Наука и Образование
МГТУ им. Н.Э. Баумана
Сетевое научное издание
ISSN 1994-0408

Ссылка на статью:
// Наука и Образование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2016. № 9. С. 111–122.

DOI: 10.7463/0916.0844976

Представлена в редакцию: 25.08.2016

УДК 519.234.3

Робастное оценивание в авторегрессионной модели со случайным коэффициентом

Горяинов В. Б.[1,*], Ермаков С. Ю.[1]

[*]vb-goryainov@bmstu.ru

[1]МГТУ им. Н.Э. Баумана, Москва, Россия

В работе исследованы робастные свойства М-оценки и оценки наименьших квадратов параметра авторегрессионного уравнения первого порядка со случайным коэффициентом. М-оценка вычислялась на основе ро-функций Хьюбера и Хампеля. Рассматривались аддитивная, замещающая и инновационная модели ошибок наблюдений. В инновационной модели предполагалось, что обновляющий процесс авторегрессионного уравнения имеет загрязненное нормальное распределение (распределение Тьюки). При помощи компьютерного моделирования изучена относительная эффективность М-оценки по отношению к оценке наименьших квадратов. В частности получена зависимость относительной эффективности от параметров распределения Тьюки.

Ключевые слова: оценка наименьших квадратов; М-оценка; авторегрессия со случайным коэффициентом; распределение Тьюки

Введение

Одним из основных инструментов описания динамических систем с дискретным временем и неопределенностью являются стохастические разностные уравнения. Наиболее распространенными разностными уравнениями являются уравнения авторегрессии

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

представляющие собой дискретные аналоги дифференциальных уравнений со случайным источником. В модели (1) состояние системы X_t в момент времени t зависит от состояния системы $X_{t-1}, \ldots, X_{t-p}$ в предыдущие моменты времени $t-1, \ldots, t-p$ и от случайных неподдающихся учету факторов, описываемых процессом ε_t, который называется инновационным или обновляющим процессом.

Важной задачей изучения систем, описываемых уравнением (1), является их идентификация, заключающаяся в оценивании коэффициентов $a_1, \ldots, a_p$ по наблюдениям $X_1, \ldots, X_n$ авторегрессионного процесса. В классическом случае, когда коэффициенты $a_1, \ldots, a_p$ неслучайны, существует достаточно хорошо разработанная теория их оценивания. Так, если распределение вероятности обновляющего процесса ε_t известно, то наилучшим методом оценивания является метод максимального правдоподобия [1], совпадающий с методом наименьших квадратов, если процесс ε_t является гауссовским (нормальным).

На практике, однако, наиболее распространена ситуация, когда среди большинства наблюдений X_t, хорошо описываемых каким-либо вероятностным распределением, например, нормальным, встречается небольшая доля нетипичных наблюдений, называемых выбросами. Основными причинами выбросов являются небрежность при сборе данных и неточность математической модели. В этом случае классические методы максимального правдоподобия и наименьших квадратов теряют свою эффективность, более того они могут выдавать абсурдные результаты [2].

Тем не менее существуют оценки, устойчивые к выбросам, они называются робастными. Наиболее распространенными робастными оценками коэффициентов $a_1, \ldots, a_p$ уравнения (1) являются M-оценки [3]. Известны аналитические выражения для предельного при $n \to \infty$ значения дисперсии M-оценок и оценки наименьших квадратов [3], эти пределы называются асимптотическими дисперсиями соответствующих оценок. В случаях, когда вероятностное распределение ε_t отлично от нормального, асимптотическая дисперсия M-оценок обычно меньше дисперсии оценки наименьших квадратов, поэтому M-оценки, как правило, предпочтительнее оценки наименьших квадратов.

Опыт практического применения модели (1) показал, что в большинстве случаев коэффициенты $a_1, \ldots, a_p$ с течением времени флуктуируют случайным образом вокруг своих средних значений под воздействием неизвестных, неучтенных внешних факторов, так что разумно коэффициенты $a_1, \ldots, a_p$ считать случайными величинами. Такие модели, называются моделями авторегрессии со случайными коэффициентами или RCA-моделями (сокращенно от random coefficient autoregressive) [4] и являются дискретными аналогами стохастических дифференциальных уравнений. RCA-модели широко применяются во многих областях науки и техники (см., например [5, 6, 7, 8, 9] и библиографию в них). В этом случае параметрами модели, подлежащими оцениванию, являются математические ожидания $\mathsf{E}[a_1], \ldots, \mathsf{E}[a_p]$ коэффициентов $a_1, \ldots, a_p$.

