Математика и Математическое моделирование
Сетевое научное издание МГТУ им. Н.Э. Баумана
http://mathm.elpub.ru

Ссылка на статью:
// Математика и математическое моделирование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2017. № 3. С. 91–104.
DOI: **10.24108/mathm.0317.0000072**
Представлена в редакцию: 07.06.2017

УДК 519.234.3

Сравнение классических и робастных оценок параметров пороговой авторегрессии

Горяинов В. Б.[1,*]
[*]vb-goryainov@bmstu.ru
[1]МГТУ им. Н.Э. Баумана, Москва, Россия

Рассматривается задача оценивания параметров пороговой авторегрессии с одним порогом, расположение порога предполагается известным. При помощи компьютерного моделирования изучается относительная эффективность М-оценки по отношению к оценке наименьших квадратов и оценке наименьших модулей для различных вероятностных распределений обновляющего процесса: нормального, логистического, двойного-экспоненциального, Коши, Тьюки и Стьюдента с различным числом степеней свободы. М-оценки вычисляются на основе ро-функций Хьюбера и Тьюки. Рассматривается инновационная модель ошибок наблюдений. Получена зависимость относительной эффективности от параметров распределения Тьюки.

Ключевые слова: пороговая авторегрессия; М-оценка; оценка наименьших квадратов; оценка наименьших модулей; загрязненное нормальное распределение

Введение

В последние годы в различных областях науки и техники при описании случайных процессов с дискретным временем большое распространение получила пороговая авторегрессионная модель [1]. В отличие о классической модели авторегрессии пороговая авторегрессионная модель позволяет обнаружить в поведении стохастических дискретных последовательностей ряд нелинейных эффектов [2]. В данной работе рассматривается наиболее простой вариант модели пороговой авторегрессии с единственным порогом, который предполагается известным и без ограничения общности равным нулю.

Важнейшей задачей, возникающей при исследовании пороговой авторегрессионной модели, является оценивание ее параметров — коэффициентов соответствующего порогового уравнения. Наиболее распространенными оценками являются оценка наименьших квадратов и оценка наименьших модулей, менее известны М-оценки. В настоящей работе проводится сравнительный анализ вероятностных свойств перечисленных оценок параметров пороговой авторегрессионной модели при наиболее распространенных вероятностных распределениях обновляющего процесса порогового уравнения. Даны рекомендации по использованию указанных оценок.

1. Пороговая модель авторегрессии

Модель пороговой авторегрессии с m режимами описывается рекуррентным уравнением [3, p. 77]

$$X_t = \sum_{j=1}^{m} (a_{1j} X_{t-1} + \ldots + a_{p_j j} X_{t-p_j}) \, I_t(\vartheta_{j-1} \le v_{t-d} < \vartheta_j) + \varepsilon_t,$$

где ε_t — обновляющий процесс; $I_t(\cdot)$ — индикаторная функция; $-\infty = \vartheta_0 < \vartheta_1 < \ldots < \vartheta_m = \infty$ — пороговые параметры; p_j и d — натуральные числа, описывающие авторегрессионные лаги и пороговое запаздывание соответственно; a_{ij} — авторегрессионные коэффициенты. Эта модель является нелинейной во времени t и кусочно-линейной в пространстве пороговых параметров Θ. Переключающая переменная v_{t-d} может иметь как внешнее (экзогенное) происхождение, так и определяться самой моделью. Если v_{t-d} полностью определяется величинами X_s, $s \le t-d$, например, совпадает с X_{t-d}, то пороговая модель называется самовозбуждающейся.

В данной работе рассматривается самовозбуждающаяся пороговая авторегрессионная модель первого порядка с одним переключением в нуле. В этой модели наблюдения X_t описываются уравнением

$$X_t = a_1 X_{t-1} I_t(0 \le X_{t-1}) + a_2 X_{t-1} I_t(X_{t-1} \le 0) + \varepsilon_t,$$

или, что равносильно, кусочно-разностным уравнением

$$X_t = \begin{cases} a_1 X_{t-1} + \varepsilon_t, & X_{t-1} \ge 0; \\ a_2 X_{t-1} + \varepsilon_t, & X_{t-1} < 0. \end{cases}$$

Эквивалентным образом уравнение пороговой модели может быть записано в виде

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

где $X_t^{+} = \max(X_t, 0)$, $X_t^{-} = \min(X_t, 0)$. Действительные числа a_1, a_2 называются параметрами авторегрессионной пороговой модели и считаются неизвестными.

