DOI: https://doi.org/10.5281/zenodo.18008794

НЕСТАНДАРТНЫЕ ОЦЕНКИ В ОДНОМЕРНЫХ И ПРОСТРАНСТВЕННЫХ AR-МОДЕЛЯХ

Мирзаев Тохиржон Салохетдинович
Доцент кафедры языков, точных и общественных наук, ф.-м.ф.н.
Институт технологий, менеджмента и коммуникаций
tsmirzayev@gmail.com
+998903382909

Кучкарова Сарвиноз Атамуратовна
Доцент кафедры языков, точных и общественных наук, PhD
Институт технологий, менеджмента и коммуникаций
quchqarova.sarvinoz@tmci.uz
+998931111199

Аннотация. *В докладе предлагаются оценки параметров авторегрессии, отличные от оценок наименьших квадратов. Оценки наименьших квадратов в неустойчивых (критических) случаях, т.е. когда корни характеристического уравнения лежат на единичной окружности, имеют, как правило, сложное предельное распределение. Предлагаемые же нестандартные оценки в большинстве критических случаев имеют более простое предельное распределение.*

Ключевые слова. *Обыкновенная авторегрессии первого порядка, пространственная авторегрессии, авторегрессионной схемой $p-$го порядка, оценка параметров, оценка методом наименьших квадратов, нестандартный подход, стандартный винеровский процесс, предельные распределения.*

ВВЕДЕНИЕ

В докладе предлагаются нестандартные подходы к построению оценок в различных моделях авторегрессии. Уравнения оценивания строятся с использованием рекуррентных связей, задающих исходный процесс, и в каждом из рассматриваемых случаев отдельно. Схожим способом можно строить и классические оценки наименьших квадратов, т.е. не опираясь на метод минимизации соответствующей суммы квадратов по исходным параметрам.

Необходимость построения такого типа оценок связана с тем, что оценки наименьших квадратов в критических (неустойчивых) случаях имеют сложное предельное распределение, выражающееся, как правило, через функционалы от стандартного винеровского процесса.

С точки зрения приложений именно критические случаи представляют значительный интерес. В связи с этим, например, в случае обыкновенной авторегрессии первого порядка, т.е. когда $X_t = \alpha X_{t-1} + \varepsilon_t$, было уделено большое внимание вопросу табулирования сложного предельного распределения при $\alpha = 1$. Если модели обыкновенной авторегрессии представляют большой интерес для экономистов [1-4], [13-17] и модели пространственной авторегрессии, т.е. когда $X_{t,s} = \alpha X_{t-1,s} + \beta X_{t,s-1} + \varepsilon_{t,s}$ представляют интерес в геологии, сельском хозяйстве, а также при обработке спутниковых изображений поверхности Земли с точки зрения прогноза аномальных явлений [5-8]. Следует также отметить, что модели пространственной авторегрессии начали интенсивно исследоваться лишь в последние десятилетия. Более подробную историю вопроса, в сравнении с приведенней здесь, по каждой из рассматриваемых моделей можно почерпнуть в соответствующих ссылках на литературные источники.

ФОРМУЛИРОВКА ПРОБЛЕМЫ

Рассматриваем следующие модели авторегрессии.

Определение [10]. Авторегрессионной схемой $p-$го порядка (AR(p)) называется соотношение вида

$$X_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t, \quad t = 1,2,... \tag{1}$$

где α_i, $i = 0,1,...,p-$постоянные, а $\varepsilon_t -$случайные величины называемые авторегрессионными шумами или просто шумами.

В случае $\alpha_0 = X_0 = 0$ модель авторегрессии первого порядка примет вид

$$X_t = \alpha X_{t-1} + \varepsilon_t, \quad t = 1,2,..., \tag{2}$$

где $\varepsilon_1, \varepsilon_2, ... -$независимые одинаково распределенные случайные величины (н.о.р.с.в.) с $E\varepsilon_1 = 0$, $E\varepsilon_1^2 = 1$, $X_0 -$начальное состояние.