При оценке параметров RCA-моделей можно применять как метод наименьших квадратов [4], так и метод построения M-оценок. Однако в отличие от классических авторегрессионных моделей с неслучайными коэффициентами, вычислить аналитически асимптотическую дисперсию M-оценок в RCA-модели пока не удается. Таким образом, сравнить M-оценки с оценкой наименьших квадратов путем сравнения их асимптотических дисперсий пока не представляется возможным.

Целью настоящей работы является сравнение оценки наименьших квадратов и M-оценки параметров RCA-модели посредством оценивания их дисперсий методом имитационного компьютерного моделирования. Для простоты рассматривалась модель первого порядка ($p = 1$).

1. Оценка наименьших квадратов параметра процесса авторегрессии

Авторегрессионный процесс первого порядка со случайным коэффициентом X_t определяется разностным уравнением

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

где η_t и ε_t — случайные величины; a — параметр авторегрессионной модели (неслучайная составляющая авторегрессионного коэффициента $a + \eta_t$).

Предположим, что (η_t, ε_t), $t = 0,\, \pm 1,\, \pm 2,\, \dots,$ — последовательность независимых одинаково распределенных случайных векторов с независимыми координатами, имеющими нулевые математические ожидания $\mathsf{E}[\eta_t] = 0$, $\mathsf{E}[\varepsilon_t] = 0$ и конечные дисперсии $\mathsf{D}[\eta_t] = \omega^2$, $\mathsf{D}[\varepsilon_t] = \sigma^2$.

Мы будем предполагать также выполненным условие $a^2 + \omega^2 < 1$, которое, как показано в [4], является необходимым и достаточным условием существования строго стационарного решения уравнения (2) с конечным вторым моментом $\mathsf{E}[X_t^2]$, которое может быть представлено в виде сходящегося с вероятностью 1 ряда

$$X_t = \sum_{i=0}^{\infty} \delta_i \varepsilon_{t-i}, \tag{3}$$

где

$$\delta_0 = 1, \quad \delta_i = \prod_{j=0}^{i-1} (a + \eta_{i-j}), \quad i = 1,\, 2,\, \dots \tag{4}$$

Обозначим через a_0 истинное значение авторегрессионного параметра a в (2), а через $\mathfrak{F}_k$ — σ-алгебру, порожденную случайными величинами $\{\eta_i, \varepsilon_i\}$, $i \le k$.

Оценка наименьших квадратов $\tilde{a}_n$ параметра a_0 по наблюдениям $X_1, \dots, X_n$ уравнения (2) определяется как точка минимума функции

$$L_{LS}(a) = \sum_{t=1}^{n} (X_t - aX_{t-1})^2$$

и равна

$$\tilde{a}_n = \frac{\sum\limits_{t=2}^{n} X_t X_{t-1}}{\sum\limits_{t=2}^{n} X_{t-1}^2}. \tag{5}$$

В [4] показано, что при сделанных выше предположениях случайная величина $\sqrt{n}(\tilde{a}_n - a_0)$ асимптотически нормальна (стремится по распределению при $n \to \infty$ к нормальной случайной величине) с нулевым математическим ожиданием и дисперсией

$$\frac{(\sigma^2 \mathsf{E}[X_0^2] + \omega^2 \mathsf{E}[X_0^4])}{(\mathsf{E}[X_0^2])^2}.$$

2. М-оценка параметра процесса авторегрессии

Определим М-оценку $\hat{a}_n$ параметра a_0 как точку минимума функции

$$L_M(a) = \sum_{t=1}^{n} \rho(X_t - aX_{t-1}), \tag{6}$$

где ρ — некоторая функция. Наиболее распространенными классами таких функций являются (см. [2]) ρ-функция Хьюбера

$$\rho_H(x) = \begin{cases} x^2, & |x| \le k; \\ 2k|x| - k^2, & |x| > k, \end{cases} \tag{7}$$

и бивес Тьюки

$$\rho_T(x) = \begin{cases} 1 - \left(1 - \left(\dfrac{x}{k}\right)^2\right)^3, & |x| \le k; \\ 1, & |x| > k. \end{cases} \tag{8}$$

Функции $\rho_H(x)$ и $\rho_T(x)$ зависят от параметра $k > 0$, изменение которого позволяет регулировать устойчивость оценки к сильно выделяющимся наблюдениям.