Далее всюду будем предполагать, что ε_t — последовательность независимых одинаково распределенных непрерывных случайных величин с плотностью f, нулевым математическим ожиданием $\mathsf{E}\varepsilon_t = 0$ и конечной дисперсией $\mathsf{D}\varepsilon_t = \sigma^2$.

Будем также считать, что процесс X_t является стационарным. Для этого, например, достаточно одновременное выполнение условий $a_1 < 1$, $a_2 < 1$, $a_1 a_2 < 1$, $f(x) > 0$ для любых $x \in \mathbb{R}$ и существование $\mathsf{E}|\varepsilon_t|^{2+\delta} < \infty$ для некоторого $\delta > 0$ [4].

2. Классические оценки авторегрессионных параметров

Важнейшей задачей, возникающей при исследовании (1), является оценивание векторного параметра $a = (a_1,\ a_2)$ по наблюдениям процесса X_t. В данной работе рассматриваются оценка наименьших квадратов, оценка наименьших модулей и М-оценки.

Оценка наименьших квадратов $a^* = (a_1^*,\ a_2^*)$ параметра a, построенная по наблюдениям $X_1, X_2, \ldots, X_n$, определяется как точка 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,

$$a_1^* = \frac{\sum\limits_{t=1}^{n} X_t X_{t-1}^+}{\sum\limits_{t=1}^{n} (X_{t-1}^+)^2}, \qquad a_2^* = \frac{\sum\limits_{t=1}^{n} X_t X_{t-1}^-}{\sum\limits_{t=1}^{n} (X_{t-1}^-)^2}. \tag{2}$$

Согласно [4], оценка наименьших квадратов a^* асимптотически нормальна с ковариационной матрицей $\sigma^2 K^{-1}$, где

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

Другими словами, при $n \to \infty$ последовательность $\sqrt{n}(a^* - a)$ сходится по распределению к двумерному нормальному вектору с нулевым математическим ожиданием и ковариационной матрицей $\sigma^2 K^{-1}$, т.е. для любых действительных x и y

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

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt \quad —$$

функция распределения вероятности стандартной нормальной случайной величины, $\sigma^2 = \mathsf{E}\varepsilon_t^2$ — общая дисперсия всех ε_t, $t = 1,\ 2,\ \ldots$

Основной недостаток метода наименьших квадратов заключается в его сильной чувствительности к большим значениям невязок $\varepsilon_t = X_t - a_1 X_{t-1}^+ - a_2 X_{t-1}^-$, поскольку они влияют на минимизируемую функцию $L_{LSQ}(a)$ квадратичным образом. Вследствие этого точность оценки наименьших квадратов достаточно сильно ухудшается при росте вероятности экстремальных значений ε_t, что происходит, если плотность распределения вероятности f случайных величин ε_t медленно убывает на бесконечности.

Этого недостатка лишен метод наименьших модулей. Оценкой наименьших модулей $\tilde{a} = (\tilde{a}_1,\ \tilde{a}_2)$ параметра a, построенной по наблюдениям $X_1, X_2, \ldots, X_n$, называется точка минимума функции

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

Согласно [5], случайная последовательность $\sqrt{n}(\tilde{a} - a)$ асимптотически нормальна с нулевым математическим ожиданием и ковариационной матрицей $\frac{1}{4f^2(0)} K^{-1}$.

В методе наименьших модулей влияние невязок $\varepsilon_t = X_t - a_1 X_{t-1}^+ - a_2 X_{t-1}^-$ на минимизируемую функцию $L_{LAD}(a)$ линейно. Поэтому при уменьшении скорости сходимости

f к нулю на бесконечности точность оценки наименьших модулей падает не так сильно как точность оценки наименьших квадратов [6]. Однако в случае нормального распределения ε_t, плотность которого $f(x)$ с ростом $|x|$ стремится к нулю достаточно быстро, метод наименьших квадратов гораздо эффективнее метода наименьших модулей.