Известная оценка параметра α, полученная по n наблюдениям методом наименьших квадратов имеет вид

$$\hat{\alpha}_n = \sum_{t=1}^{n} X_{t-1} X_t \Big/ \sum_{t=1}^{n} X_{t-1}^2 . \tag{3}$$

В работе [11] эта оценка приводится как коэффициент сериальной корреляции.

Если $\{\varepsilon_i\}$ нормально распределены, то оценка (3) совпадает с оценкой максимального правдоподобия.

Сложная структура оценки $\hat{\alpha}_n$ затрудняет нахождение предельного распределения даже в случае нормально распределенных шумов $\{\varepsilon_i\}$.Известно [12[, [15], что $\sqrt{n}\left(\hat{\alpha}_n - \alpha\right)$ имеет нормальное предельное распределение при $\alpha \in (-1, 1)$, а для $|\alpha| \ge 1$ предельное распределение является сложным и, более того, при $|\alpha| > 1$ оно становится зависящим от распределения шумов $\{\varepsilon_i\}$. Например, при $\alpha = 1$ имеет место утверждение [14-15] при $n \to \infty$

$$n\left(\widehat{\alpha}_n-1\right)\Rightarrow\frac{1}{2}\left(w^2(1)-1\right)\Big/\int_0^1 w^2(t)dt, \qquad (4)$$

где $w(t)$ – стандартный винеровский процесс, а символ $\Rightarrow$ означает слабую сходимость соответствующих распределений.

В работе [16] была предложена другая, более простая по структуре, оценка параметра α. Простое суммирование (2) по t от k до n приводит к следующей оценке

$$\alpha_{n,k}^{*}=\frac{X_k+\ldots+X_n}{X_{k-1}+\ldots+X_{n-1}}=\alpha+\frac{\varepsilon_k+\ldots+\varepsilon_n}{X_k+\ldots+X_n}. \qquad (5)$$

Показано, что при $n\to\infty$

$$P\left(n\beta(c)\left(\alpha_{n,k}^{*}-1\right)<x\right)\Rightarrow\frac{1}{2}+\frac{1}{\pi}arctg\left[(x-\rho_c)\Big/\sqrt{1-\rho_c^2}\right],$$

где $c=\lim\limits_{n\to\infty}\dfrac{k}{n}<1,\ \ \beta(c)=\sqrt{(1-c)(1+2c)/3},\ \ \rho_c=\dfrac{1}{2}\left(\dfrac{3(1-c)}{1+2c}\right)^{1/2}.$

Если же $c=1$, то

$$P\left(\sqrt{k(n-k)}\left(\alpha_{n,k}^{*}-1\right)<x\right)\Rightarrow\frac{1}{\pi}\int_{-\infty}^{x}\frac{1}{1+u^2}du.$$

Заметим, что в случае $\alpha=-1$ какие-либо результаты отсутствуют. Принцип инвариантности, использованный при изучении оценки (3) в работе [15], при $\alpha=-1$ результата не даёт, а что касается оценки (5), то она в этом случае является несостоятельной, что нетрудно показать.

Модели пространственной авторегрессии начали интенсивно исследоваться сравнительно недавно и еще не вошли в монографическую и учебную литературу. Некоторый обзор результатов мы дадим следуя работе [9].

Анализ пространственных моделей представляет интерес во многих областях, таких как география, геология, биология и сельское хозяйство. Дискуссию по этому поводу см. в работе [18]. Эти авторы рассматривали случай, так называемой, односторонней модели AR(p), имеющей форму

$$X_{k,l}=\sum_{i=0}^{p_1}\sum_{j=0}^{p_2}\alpha_{i,j}X_{k-i,l-j}+\varepsilon_{k,l}, \quad \alpha_{0,0}=0.$$

В работе [9] рассматривается особый случай этой модели, а именно, когда $p_1=p_2=1$, $\alpha_{0,1}=\alpha_{1,0}=:\alpha$, $\alpha_{1,1}=0$, и получены специфические результаты об асимптотическом поведении оценки α в неустойчивом случае. В литературе очень немного результатов этого типа для пространственных моделей. С общей точки зрения, желательно иметь дело с моделями, где $X_{k,l}$ является линейной комбинацией всех соседей по решетки. В частности, было бы интересно рассмотреть обобщение модели S. Baran, когда $X_{k-1,l}$ и $X_{k,l-1}$ имеют различные веса α и β. Но даже в данной модели с $\alpha=\beta$ возникают сложные математические проблемы с достаточно нестандартными результатами [9].