Заметим, что при $\rho(x) = x^2$ М-оценка совпадает с оценкой наименьших квадратов.

Предположим, что $\rho(x)$ достаточно гладкая функция. Тогда, в частности, точка минимума функции L_M является решением уравнения

$$L'_M(a) = \sum_{t=1}^{n} \rho'(X_t - aX_{t-1})(-X_{t-1}) = 0.$$

Заметим, что

$$X_t - aX_{t-1} = (X_t - a_0 X_{t-1}) + (a_0 - a)X_{t-1} = \eta_t X_{t-1} + \varepsilon_t - (a_0 - a)X_{t-1}.$$

Разложим $L'_M(a)$ по формуле Тейлора в окрестности точки a_0 и обозначим через $\hat{\alpha}_n$ решение относительно a уравнения

$$L'_M(a_0) + L''_M(a_0)(a - a_0) = 0,$$

что равносильно решению уравнения

$$\sum_{t=1}^{n} \rho'(\eta_t X_{t-1} + \varepsilon_t)X_{t-1} - \sum_{t=1}^{n} \rho''(\eta_t X_{t-1} + \varepsilon_t)(a - a_0)X_{t-1}^2 = 0.$$

Рассуждая так же как и при доказательстве теоремы 1 из [10], получим, что если функция ρ выпуклая, то $\sqrt{n}(\hat{a}_n - \hat{\alpha}_n) \to 0$ по вероятности при $n \to \infty$, откуда следует, что асимптотические распределения оценок $\hat{a}_n$ и $\hat{\alpha}_n$ совпадают. Поэтому асимптотическое распределение $\sqrt{n}(\hat{a}_n - a_0)$ совпадает с асимптотическим распределением

$$\sqrt{n}(\hat{\alpha}_n - a_0) = \frac{\dfrac{1}{\sqrt{n}} \sum\limits_{t=1}^{n} \rho'(\eta_t X_{t-1} + \varepsilon_t)X_{t-1}}{\dfrac{1}{n} \sum\limits_{t=1}^{n} \rho''(\eta_t X_{t-1} + \varepsilon_t)X_{t-1}^2}.$$

Согласно закону больших чисел

$$\frac{1}{n}\sum_{t=1}^{n}\rho''(\eta_t X_{t-1}+\varepsilon_t)X_{t-1}^2 \to \mathsf{E}[\rho''(\eta_1 X_0+\varepsilon_1)X_0^2], \quad n\to\infty,$$

а согласно центральной предельной теореме при $n\to\infty$ случайная величина

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\rho'(\eta_t X_{t-1}+\varepsilon_t)X_{t-1}$$

асимптотически нормальна с нулевым математическим ожиданием и дисперсией

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

Следовательно, случайная величина $\sqrt{n}(\hat{\alpha}_n - a_0)$ (а стало быть, и случайная величина $\sqrt{n}(\hat{a}_n - a_0)$) асимптотически нормальна с нулевым математическим ожиданием и дисперсией

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

В частности, для оценки наименьших квадратов $\rho(x) = x^2$, $\rho'(x) = 2x$, $\rho''(x) = 2$. Поэтому

$$\frac{\mathsf{E}[(\rho'(\eta_1 X_0+\varepsilon_1)X_0)^2]}{(\mathsf{E}[\rho''(\eta_1 X_0+\varepsilon_1)X_0^2])^2} = \frac{\omega^2\mathsf{E}X_0^4+\sigma^2\mathsf{E}[X_0^2]}{(\mathsf{E}[X_0^2])^2}.$$

3. Модели загрязнения наблюдений

В теории временных рядов наиболее распространены три модели ошибок наблюдений: аддитивная, замещающая и инновационная [3].

В аддитивной модели вместо X_t наблюдается процесс Y_t вида

$$Y_t = X_t + \nu_t\zeta_t, \tag{9}$$

где ν_t — случайный процесс с независимыми значениями и

$$\mathsf{P}\{\nu_t = 1\} = \delta, \quad \mathsf{P}\{\nu_t = 0\} = 1-\delta, \quad 0<\delta<1. \tag{10}$$

Другими словами, на наблюдение X_t случайным образом с вероятностью δ накладывается выброс ζ_t, который можно интерпретировать как результат сбоя некоторых узлов измерительного устройства. Будем предполагать, что ζ_t — процесс с независимыми значениями, общей для всех ζ_t функцией распределения F_ζ и конечной дисперсией $\sigma_\zeta^2 = \mathsf{D}[\zeta_t]$.

