution of the innovation process slightly deviates from the normal distribution. Finally, we give recommendations on the use of these estimates in practical applications

Received 20.02.2017

REFERENCES

[1] Franses P.H., Dijk D.V., Opschoor A. Time series models for business and economic forecasting. Cambridge, Cambridge University Press, 2014. 300 p.

[2] Tong H. Nonlinear time series: A dynamical approach. New York, Oxford University Press, 1990. 564 p.

[3] Tong H. Threshold models in time series analysis — 30 years on. *Statistics and its Interface*, 2011, vol. 4, no. 2, pp. 107–118. DOI: 10.4310/SII.2011.v4.n2.a1

[4] Petruccelli J.D., Woolford S.W. A threshold AR(1) model. *J. Appl. Probab.*, 1984, vol. 21, iss. 2, pp. 270–286. DOI: 10.1017/S0021900200024670

[5] Wang L., Wang J. The limiting behavior of least absolute deviation estimators for threshold autoregressive models. *J. Multivariate Anal.*, 2004, vol. 89, iss. 2, pp. 243–260.
DOI: 10.1016/j.jmva.2004.02.006

[6] Pan P.-Q. Linear programming computation. Heidelberg, Springer, 2014. 747 p.

[7] Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. Numerical recipes: The art of scientific computing. New York, CAP, 2007. 1235 p.

[8] Bissantz N., Dümbgen L., Munk A., Stratmann B. Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces. *SIAM J. Optim.*, 2009, vol. 19, iss. 4, pp. 1828–1845. DOI: 10.1137/050639132
Available at: http://epubs.siam.org/doi/abs/10.1137/050639132

[9] Mudrov V.I., Kushko V.L. Metod naimen'shikh moduley [Least modules method]. Moscow, Znanie Publ., 1971. 64 p.

[10] Huber P., Ronchetti E.M. Robust statistics. Hoboken, Wiley, 2009. 360 p.

[11] Goryainov V.B. M-estimates of the spatial autoregression coefficients. *Automation and Remote Control*, 2012, vol. 73, iss. 8, pp. 1371–1379. DOI: 10.1134/S0005117912080103

[12] Häusler E., Luschgy H. Stable convergence and stable limit theorems. Heidelberg, Springer, 2015. 228 p.

[13] Goryainov V.B. Least-modules estimates for spatial autoregression coefficients. *Journal of Computer and Systems Sciences International*, 2011, vol. 50, no. 4, pp. 565–572.
DOI: 10.1134/S1064230711040101

[14] 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. *Automation and Remote Control*, 2016, vol. 77, iss. 9, pp. 1579–1588.
DOI: 10.1134/S000511791609006X

[15] Andersen P.K., Gill R.D. Cox's regression model for counting processes: A large sample study. *Ann. Statist.*, 1982, vol. 10, no. 4, pp. 1100–1120. DOI: 10.1214/aos/1176345976
Available at: https://projecteuclid.org/euclid.aos/1176345976

[16] 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. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Estestv. Nauki* [Herald of the Bauman Moscow State Tech. Univ., Nat. Sci.], 2016, no. 2, pp. 16–24 (in Russ.). DOI: 10.18698/1812-3368-2016-2-16-24

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

Goryainova E.R. — Cand. Sc. (Phys.-Math.), Assoc. Professor of Faculty of Economic Sciences, Department of Mathematics, National Research University Higher School of Economics (Myasnitskaya ul. 20, Moscow, 101000 Russian Federation).

Please cite this article in English as:

Goryainov V.B., Goryainova E.R. Robust Estimation in Threshold Autoregression. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Estestv. Nauki* [Herald of the Bauman Moscow State Tech. Univ., Nat. Sci.], 2017, no. 6, pp. 19–30 (in Russ.). DOI: 10.18698/1812-3368-2017-6-19-30