В данной работе предлагается оценка более простой структуры и использующая только часть наблюдений, расположенных вдоль диагонали в прямоугольнике $R_{m,n}$. Такой подход с существенным сокращением числа наблюдений является актуальным в задачах геологии и некоторых других областях применения моделей пространственной авторегрессии. Более простая структура оценки позволяет также снизить моментные ограничения на шумы.

РЕШЕНИЕ ПРОБЛЕМЫ И РЕЗУЛЬТАТЫ

Модель обыкновенной одномерной авторегрессии первого порядка. Близкий к критическому случай

В настоящем работе рассматривается модель авторегрессии первого порядка. Процесс (2) является устойчивым в случае, когда $|\alpha|<1$, неустойчивым при $|\alpha|=1$ и относится к взрывному типу при $|\alpha|>1$.

Заметим, что в случае $\alpha(n)\to 1$ при $n\to\infty$ какие-либо результаты отсутствуют. В настоящей работе исследуется близкий к критическому случай, когда $\alpha(n)\to 1$ при $n\to\infty$ для частного вида оценки (5) $(k=1)$.

Построение оценки и её предельное поведение

Уравнение оценивания составим путем суммирования соотношения (2) по t от 1 до n

$$\sum_{t=1}^{n} X_t = \alpha \sum_{t=1}^{n} X_{t-1} + \sum_{t=1}^{n} \varepsilon_t$$

или в сокращенном виде

$$Y_n = \alpha Y_{n-1} + E_n. \tag{6}$$

Решая уравнение (6) без учёта E_n, получим оценку

$$\alpha_{n,1}^{*} = \frac{Y_n}{Y_{n-1}}.$$

Теперь из (6) найдем отклонение

$$\alpha_{n,1}^{*} - \alpha = \frac{E_n}{Y_{n-1}}. \tag{7}$$

Теперь заметим, из (2) имеем

$$X_{t-1} = \frac{1}{\alpha}\left(X_t - \varepsilon_t\right)$$

и следовательно

$$Y_{n-1} = \frac{1}{\alpha}\left(Y_n - E_n\right). \tag{8}$$

Теорема 1. Если $\varepsilon_1, \varepsilon_2, \ldots$ – н. о. р. с. в. с $E(\varepsilon_1)=0$, $E\left(\varepsilon_1^2\right)=1$ и $X_0=0$,

1. При $n\to\infty$, когда $\alpha = 1-\dfrac{c}{n}$, $|c|<\infty$

$$P\left(n\sigma_c\left(\alpha_{n,1}^*-\alpha\right)<x\right)\Rightarrow\frac{\sqrt{1-\delta_c^2}}{\pi}\int_{-\infty}^{x}\frac{du}{u^2+2\delta_c u+1}=\frac{1}{2}+\frac{1}{\pi}arctg\frac{x+\delta_c}{\sqrt{1-\delta_c^2}},$$

где $\sigma_c^2=\frac{1}{2|c|^3}\left[2|c|-3+4e^{-|c|}-e^{-2|c|}\right],\qquad \delta_c=\frac{1}{c^2\sigma_c}\left(e^{-|c|}+|c|-1\right).$

2. Если $\alpha(n)=1-\frac{c}{\gamma_n}$, $\gamma_n\to\infty$, $\gamma_n=o(n)$, при $n\to\infty$ имеет место соотношение

$$\frac{\sqrt{2}\gamma_n}{|c|}\left(\alpha_{n,1}^*-\alpha\right)\Rightarrow\frac{\xi_1}{\xi_2},$$

где (ξ_1,ξ_2) – нормальный случайный вектор с нулевым средним и ковариационной матрицей $\begin{pmatrix}1 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 1\end{pmatrix}$.