В замещающей модели наблюдения Y_t имеют вид

$$Y_t = (1-\nu_t)X_t + \nu_t\zeta_t, \tag{11}$$

где ν_t и δ описываются (10), т.е. с вероятностью δ вместо X_t наблюдается процесс ζ_t. Таким образом, замещающая модель имитирует полный отказ с вероятностью δ измерительной аппаратуры.

Обычно в моделях (9)–(11) процесс ζ_t гауссовский с дисперсией, значительно большей, чем дисперсия наблюдаемого временного ряда X_t. Мы будем предполагать, что случайные процессы X_t, ν_t и ζ_t не зависят друг от друга и являются стационарными в широком смысле. Отметим, что в моделях (9)–(11) выброс ζ_t в фиксированный момент времени t влияет только на наблюдаемый процесс в этот же момент времени и не влияет на все последующие наблюдения.

В инновационной модели выброс воздействует на обновляющий процесс ε_t. Обозначим через f_0 плотность распределения вероятности случайной величины ε_t в отсутствие выбросов. Тогда при наличии выбросов с небольшой вероятностью δ случайные величины ε_t меняют плотность с f_0 на некоторую другую плотность f_1, т.е. плотность распределения вероятности f_ε случайной величины ε_t при наличии выбросов имеет вид

$$f_\varepsilon(x) = (1-\delta)f_0(x) + \delta f_1(x). \tag{12}$$

Можно также представлять себе инновационный выброс как импульс на входе динамической системы (2), а X_t — как реакцию системы на этот импульс. Заметим, что инновационный выброс воздействует не только на текущее наблюдение, но и на все последующие. Таким образом, в инновационной модели $Y_t = X_t$, где X_t удовлетворяет (2), в котором плотность распределения вероятности f_ε случайной величины ε_t имеет вид (12).

Важным примером инновационной модели выбросов является загрязнённое нормальное распределение [11], называемого также распределением Тьюки, плотность которого имеет вид

$$f_\varepsilon(x) = (1-\delta)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} + \delta\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{x^2}{2\tau^2}}, \quad 0 \le \delta \le 1. \tag{13}$$

Последовательность случайных величин, имеющих распределение Тьюки, имитирует типичное на практике загрязнение последовательности центрированных нормальных величин с дисперсией 1 небольшой (0,01–0,15) долей δ центрированных нормальных величин с дисперсией $\tau^2 > 1$.

4. Асимптотическая относительная эффективность М-оценки по отношению к оценке наименьших квадратов

Одним из стандартных способов сравнения качества двух асимптотически нормальных оценок является вычисление их асимптотической относительной эффективности, которая определяется как величина, равная обратному отношению их асимптотических дисперсий. Таким образом, асимптотическая относительная эффективность М-оценки относительно оценки наименьших квадратов есть

$$e = \frac{(\omega^2 \mathsf{E}[X_0^4] + \sigma^2 \mathsf{E}[X_0^2])\,(\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2])^2}{(\mathsf{E}[X_0^2])^2\,\mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2]}. \tag{14}$$

Асимптотическая относительная эффективность показывает, во сколько раз больше наблюдений нужно взять для того, чтобы оценка наименьших квадратов достигла такой же точности, что и М-оценка. Если $e > 1$, то М-оценка точнее, чем оценка наименьших квадратов. Если же $e < 1$, то, наоборот, оценка наименьших квадратов точнее М-оценки.

Случай постоянного коэффициента авторегрессии. В частном случае, когда авторегрессионный коэффициент $a+\eta_t$ в модели (2) является неслучайным, т.е. когда $\eta_t \equiv 0$, $\omega = 0$, формула (14) заметно упрощается:

$$e = \sigma^2 \frac{(\mathsf{E}[\rho''(\varepsilon_1)])^2}{\mathsf{E}[(\rho'(\varepsilon_1))^2]} = \frac{\int\limits_{-\infty}^{\infty} x^2 f_\varepsilon(x)\,dx \left(\int\limits_{-\infty}^{\infty} \rho''(x) f_\varepsilon(x)\,dx \right)^2}{\int\limits_{-\infty}^{\infty} (\rho'(x))^2 f_\varepsilon(x)\,dx} \tag{15}$$

и для различных $\rho(x)$ и $f_\varepsilon(x)$ может быть вычислена аналитически или при помощи численного интегрирования.

