МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет

# Всероссийская молодежная научная конференция «Все грани математики и механики»

(25–29 апреля 2016 г.)

## Сборник статей

Под редакцией
д-ра физ.-мат. наук, профессора А.В. Старченко

Томск
Издательский Дом Томского государственного университета
2016

# Сравнение различных подходов к оцениванию параметра модели устойчивой авторегрессии первого порядка с дискретным временем

**Иващенко А. О., Емельянова Т. В.**

Томский государственный университет
e-mail: anutka0694@gmail.com

**Аннотация**

Настоящая работа посвящена оцениванию параметра модели устойчивой авторегрессии первого порядка с дискретным временем. В работе проводится сравнительный анализ процедур оценивания параметра модели устойчивой авторегрессии первого порядка AR(1): последовательная процедура оценивания, усеченная последовательная процедура оценивания. Проведено имитационное моделирование, с помощью которого установлено, что обе процедуры позволяют получить оценки с заданной среднеквадратической точностью.

**Ключевые слова:** параметрическое оценивание; модель авторегрессии первого порядка (AR(1)); последовательный подход к оцениванию; усеченная процедура последовательного оценивания.

В задачах обработки временных рядов широко используются авторегрессионные модели, описывающие стационарные случайные процессы. Как правило, параметры таких моделей неизвестны, поэтому требуется идентифицировать параметры модели перед ее использованием. В работе рассматривается задача оценивания параметра. Проводится исследование качества оценки модели авторегрессии первого порядка с дискретным временем.

Для получения последовательных оценок с произвольной точностью необходимо иметь выборку неограниченного размера. Однако на практике время наблюдения системы, как правило, фиксировано. Последовательное оценивание и усеченное последовательное оценивание являются подходами для нахождения оценок с гарантированной точностью.

Целью исследования является сравнение последовательной и усеченной последовательной процедуры оценивания параметра модели устойчивой авторегрессии первого порядка.

Рассматривается задача оценивания параметра процесса $x_t$, заданного стохастическим разностным уравнением

$$X_i = \lambda X_{i-1} + \sigma\varepsilon_i, i = 1, 2, ..., \quad (1)$$

где $X_0 = 0$, $\varepsilon_i$ – независимые одинаково распределенные случайные величины. $E\varepsilon_i = 0$, $Var\varepsilon_i = \sigma^2 < +\infty$.

Чаще всего для оценивания параметра модели AR(1) применяют метод наименьших квадратов (МНК) [1]. Можно рассмотреть процедуру оценивания по методу наименьших квадратов, в которой число наблюдений не фиксируется заранее. Процесс наблюдений останавливается в некоторый момент времени.

При последовательном оценивании число наблюдений заранее неизвестно, оно определяется в ходе наблюдения процесса [3]. Построим последовательную оценку, представляющую собой оценку по методу наименьших квадратов, вычисленную в момент остановки, когда значения дисперсии и параметра $\lambda$ неизвестны. При этом пользуются правилом остановки, построенным по наблюдаемому процессу [4]

$$t_A = inf(n \geq m_A : n \geq A^{1/2}\hat{\sigma}_n), \quad (2)$$

где $\hat{\sigma}_n = n^{-1}\sum_{i=1}^{n}(X_i - \hat{\lambda}_n X_{i-1})^2$ – оценка дисперсии по выборочным данным и $m_A$ - заранее заданный объем выборки. Таким образом, последовательной оценкой параметра модели авторегрессии первого порядка является оценка по МНК [5]

$$\hat{\lambda}_{t_A} = \sum_{i=1}^{t_A} \frac{X_i X_{i-1}}{X_{i-1}^2}. \quad (3)$$

Оценка МНК нелинейная, что затрудняет исследования, хорошо исследованы ее асимптотические свойства: полученная оценка является асимптотически несмещенной и асимптотически нормальной [1].

При решении практических задач может оказаться, что при заданном объеме выборки момент остановки не достигается. В этом случае используют усеченную последовательную процедуру.

Определим в общей процедуре оценивания весовые функции [2]

$$w(n) = \begin{cases} 0, & 1 \leq n \leq m; \\ \chi[\sigma_m^2 > (logm)^{-1}], & m < n \leq N. \end{cases} \quad (4)$$

Длительность процедуры определяется правилом

$$\tau_n = \begin{cases} inf\{k \in [1, N] : \sum_{n=m+1}^{k} x_{n-1}^2 \geq H_N\}, & \sum_{n=m+1}^{N} x_{n-1}^2 \geq H_N; \\ N, & \sum_{n=m+1}^{N} x_{n-1}^2 < H_N; \end{cases} \quad (5)$$

где $H_N = h\sigma_m^2 N$ – пороговое значение.