М-оценка

М-оценка вектора a по наблюдениям $X_1, X_2, \ldots, X_n$ определяется как точка минимума $\hat{a} = (\hat{a}_1,\ \hat{a}_2)$ функции

$$L_M(a) = \sum_{t=1}^{n} \rho(X_t - a_1 X_{t-1}^{+} - a_2 X_{t-1}^{-}), \tag{3}$$

где ρ — некоторая функция. М-оценки образуют целое семейство, зависящее от вида функции ρ. Свое название М-оценка получила из-за того, что если $\rho(x) = -\ln f(x)$, то М-оценка совпадает с оценкой максимального правдоподобия. Оценки наименьших квадратов и наименьших модулей также являются частными случаями М-оценок, поскольку получаются из них при $\rho(x) = x^2$ и $\rho(x) = |x|$ соответственно.

М-оценки являются разумным компромиссом между оценкой наименьших квадратов и оценкой наименьших модулей. Например, ρ-функция

$$\rho_H(x) = \begin{cases} x^2, & |x| \le k; \\ 2k|x| - k^2, & |x| > k, \end{cases} \tag{4}$$

называемая ρ-функцией Хьюбера (см. [7]), совпадает с x^2 в окрестности $(-k,\ k)$ начала координат и ведет себя линейно вне этой окрестности. В этом случае, вклад в сумму (3) наблюдений X_t, сформировавшихся под влиянием экстремальных значений ε_t, будет принижен по сравнению с вкладом остальных наблюдений. Параметр k можно изменять от нуля до бесконечности, подстраиваясь под конкретный вид f для достижения максимальной эффективности оценки.

Функция (4) является наиболее распространенным классом ρ-функций. Еще одним важным классом является (см. [7]) бивес Тьюки

$$\rho_T(x) = \begin{cases} 1 - \left(1 - \left(\dfrac{x}{k}\right)^2\right)^3, & |x| \le k; \\ 1, |x| > k. \end{cases} \tag{5}$$

Как видно из (3), (5), ρ-функция Тьюки игнорирует резко выделяющиеся невязки, заменяя их единицей.

3. Модель загрязнения наблюдений

Будем предполагать, что обновляющий процесс ε_t описывается инновационной загрязняющей моделью, суть которой заключается в том, что последовательность одинаково распределенных случайных величин засоряется случайными величинами (выбросами) с другим вероятностным распределением. Обозначим через f_0 плотность распределения вероятности случайной величины ε_t при отсутствии выбросов. Предположим, что в результате выбросов с небольшой вероятностью δ случайные величины ε_t меняют плотность с f_0 на некоторую другую плотность f_1. Таким образом, будем считать, что плотность распределения вероятности f случайной величины ε_t имеет вид

$$f(x) = (1-\delta)f_0(x) + \delta f_1(x).$$

Можно представлять себе выброс как импульс на входе динамической системы (1), а X_t — как реакцию системы на этот импульс. Отметим, что выброс в инновационной модели воздействует не только на текущее наблюдение $X(t)$, но и на все последующие.

Важным примером инновационной загрязняющей модели является загрязненное нормальное распределение [8], называемое также распределением Тьюки, плотность которого имеет вид

$$f(x) = (1-\delta)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} + \delta\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{x^2}{2\tau^2}}, \quad 0 \le \delta \le 1. \tag{6}$$

Последовательность случайных величин, имеющих распределение Тьюки, имитирует типичное на практике загрязнение последовательности центрированных нормальных величин с дисперсией 1 небольшой долей δ центрированных нормальных величин с дисперсией $\tau^2 > 1$. На практике обычно $\delta \in (0.01,\ 0.15)$.

4. Сравнение оценок

Среди двух оценок наилучшей логично считать ту, рассеяние которой вокруг оцениваемого параметра меньше. Если оценка является несмещенной, то рассеяние оценки измеряется ее дисперсией, и поэтому из двух несмещенных оценок лучшей будет оценка с меньшей дисперсией. К сожалению, для фиксированного объема наблюдений дисперсию оценок вычислить можно лишь в самых простых случаях.

Преодолевается эта трудность обычно одним из следующих способов. Во-первых, можно попытаться найти асимптотическое распределение сравниваемых оценок. Часто сравниваемые оценки являются асимптотически нормальными. В этом случае сравнить точность двух оценок можно, сравнив их асимптотические дисперсии. Во-вторых, для оценивания дисперсий оценок можно прибегнуть к компьютерному моделированию.

Как уже упоминалось выше, оценки наименьших квадратов и наименьших модулей являются асимптотически несмещенными и асимптотически нормальными. К сожалению, асимптотическое распределение M-оценок параметров пороговой авторегрессионной модели неизвестно. Поэтому сравним точность M-оценок по отношению к оценкам наименьших квадратов и наименьших модулей при помощи компьютерного моделирования.