3. Если $\alpha(n)=1-\frac{c}{\gamma_n}$, $\gamma_n\to\infty$, $\frac{n}{\gamma_n}\to 0$, при $n\to\infty$ имеет место утверждение

$$\frac{2n}{\sqrt{3}}\left(\alpha_{n,1}^*-\alpha\right)\Rightarrow\frac{\xi_1}{\xi_2},$$

где (ξ_1,ξ_2) – нормальный случайный вектор с нулевым средним и ковариационной матрицей $\begin{pmatrix}1 & \frac{\sqrt{3}}{2\sqrt{2}}\\ \frac{\sqrt{3}}{2\sqrt{2}} & 1\end{pmatrix}$.

Модель пространственной авторегрессии первого порядка с одним параметром

Далее рассматриваемый пространственный авторегрессионный процесс $\{X_{k,l}:k,l\in Z_+\}$ определяется следующим образом

$$X_{k,l}=\begin{cases}\alpha\left(X_{k-1,l}+X_{k,l-1}\right)+\varepsilon_{k,l},\ если\ \ k,l\ge 1,\\ 0,\qquad\qquad\qquad остальных\,случаях.\end{cases}\tag{9}$$

Эта модель устойчива (асимптотически стационарна), если $|\alpha|<1/2$ и неустойчива, если $|\alpha|=1/2$.

Рассмотрим прямоугольник

$$R_{m,n}:=\left\{(k,l)\in N^2:1\le k\le m\ \ and\ \ 1\le l\le n\right\}.$$

Оценка наименьших квадратов $\hat{\alpha}_{m,n}$, основанная на наблюдениях $\{X_{k,l} : (k,l) \in R_{m,n}\}$, минимизирует сумму квадратов

$$\sum_{(k,l)\in R_{m,n}} \left(X_{k,l} - \alpha\left(X_{k-1,l} + X_{k,l-1}\right)\right)^2$$

и имеет вид

$$\hat{\alpha}_{m,n} = \frac{\sum_{(k,l)\in R_{m,n}} \left(X_{k-1,l} + X_{k,l-1}\right) X_{k,l}}{\sum_{(k,l)\in R_{m,n}} \left(X_{k-1,l} + X_{k,l-1}\right)^2}.$$

Относительно асимптотического поведения этой оценки имеет место следующее утверждение.

Теорема [9]. Пусть $\{\varepsilon_{k,l} : k,l \in N\}$ – независимые случайные величины с $\sup\{E\varepsilon_{k,l}^4 : k,l \in N\} < \infty$.

Если $|\alpha| < 1/2$, то при $m,n \to \infty$, $m/n \to const > 0$

$$(mn)^{1/2}\left(\hat{\alpha}_{m,n} - \alpha\right) \xrightarrow{D} N\left(0, \sigma_\alpha^2\right),$$

где $$\sigma_\alpha^2 := \begin{cases} \dfrac{\alpha^2}{\left(1-4\alpha^2\right)^{-1/2} - 1}, & \text{если } \alpha \neq 0, \\ 1/2, & \text{если } \alpha = 0. \end{cases}$$

Если $|\alpha| = 1/2$, то при $m,n \to \infty$, $m/n \to const > 0$,

$$(mn)^{5/8}\left(\hat{\alpha}_{m,n} - \alpha\right) \xrightarrow{D} N\left(0, \sigma^2\right),$$

где $\sigma^2 := \dfrac{15\sqrt{\pi}}{32\left(2^{5/2}+3\right)}$.

Заметим, что $\sigma_0^2 = \lim\limits_{\alpha\to 0} \sigma_\alpha^2$, $\sigma^2 \neq \lim\limits_{\alpha\to 1/2} \sigma_\alpha^2 = 0$.

Далее предлагается оценка более простой структуры и использующая только часть наблюдений, расположенных вдоль диагонали в прямоугольнике $R_{m,n}$. Такой подход с существенным сокращением числа наблюдений является актуальным в задачах геологии и некоторых других областях применения моделей пространственной авторегрессии. Более простая структура оценки позволяет также снизить моментные ограничения на шумы.