Весовая функция имеет вид

$$\beta_n = \begin{cases} 1, & 1 \le n < \tau_N; \\ 1, & n = \tau_N, \sum_{n=m+1}^{N} x_{n-1}^2 < H_N; \\ \alpha_N, & n = \tau_N, \sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N; \end{cases} \tag{6}$$

где $\alpha_N = (H_N - \sum_{n=m+1}^{\tau_N - 1} x_{n-1}^2)/x_{\tau_N - 1}^2$.

Тогда усеченная последовательная оценка модели (1) имеет вид

$$\tilde{\lambda}_N = \frac{1}{H_N} \sum_{n=m+1}^{\tau_N} \beta_n x_n x_{n-1} \chi[\sum_{n=m+1}^{N} x_{n-1}^2 \ge H_N, \sigma_m^2 > > (log m)^{-1}]. \tag{7}$$

Приведем формулировку теоремы о свойствах усеченной последовательной оценки [2].

**Теорема 1.** *Пусть модель (1) устойчива, тогда усеченная последовательная оценка (7) удовлетворяет следующему неравенству*

$$E_\mu(\tilde{\lambda}_N - \lambda)^2 \ge \frac{1}{Nh} + \varepsilon_N. \tag{8}$$

Проведено имитационное моделирование, в ходе которого установлено, что обе процедуры позволяют получить оценки с заданной среднеквадратической точностью. Моделирование проведено для случая гауссовских ошибок (ошибок из стандартного нормального распределения). Для определенности значение параметра $\lambda$ приняли равным 0.1. $X_0 = 0$, $A = 400, 1000, 4000$, $m_A = 5$.

Приведем результаты численного моделирования для последовательной процедуры оценивания. Здесь $\hat{\lambda}_n$ – оценка параметра модели, вычисленной по МНК при фиксированном объеме выборки; $\hat{\sigma}_n$ – оценка дисперсии при фиксированном объеме выборки; $\hat{\lambda}_{t_A}$ – последовательная оценка параметра, вычисленная в момент остановки $t_A$; $\hat{\sigma}_{t_A}$ – оценка дисперсии, вычисленная в момент остановки $t_A$; $S_0^2 = \frac{1}{t_A} \sum_{i=1}^{t_A} (\hat{\lambda}_{t_A} - \lambda)^2$ – среднеквадратическое отклонение последовательной оценки от истинного значения параметра.

Таблица 1

Результаты численного моделирования последовательной процедуры оценивания

| | | | |
|---|---|---|---|
| **$A = 400$** | | | |
| $t_A$ | | 23 | |
| $\hat{\lambda}_n$ | 0,08 | $S_0^2$ | 0,18 |
| $\hat{\sigma}_n$ | 1,23 | | |
| $\hat{\lambda}_{t_A}$ | 0,15 | | |
| $\hat{\sigma}_{t_A}$ | 0,09 | | |
| **$A = 1\,000$** | | | |
| $t_A$ | | 36 | |
| $\hat{\lambda}_n$ | 0,08 | $S_0^2$ | 0,11 |
| $\hat{\sigma}_n$ | 1,23 | | |
| $\hat{\lambda}_{t_A}$ | 0,16 | | |
| $\hat{\sigma}_{t_A}$ | 0,06 | | |
| **$A = 4\,000$** | | | |
| $t_A$ | | 71 | |
| $\hat{\lambda}_n$ | 0,08 | $S_0^2$ | 0,06 |
| $\hat{\sigma}_n$ | 1,23 | | |
| $\hat{\lambda}_{t_A}$ | 0,1 | | |
| $\hat{\sigma}_{t_A}$ | 0,02 | | |
| **$A = 10\,000$** | | | |
| $t_A$ | | 111 | |
| $\hat{\lambda}_n$ | 0,08 | $S_0^2$ | 0,03 |
| $\hat{\sigma}_n$ | 1,23 | | |
| $\hat{\lambda}_{t_A}$ | 0,11 | | |
| $\hat{\sigma}_{t_A}$ | 0,01 | | |

По имеющимся данным построен график, иллюстрирующий изменения среднеквадратического отклонения.

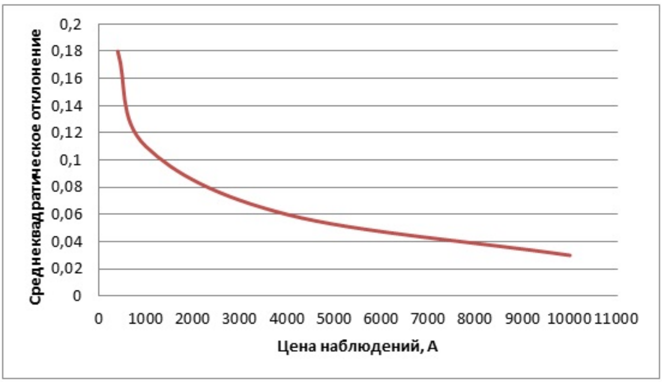


Рис. 8. Изменение среднеквадратического отклонения

Таким образом, с увеличением цены наблюдений точность оценивания становится выше, а значит и оценки становятся более эффективными.