Распределение вероятности обновляющего процесса ε_t будем моделировать типичными вероятностными распределениями, они перечислены в левом столбце табл. 1. А именно, сравним изучаемые оценки для двойного экспоненциального распределения (распределения

Т а б л и ц а 1

Оценки дисперсий оценок наименьших квадратов (ОНК), наименьших модулей (ОНМ), М-оценок с ρ-функцией Хьюбера (М-Х) и ρ-функцией Тьюки (М-Т) при различных распределениях ε_t

Распределение ε_t	ОНК	ОНМ	М-Х	М-Т
Нормальное	0.04066	0.06389	0.04108	0.04315
Лапласа	0.08168	0.04104	0.06497	0.05800
Логистическое	0.13415	0.16266	0.12323	0.14899
Стьюдента (20)	0.04618	0.06555	0.04517	0.04698
Стьюдента (10)	0.04799	0.06642	0.04678	0.04836
Стьюдента (5)	0.06810	0.07054	0.05694	0.05496
Коши	0.39403	0.10054	0.14333	0.09466

Лапласа) с плотностью $f(x) = \dfrac{1}{2}e^{-|x|}$, распределения Коши с плотностью $f(x) = \dfrac{1}{\pi(x^2+1)}$, логистического распределения с плотностью $f(x) = \dfrac{e^{-x}}{(1+e^{-x})^2}$ и распределения Стьюдента с плотностью

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

и числом степеней свободы m, равным 5, 10 и 20.

Реализации $X_1, \ldots, X_n$ длины $n = 500$ процесса X_t построим по рекуррентной формуле (1) с начальным условием $X_0 = 0$. Авторегрессионные параметры для определенности положим равными $a = (-0.3,\, 0.5)$. Реализации ε_t смоделируем при помощи датчика псевдослучайных чисел для соответствующей плотности f. Минимум функции (3) с ρ-функциями Хьюбера (4) и Тьюки (5) для вычисления М-оценки найдем методом Нелдера — Мида [9]. Оценку наименьших квадратов будем вычислять по формуле (2), а оценку наименьших модулей — при помощи итерационного взвешенного метода наименьших квадратов, при котором $L_{LAD}(a)$ представляется в виде

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

с весами $w_t(a_1, a_2) = |X_t - a_1 X_{t-1}^{+} - a_2 X_{t-1}^{-}|^{-1}$. Точка минимума $\tilde{a}$ функции $L_{LAD}(a)$ является пределом последовательности $a^{(k)} = \left(a_1^{(k)},\ a_2^{(k)}\right)$, k-й член которой находится при помощи минимизации функции

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

В качестве начального приближения $a^{(0)} = (a_1^{(0)}, a_2^{(0)})$ будем использовать оценку наименьших квадратов.

Описанный эксперимент повторялся $N = 100000$ раз. Моделирование показало, что М-оценки можно считать асимптотически несмещенными. Поэтому дисперсия всех оценок параметра a_j оценивалась величиной $\frac{1}{N}\sum\limits_{i=1}^{N}(a_{ij}-a_j)^2$, где a_{ij} — оценка параметра a_j в i-м эксперименте, $i = 1, \ldots, N$, $j = 1, 2$.

В табл. 1 приведены оценки дисперсий оценок наименьших квадратов, оценок наименьших модулей, М-оценок с ρ-функцией Хьюбера и ρ-функцией Тьюки для перечисленных выше вероятностных распределений ε_t.

Отметим, что дисперсия М-оценок зависит от параметра k в формулах (4), (5) а значение k, при котором дисперсия минимальна, зависит от f. Величины дисперсий в табл. 1 соответствуют компромиссным значениям k, равным 2 и 4,5 для М-оценок с ρ-функцией Хьюбера и ρ-функцией Тьюки соответственно.

Видно, что оценка наименьших квадратов является лучшей только при нормальном распределении ε_t. Из распределений, указанных в табл. 1, наибольшее сходство с нормальным имеют логистическое распределение и распределение Стьюдента с большим числом степеней свободы, плотности которых являются гладкими и достаточно быстро убывают на бесконечности. И уже в этом случае М-оценки составляют конкуренцию оценке наименьших квадратов, причем М-оценка с ρ-функцией Хьюбера превосходит оценку наименьших квадратов в случае обоих распределений.