Например, если ε_t имеют распределение Тьюки (13), а ρ-функция является ρ-функцией Хьюбера (7), то

$$e = \frac{4(1+\tau^2\delta-\delta)\left((1-\delta)\Phi_0(k)+\delta\Phi_0\left(\frac{k}{\tau}\right)\right)^2}{k^2+2(1-\delta)(1-k^2)\Phi_0(k)+2\delta(\tau^2-k^2)\Phi_0\left(\frac{k}{\tau}\right)+\frac{k\sqrt{2}}{\sqrt{\pi}}\left((\delta-1)e^{-\frac{k^2}{2}}-\tau\delta e^{-\frac{k^2}{2\tau^2}}\right)}, \tag{16}$$

где

$$\Phi_0(x) = \int\limits_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

— функция Лапласа. Аналогично можно получить зависимость e от τ, δ и k для ρ-функция Хампеля (8).

Случай стохастического коэффициента авторегрессии. Если $\eta_t \not\equiv 0$, то найти распределение вероятности случайных величин X_t, зная распределение вероятности случайных величин (η_t, ε_t) не представляется возможным ввиду сложной зависимости (3)–(4) процесса X_t от (η_t, ε_t). Следовательно, нельзя найти распределение вероятности случайных величин $\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2$ и $\rho'(\eta_1 X_0 + \varepsilon_1)X_0^2$ и выразить правую часть (14) через интегралы. Поэтому для оценивания относительной эффективности М-оценки по отношению к оценке наименьших квадратов использовалось компьютерное моделирование.

Реализации $X_1, \ldots, X_n$ длины $n = 500$ процесса X_t строились по рекуррентной формуле (2) с начальным условием $X_0 = 0$. Случайные величины η_t предполагались гауссовскими с математическим ожиданием $\mathsf{E}[\eta_t] = 0$ и дисперсией $\mathsf{D}[\eta_t] = 0{,}01$. Плотность f_ε распределения вероятности случайных величин ε_t задавалась выражением (12). Реализации η_t и ε_t получались с использованием датчика нормально распределенных псевдослучайных чисел. Минимум функции (6) с ρ-функцией Хьюбера (7) для вычисления М-оценки находился методом Нелдера — Мида. Оценка наименьших квадратов вычислялась по формуле (5).

Описанный эксперимент повторялся $N = 20000$ раз; по последовательности оценок наименьших квадратов $\tilde{a}_{n1}, \ldots, \tilde{a}_{nN}$ и последовательности М-оценок $\hat{a}_{n1}, \ldots, \hat{a}_{nN}$ величиной

$$e_{nN} = \frac{\sum\limits_{i=1}^{N} (\tilde{a}_{ni} - a_0)^2}{\sum\limits_{i=1}^{N} (\hat{a}_{ni} - a_0)^2},$$

оценивалась относительная эффективность М-оценки по отношению к оценке наименьших квадратов, поскольку e_{nN} при $N \to \infty$ стремится к относительной эффективности М-оценки по отношению к оценке наименьших квадратов для фиксированного объема наблюдений n. Результаты эксперимента приведены на рис. 1, 2.