Приведем результаты численного моделирования для усеченной процедуры оценивания. Здесь $\hat{\lambda}_N$ – оценка параметра модели, вычисленной по МНК при фиксированном объеме выборки; $\sigma_m^2$ – оценка дисперсии; $\tilde{\lambda}_{\tau_N}$ – усеченная последовательная оценка параметра, вычисленная в момент остановки $\tau_N$; $\hat{\sigma}_{\tau_N}$ – оценка дисперсии, вычисленная в момент остановки $\tau_N$. $S_0'^2 = \frac{1}{\tau_N}\sum_{i=1}^{\tau_N}(\tilde{\lambda}_{\tau_N} - \lambda)^2$ – среднеквадратическое отклонение усеченной последовательной оценки от истинного значения параметра.

Таблица 2

Результаты численного моделирования усеченной последовательной процедуры оценивания

| | | | |
|---|---|---|---|
| $N$=500, $m$=167 | | | |
| $\tau_N$ | | 301 | |
| $\hat{\lambda}_N$ | 0,08 | $S_0'^2$ | 0,009 |
| $\hat{\sigma}_m^2$ | 1,59 | | |
| $\tilde{\lambda}_{\tau_N}$ | 0,06 | | |
| $\hat{\sigma}_{\tau_N}^2$ | 0,53 | | |
| $N = 1000, m = 333$ | | | |
| $\tau_N$ | | 577 | |
| $\hat{\lambda}_N$ | 0,14 | $S_0'^2$ | 0,008 |
| $\hat{\sigma}_m^2$ | 1,36 | | |
| $\tilde{\lambda}_{\tau_N}$ | 0,09 | | |
| $\hat{\sigma}_{\tau_N}^2$ | 0,48 | | |
| $N = 2000, m = 667$ | | | |
| $\tau_N$ | | 1041 | |
| $\hat{\lambda}_N$ | 0,15 | $S_0'^2$ | 0,007 |
| $\hat{\sigma}_m^2$ | 1,25 | | |
| $\tilde{\lambda}_{\tau_N}$ | 0,12 | | |
| $\hat{\sigma}_{\tau_N}^2$ | 0,47 | | |

По имеющимся данным построен график, иллюстрирующий изменения среднеквадратического отклонения.

Уменьшение среднеквадратического отклонения при увеличении объема выборки доказывает, что усеченная последовательная процедура оценивания является эффективной.

Таким образом, как последовательная процедура, так и усеченная процедура оценивания, дают надежные оценки с заданной среднеквадратической точностью. В случае неограниченного объема выборки удобно пользоваться последовательной процедурой

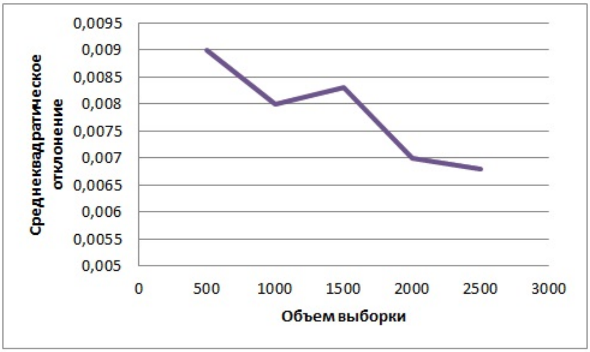


Рис. 9. Изменение среднеквадратического отклонения

оценивания. Если при заданном объеме выборки момент остановки не достигается, оценивание следует проводить с помощью усеченной последовательной процедуры.

**Литература**

1. Андерсон Т. Статистический анализ временных рядов: пер. с англ. / Т. Андерсон; под ред. Ю. К. Беляева. М.: Мир, 1976. 755 с.

2. В.А. Васильев, Т.В. Догадова Гарантированное оценивание параметров стохастической линейной регрессии по выборке фиксированного размера. Вестник Томского государственного университета. 2014. №1(26). С.39-52.

3. Вальд А. Последовательный анализ: пер. с англ. / А. Вальд; под ред. Б. А. Севастьянова. М.: Государственное изд. Физико-математической лит-ры, 1960. 329 с.

4. Sriram T. Sequential Estimation for Time Series Models / T.N.Sriram, R.Iaci // Sequential Analysis: Design Methods and Applications. 2014. V. 33. P. 136-157.

5. Sriram T. Sequential Estimation of the autoregressive parameter in a first order autoregressive process / T.N.Sriram// Sequential Analysis: Design Methods and Applications. 1988. V. 7(1). P. 53-74.