Для распределения Лапласа оценка наименьших модулей совпадает с оценкой максимального правдоподобия и поэтому является наилучшей среди всех оценок.

Распределение Коши имеют бесконечную дисперсию, что приводит к несравнимо низкой эффективности оценки наименьших квадратов по отношению к остальным оценкам.

Семейство распределений Стьюдента с различным числом m степеней свободы служит моделью гладких распределений, плотность которых на бесконечности убывает с различными скоростями — от квадратичной (у распределения Стьюдента с одной степенью свободы, совпадающим с распределением Коши) до практически экспоненциальной при достаточно больших m. Из табл. 1 видно, что с уменьшением числа степеней свободы оценка наименьших квадратов сначала проигрывает только М-оценке с ρ-функцией Хьюбера, потом обеим М-оценкам, а затем и всем оценкам.

Данные табл. 1 также свидетельствуют в пользу того, что, если распределение ε_t практически не отличается от нормального, то целесообразно использовать М-оценку с ρ-функцией Хьюбера, при более сильном отклонении распределения от нормального — применять М-оценку с ρ-функцией Тьюки, а при значительном отклонении — оценку наименьших модулей. Отметим, что при еще большем отклонении распределения от нормального следует использовать знаковую оценку [11].

Предположим теперь, что плотность распределения вероятности обновляющего процесса ε_t имеет загрязненное нормальное распределение (6), и исследуем зависимость точности оценок от параметров δ и τ этого распределения. В качестве сравнительной характеристики точности двух оценок будем использовать относительную эффективность оценок, которая определяется как обратное отношение их дисперсий. Таким образом, например, относительная эффективность М-оценки параметра a_j по отношению к оценке наименьших квадратов этого же параметра будет оцениваться по формуле

$$e_j = \frac{\sum\limits_{i=1}^{N}(\tilde{a}_{ij} - a_j)^2}{\sum\limits_{i=1}^{N}(\hat{a}_{ij} - a_j)^2}, \qquad j = 1,\ 2,$$

где $\tilde{a}_{ij}$ и $\hat{a}_{ij}$ — оценка наименьших квадратов и М-оценка параметра a_j соответственно; $i = 1, \ldots, N$ — номер эксперимента. При этом неравенство $e_j > 1$ будет означать, что М-оценка точнее оценки наименьших квадратов и что для достижения такой же точности, как у М-оценки, оценке наименьших квадратов потребуется в e_j раз больше наблюдений. Неравенство $e_j < 1$ интерпретируется аналогичным образом.

Результаты эксперимента приведены на рис. 1, 2.

На рис. 1 показана зависимость относительной эффективности (АОЭ) от δ при $\tau = 3$, на рис. 2 — зависимость относительной эффективности от τ при $\delta = 0.15$. Красной и фиолетовой линиями обозначена эффективность М-оценки с ρ-функцией Хьюбера по отношению к оценке наименьших квадратов и оценке наименьших модулей соответственно, зеленой и черной — эффективность М-оценки с ρ-функцией Тьюки по отношению к тем же оценкам. Для сравнения синей линией отмечена эффективность оценки наименьших модулей по отношению к оценке наименьших квадратов. Сплошная линия везде относится к оцениванию коэффициента $a_1 = -0.3$, пунктирная — к оцениванию коэффициента $a_2 = 0.5$.