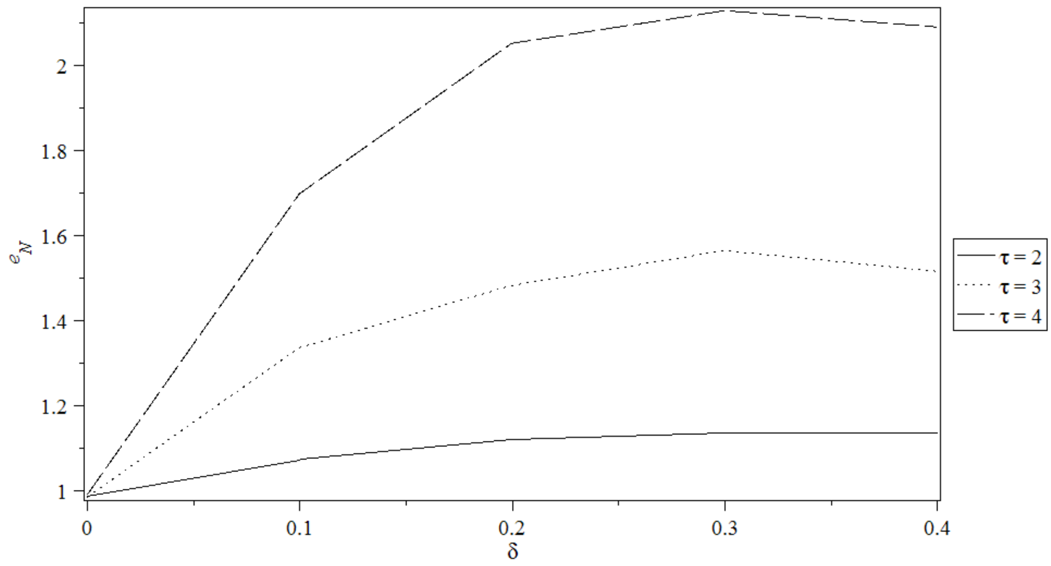

Рис. 1. Зависимость e_{nN} от δ при различных τ

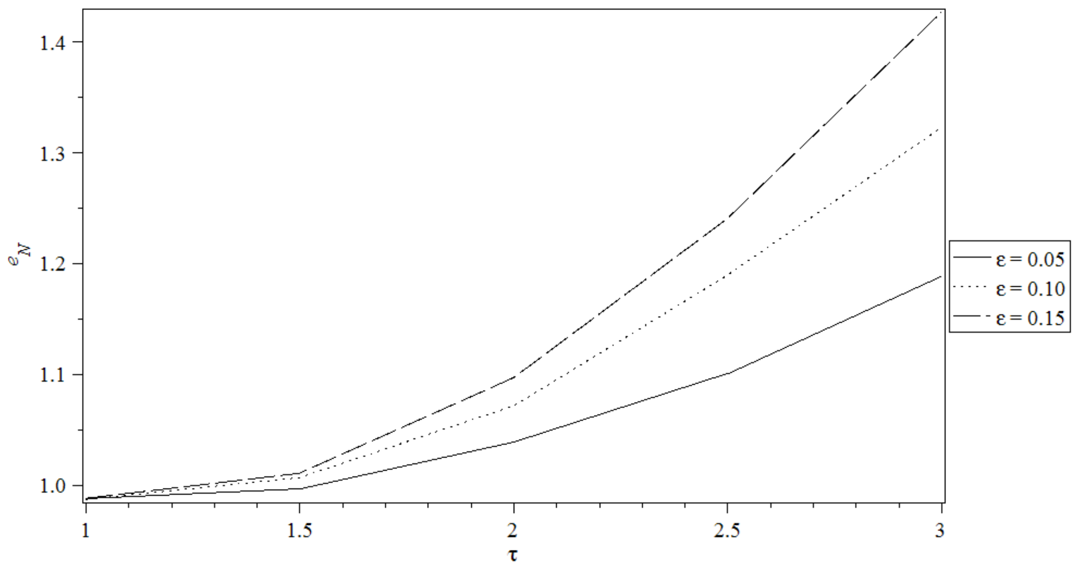

Рис. 2. Зависимость e_{nN} от τ при различных δ

На рис. 1 приведена зависимость e_{nN} от δ при различных τ, на рис. 2 — зависимость e_{nN} от τ при различных δ. Постоянная k ρ-функции Хьюбера во всех экспериментах равнялась 2, поскольку опытным путем было установлено, что если f_ε задается (13) и $\delta \in (0; 0{,}3)$, $\tau \in (1; 5)$, то выборочная дисперсия М-оценки принимает минимальные значения при $k \in (1{,}5; 2{,}5)$.

Из рисунков видно, что в отсутствие загрязнений оценка наименьших квадратов немного эффективнее М-оценки. Однако с ростом δ и τ относительная эффективность М-оценки по отношению к оценке наименьших квадратов увеличивается, становясь больше единицы. Например, если $\delta = 0{,}01$, то М-оценка эффективнее оценки наименьших квадратов уже при $\tau > 2$.