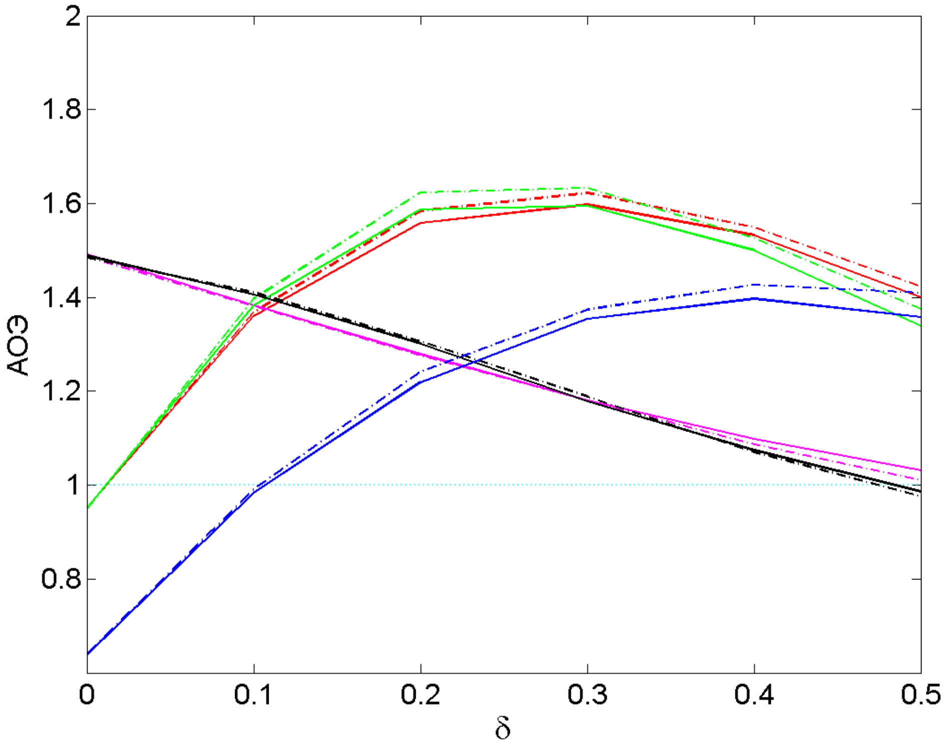

Рис. 1. Зависимость относительной эффективности от δ

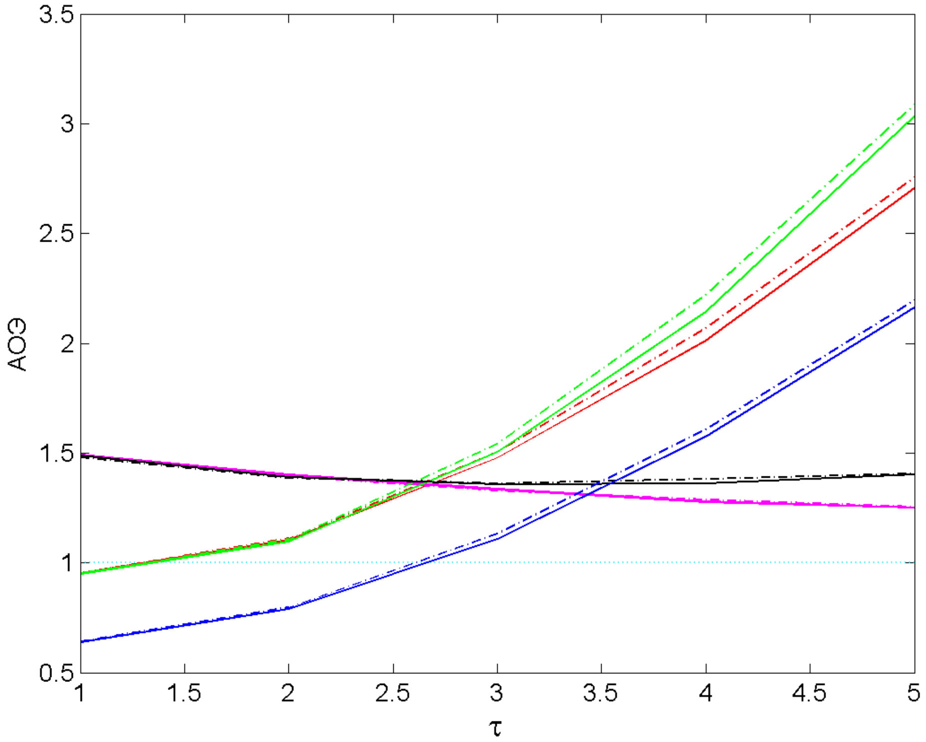

Рис. 2. Зависимость относительной эффективности от τ

Для удобства анализа графиков голубой горизонтальной линией отмечена относительная эффективность, равная единице.

Постоянная k ρ-функции Хьюбера во всех экспериментах равнялась 2, поскольку опытным путем было установлено, что если f задается равенством (6), а δ и τ принимают типичные на практике значения $\delta \in (0,\ 0{,}3)$, $\tau \in (1,\ 5)$, то выборочная дисперсия М-оценки минимальна при $k \in (1{,}5,\ 2{,}5)$.

Из рисунков видно, что при отсутствии 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.

Заключение

В работе изучены свойства М-оценки, оценки наименьших квадратов и оценки наименьших модулей параметров пороговой авторегрессионной модели с одним порогом, который предполагается известным и равным нулю. Методом компьютерного моделирования проведена оценка дисперсий всех указанных оценок для типичных вероятностных распределений обновляющей последовательности авторегрессионного уравнения, в частности, для загрязненного нормального распределения.