Кроме того, выяснилось, что относительная эффективность практически не зависит от дисперсии авторегрессионного коэффициента, если ее изменение удовлетворяет условию стационарности $a^2 + \omega^2 < 1$.

Исследования также показали, что в условиях аддитивной и замещающей моделей наблюдения (9)–(11) авторегрессионного процесса обе оценки быстро теряют не только свою эффективность, но и состоятельность.

Заключение

В работе изучены свойства М-оценки и оценки наименьших квадратов параметра уравнения авторегрессии со случайным коэффициентом в предположении, что расспределение вероятности обновляющей авторегрессионной последовательности отклоняется от гауссовского. Нарушения в предположении гауссовости заключались в том, что исходная обновляющая последовательность стандартных нормальных величин «загрязнялась» случайным образом небольшой долей δ центрированных гауссовских случайных величин с бóльшей дисперсией τ^2 ($\tau > 1$). Оказалось, что в отсутствии загрязнений ($\delta = 0$ или $\tau = 1$) оценка наименьших квадратов лишь незначительно превосходит М-оценку (в $0{,}987$ раз). Другими словами, в отсутствие загрязнений для достижения одинаковой точности М-оценке нужно всего лишь в $1/0{,}987 = 1{,}013$ раз больше наблюдений, чем оценке наименьших квадратов. Однако с увеличением δ и τ М-оценка становится эффективнее оценки наименьших квадратов. Например, если $\delta = 0{,}05$, то М-оценка эффективнее оценки наименьших квадратов при $\tau > 1{,}53$, а если же $\delta = 0{,}15$, то при $\tau > 1{,}27$. Поскольку, как показывает практика, в реальных данных обычно присутствует до 15% загрязнений, то при оценивании параметра авторегрессии со случайным коэффициентом М-оценки следует предпочесть оценке наименьших квадратов.

Список литературы

1. Kirchgassner G., Wolters J. Introduction to Modern Time Series Analysis. Berlin: Springer, 2007. 277 p. DOI: 10.1007/978-3-540-73291-4
2. Huber P., Ronchetti E.M. Robust Statistics, Second Edition. Hoboken: Wiley, 2009. 360 p. (Ser. Wiley Series in Probability and Statistics). DOI: 10.1002/9780470434697

3. Maronna R.A., Martin D., Yohai V. Robust Statistics: Theory and Methods. Chichester: Wiley, 2006. 403 p. (Ser. Wiley Series in Probability and Statistics). DOI: 10.1002/0470010940
4. Nicholls D.F., Quinn B.G. Random Coefficient Autoregressive Models: An Introduction. New York: Springer, 1982. 154 p. (Lecture Notes in Statistics; vol. 11). DOI: 10.1007/978-1-4684-6273-9
5. Bera A.K., Garcia P., Roh J.-S. Estimation of Time-varying Hedge Ratios for Corn and Soybeans: Bgarch and Random Coefficient Approaches // Sankhya: The Indian Journal of Statistics, Series B. 1997. Vol. 59, № 3. P. 346–368.
6. Rahiala M. Miscellanea. Random coefficient autoregressive models for longitudinal data // Biometrika. 1999. Vol. 86, № 3. P. 718–722. DOI: 10.1093/biomet/86.3.718
7. Sàfadi T., Morettin P.A. Bayesian analysis of autoregressive models with random normal coefficients // J. Stat. Comput. Simul. 2003. Vol. 73, № 8. P. 563–573. DOI: 10.1080/0094965031000136003
8. Lee H.T., Yoder K.J., Mittlehammer R.C., McCluskey J.J. A Random Coefficient Autoregressive Markov Regime Switching Model for Dynamic Futures Hedging // The Journal of Futures Markets. 2006. Vol. 26, № 2. P. 103–129. DOI: 10.1002/fut.20193
9. Tong H. Non-linear Time Series: A Dynamical System Approach. Oxford: Clarendon Press, 1990. 564 p.
10. Goryainov V.B. M-estimates of the spatial autoregression coefficients // Automation and Remote Control. 2012. Vol. 73, № 8. P. 1371–1379. DOI: 10.1134/S0005117912080103
11. Wilcox R.R. Introduction to Robust Estimation and Hypothesis Testing. Amsterdam: Elsevier, 2012. 690 p.

Science & Education of the Bauman MSTU
Electronic journal
ISSN 1994-0408

Science and Education of the Bauman MSTU, 2016, no. 9, pp. 111–122.
DOI: **10.7463/0916.0844976**
Received: 25.08.2016

Robust estimation in random coefficient autoregression model

Goryainov V. B.[1,*], Ermakov S. Yu.[1]
[*]vb-goryainov@bmstu.ru
[1]Bauman Moscow State Technical University, Russia

Keywords: least square estimate, M-estimate, autoregression with random coefficient, Tukey distribution

We study the robust properties of the M-estimate and least square estimate of parameter of autoregressive equation of the first order with a random coefficient. The autoregressive equations with random coefficients are discrete analogues of stochastic differential equations, and estimation of autoregressive parameters is important task for identification of dynamic systems described by these equations.

The objective is to compare the least square estimate and M-estimate of the autoregressive parameter. For simplicity, we consider the equation of the first order the update sequence of which is a sequence of independent identically distributed random variables with zero mean and finite variances. An auto-regression coefficient is a sequence of independent and identically distributed normal random variables with constant expectation and finite variances. The parameter of the autoregressive equation to be estimated is a mean value of the auto-regression coefficient of equation. A solution of the autoregressive equation is assumed to be stationary in the broad sense.

We propose to use the Huber and Hampel co-functions to provide M-estimate and a Nelder - Mead method to minimize the objective function for the M-estimation. We have considered the additive, replacement and innovative models of observation errors. The innovative model assumes that the updating process of the autoregressive equation has Tukey distribution.

Using a computer simulation with a large number of implementations of the autoregressive process allowed us to calculate the sampling variances of M-estimation and least square estimate; the relative efficiency of estimations was approximated by the inverse ratio of sampling variances.

A dependence of the relative effectiveness on the Tukey distribution parameters was defined. The paper explores how the auto-regression coefficient, observation errors of the additive and replacement models influence on the relative efficiency and also gives recommended practices of M-estimate and least square estimate.

References

1. Kirchgassner G., Wolters J. *Introduction to Modern Time Series Analysis*. Berlinб Springer, 2007, 277 p. DOI: 10.1007/978-3-540-73291-4
2. Huber P., Ronchetti E.M. *Robust Statistics, Second Edition*. Hoboken, Wiley, 2009, 360 p. (Ser. Wiley Series in Probability and Statistics). DOI: 10.1002/9780470434697
3. Maronna R.A., Martin D., Yohai V. *Robust Statistics: Theory and Methods*. Chichester, Wiley, 2006, 403 p. (Ser. Wiley Series in Probability and Statistics). DOI: 10.1002/0470010940
4. Nicholls D.F., Quinn B.G. *Random Coefficient Autoregressive Models: An Introduction*. New York, Springer, 1982, 154 p. (Lecture Notes in Statistics; vol. 11). DOI: 10.1007/978-1-4684-6273-9
5. Bera A.K., Garcia P., Roh J.-S. Estimation of Time-varying Hedge Ratios for Corn and Soybeans: Bgarch and Random Coefficient Approaches. *Sankhya: The Indian Journal of Statistics, Series B*, 1997, vol. 59, no. 3, pp. 346–368.
6. Rahiala M. Miscellanea. Random coefficient autoregressive models for longitudinal data. *Biometrika*, 1999, vol. 86, no. 3, pp. 718–722. DOI: 10.1093/biomet/86.3.718
7. Sàfadi T., Morettin P.A. Bayesian analysis of autoregressive models with random normal coefficients. *J. Stat. Comput. Simul.*, 2003, vol. 73, no. 8, pp. 563–573. DOI: 10.1080/0094965031000136003
8. Lee H.T., Yoder K.J., Mittlehammer R.C., McCluskey J.J. A Random Coefficient Autoregressive Markov Regime Switching Model for Dynamic Futures Hedging. *The Journal of Futures Markets*, 2006, vol. 26, no. 2, pp. 103–129. DOI: 10.1002/fut.20193
9. Tong H. *Non-linear Time Series: A Dynamical System Approach*. Oxford, Clarendon Press, 1990, 564 p.
10. Goryainov V.B. M-estimates of the spatial autoregression coefficients. *Automation and Remote Control*, 2012, vol. 73, no. 8, pp. 1371–1379. DOI: 10.1134/S0005117912080103
11. Wilcox R.R. *Introduction to Robust Estimation and Hypothesis Testing*. Amsterdam, Elsevier, 2012, 690 p.