Оказалось, что оценка наименьших квадратов является лучшей только при нормальном распределении ε_t. Уже для логистического распределения и распределения Стьюдента с большим числом степеней свободы, плотности которых являются гладкими и достаточно быстро убывают на бесконечности, М-оценки составляют конкуренцию оценке наименьших квадратов, причем М-оценка с ρ-функцией Хьюбера превосходит оценку наименьших квадратов в случае обоих распределений.

Для распределения Лапласа оценка наименьших квадратов является наихудшей, а оценка наименьших модулей наилучшей среди всех оценок.

Для распределения Коши, которые имеют бесконечную дисперсию, оценка наименьших квадратов имеет несравнимо низкую эффективность по отношению к остальным оценкам.

С уменьшением числа степеней свободы у распределения Стьюдента оценка наименьших квадратов сначала проигрывает только М-оценке с ρ-функцией Хьюбера, потом обеим М-оценкам, а затем и оценке наименьших модулей.

Если обновляющий процесс имеет загрязненное нормальное распределение, то М-оценка уступает, причем не намного, оценке наименьших квадратов только при практически полном отсутствии загрязнений.

С ростом доли и уровня загрязнения относительная эффективность М-оценки по отношению к оценке наименьших квадратов увеличивается, становясь больше единицы для типичного на практике загрязнения.

Таким образом, при отсутствии априорной информации о вероятностном распределении обновляющего процесса порогового авторегрессионного уравнения для оценивания параметров этого уравнения целесообразно использовать М-оценки, предпочтя их, в частности, оценкам наименьших квадратов и наименьших модулей.

Список литературы

1. Hansen B.E. Threshold autoregression in economics // Statistics and Its Interface. 2011. Vol. 4, no. 2. Pp. 123–127. DOI: 10.4310/SII.2011.v4.n2.a4
2. Tong H. Non-linear time series: A dynamical system approach. Oxf.: Clarendon Press; N.Y.: Oxf. Univ. Press, 1990. 564 p.
3. Douc R., Moulines E., Stoffer D. Nonlinear time series: Theory, methods and applications with R examples. Boca Raton: CRC Press, 2014. 531 p.
4. Li D., Ling Sh. On the least squares estimation of multiple-regime threshold autoregressive models // J. of Econometrics. 2012. Vol. 167, no. 1. Pp. 240–253. DOI: 10.1016/j.jeconom.2011.11.006
5. Wang L., Wang J. The limiting behavior of least absolute deviation estimators for threshold autoregressive models // J. of Multivariate Analysis. 2004. Vol. 89. No. 2. Pp. 243-260. DOI: 10.1016/j.jmva.2004.02.006
6. 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.

7. Huber P., Ronchetti E.M. Robust statistics. 2nd ed. Hoboken: Wiley, 2009. 370 p. DOI: 10.1002/9780470434697
8. Wilcox R.R. Introduction to robust estimation and hypothesis testing. 3rd ed. Amst.; Boston: Academic Press, 2012. 690 p.
9. Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. Numerical Recipes: The Art of Scientific Computing. 3rd ed. Camb.; N.Y.: Camb. Univ. Press, 2007. 1256 p.
10. Bissantz N., Dumbgen L., Munk A., Stratmann B. Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces // SIAM J. of Optimization. 2009. Vol. 19, no. 4. Pp. 1828–1845. DOI: 10.1137/050639132
11. Горяинов В.Б., Горяинова Е.Р. Асимптотические свойства знаковой оценки коэффициентов авторегрессионного поля // Автоматика и телемеханика. 2015. № 3. С. 62–78.

Mathematics & Mathematical Modelling
Electronic journal of the Bauman MSTU
http://mathm.elpub.ru

Mathematics and Mathematical Modelling of the Bauman MSTU, 2017, no. 3, pp. 91–104.
DOI: **10.24108/mathm.0317.0000072**
Received: 07.06.2017

Comparison of Classical and Robust Estimates of Threshold Auto-regression Parameters

Goryainov V. B.[1,*]

[*] vb-goryainov@bmstu.ru

[1] Bauman Moscow State Technical University, Russia

Keywords: threshold autoregression, M-estimate, least square estimate, least absolute deviation estimate, contaminated normal distribution

The study object is the first-order threshold auto-regression model with a single zero-located threshold. The model describes a stochastic temporal series with discrete time by means of a piecewise linear equation consisting of two linear classical first-order autoregressive equations. One of these equations is used to calculate a running value of the temporal series. A control variable that determines the choice between these two equations is the sign of the previous value of the same series.

The first-order threshold autoregressive model with a single threshold depends on two real parameters that coincide with the coefficients of the piecewise linear threshold equation. These parameters are assumed to be unknown. The paper studies an estimate of the least squares, an estimate the least modules, and the M-estimates of these parameters. The aim of the paper is a comparative study of the accuracy of these estimates for the main probabilistic distributions of the updating process of the threshold autoregressive equation. These probability distributions were normal, contaminated normal, logistic, double-exponential distributions, a Student's distribution with different number of degrees of freedom, and a Cauchy distribution.

As a measure of the accuracy of each estimate, was chosen its variance to measure the scattering of the estimate around the estimated parameter. An estimate with smaller variance made from the two estimates was considered to be the best. The variance was estimated by computer simulation. To estimate the smallest modules an iterative weighted least-squares method was used and the M-estimates were done by the method of a deformable polyhedron (the Nelder-Mead method). To calculate the least squares estimate, an explicit analytic expression was used.

It turned out that the estimation of least squares is best only with the normal distribution of the updating process. For the logistic distribution and the Student's distribution with the large number of degrees of freedom, the M-estimate with the Huber rho-function exceeds the least squares estimate in the case of both distributions.

For the Laplace distribution, the least squares estimate is the worst, and the least modulus estimate is the best among all estimates.

For the Cauchy distribution, the least-squares estimate has incomparably low efficiency with respect to the remaining estimates.

With decreasing number of degrees of freedom in the Student's distribution, the least squares estimate at first loses only the M-estimate with the Huber rho-function, then both M-estimates, and then the least moduli estimate.

If the updating process has a contaminated normal distribution, then the M-estimate is a little bit lower than the least squares estimate only in case there is absolutely no contaminants.

With increasing contamination share and level, relative effectiveness of the M-estimate with respect to the estimation of least squares grows, becoming above unit for typical contamination in practice.

References

1. Hansen B.E. Threshold autoregression in economics. *Statistics and Its Interface*, 2011, vol. 4, no. 2, pp. 123-127. DOI: 10.4310/SII.2011.v4.n2.a4

2. Tong H. *Non-linear Time Series: A Dynamical System Approach*. Oxf.: Clarendon Press; N.Y.: Oxf. Univ. Press, 1990. 564 p.

3. Douc R., Moulines E., Stoffer D. *Nonlinear Time Series: Theory, Methods and Applications with R Examples*. Boca Raton: CRC Press, 2014. 531 p.

4. Li D., Ling Sh. On the least squares estimation of multiple-regime threshold autoregressive models. *J. of Econometrics*, 2012, vol. 167, no. 1, pp. 240–253. DOI: 10.1016/j.jeconom.2011.11.006

5. Wang L., Wang J. The limiting behavior of least absolute deviation estimators for threshold autoregressive models. *J. of Multivariate Anal.*, 2004, vol. 89, no. 2, pp. 243–260. DOI: 10.1016/j.jmva.2004.02.006

6. 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, no. 9, pp. 1579–1588 DOI: 10.1134/S000511791609006X (Russian version of journal: *Avtomatika i telemehanika*, no. 9, 2016, pp. 84–95).

7. Huber P., Ronchetti E.M. *Robust statistics*. 2nd ed. Hoboken: Wiley, 2009. 370 p. DOI: 10.1002/9780470434697

8. Wilcox R.R. *Introduction to Robust Estimation and Hypothesis Testing*. 3rd ed. Amst.; Boston: Academic Press, 2012. 690 p.

9. Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. Camb.; N.Y.: Camb. Univ. Press, 2007. 1256 p.

10. Bissantz N., Dumbgen L., Munk A., Stratmann B. Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces. *SIAM J. of Optimization*, 2009, vol. 19, no. 4, pp. 1828–1845. DOI: 10.1137/050639132

11. Goryainov V.B., Goryainova E.R. Asymptotic properties of the sign estimate of autoregression field coefficients. *Automation and Remote Control*, 2015, vol. 76, no. 3, pp. 419–432. DOI: 10.1134/S0005117915030066 (English version of journal: *Avtomatika i telemehanika*, 2015, no. 3, pp. 62–78).