Пусть для простоты $m = n$. Оценку будем строить основываясь на наблюдениях, расположенных вдоль диагонали квадрата $R_{n,n} = \{X_{k,l} : 1 \le k \le n, 1 \le l \le n\}$. В соответствии с (9) эти наблюдения удовлетворяют следующему рекуррентному соотношению

$$X_{k,k}=\begin{cases}\alpha(X_{k-1,k}+X_{k,k-1})+\varepsilon_{k,k}, & \textit{если} \quad k\ge 1,\\ 0, & \textit{в остальных случаях.}\end{cases} \qquad (10)$$

Суммируя (10) по k от 1 до n, получим уравнение оценивания

$$\sum_{k=1}^{n} X_{k,k}=\alpha\sum_{k=1}^{n}\left(X_{k-1,k}+X_{k,k-1}\right)+\sum_{k=1}^{n}\varepsilon_{k,k}.$$

Разделив обе части полученного уравнения на коэффициент при α, получим

$$\alpha_{n,n}^{*}-\alpha=Z_{n,n}/Y_{n,n}, \qquad (11)$$

где $Z_{n,n}=\sum_{k=1}^{n}\varepsilon_{k,k}$, $Y_{n,n}=\sum_{k=1}^{n}\left(X_{k-1,k}+X_{k,k-1}\right)$, $\alpha_{n,n}^{*}=\mathrm{X}_{n,n}/Y_{n,n}$ - предлагаемая оценка параметра α, а $\mathrm{X}_{n,n}=\sum_{k=1}^{n}X_{k,k}$.

Для построенной оценки имеет место следующее утверждение.

Теорема 2. Если $\left\{\varepsilon_{i,j}:\ i,j\ge 1\right\}$ – независимые одинаково распределенные случайные величины с $E\left(\varepsilon_{i,j}\right)=0$ и $E\left(\varepsilon_{i,j}\right)^2=1$.

1. В критическом случае $\left(\alpha=\pm 1/2\right)$, то при $n\to\infty$

$$2\sigma^{-1}n^{3/4}\left(\alpha_{n,n}^{*}-\alpha\right)\Rightarrow\frac{\xi_1}{\xi_2}\,sign\alpha,$$

где $\sigma^2=\dfrac{16}{15\sqrt{\pi}}\left(7+8\sqrt{2}\right)$, а ξ_1 и ξ_2 – независимые случайные величины, имеющие стандартное нормальное распределение.

2. Если $\alpha\left(n\right)=1/2-c/n$, $c>0$, то при $n\to\infty$

$$P\left(\sigma_c^{-1}n^{3/4}\left(\alpha_{n,n}^{*}-\alpha\right)<x\right)\Rightarrow\frac{1}{\pi}\int_{-\infty}^{x}\frac{du}{1+u^2},$$

где $\sigma_c^2=2/\sqrt{c}$.

Таким образом предложенные оценки имеют предельное распределение типа Коши, но моментные ограничения доведены до второго момента, в сравнении с теоремой [9], где требуется существование 4-го момента, но с нормальным предельном законом.

ЗАКЛЮЧЕНИЕ

Для модели обыкновенной авторегрессии первого порядка $X_k=\alpha X_{k-1}+\varepsilon_k$ найдено предельное распределение нестандартных оценок параметра α в случаях $\alpha=1-c/n$ и $\alpha=-1+c/n$.

Построенные оценки имеют, как правило, более простые предельные распределения в сравнении с традиционными оценками наименьших квадратов. При этом более простая структура предложенных оценок позволяет также при этом снизить моментные ограничения на «шумы», задающими стохастическую структуру моделей.

ЛИТЕРАТУРА

1. Peter C. B. Phillips (2021). Estimation and Inference with Near Unit Roots. Cowles foundation for research in economics yale university Box 208281 New Haven, Connecticut 06520-8281.
2. Shi, S. and P. C. B. Phillips (2021). Diagnosing housing fever with an econometric thermometer. *Journal of Economic Surveys, forthcoming*.
3. Yuhao Liu (2015). Finding moments of AR(k)-model parameter estimators. Brock Reports in Mathematics and Statistics No. 150504.
4. Jan Vrbik (2015). Moments of AR(k) parameter estimators. Communications in Statistics - Simulation and Computation 44 (2015) 1239-1252
5. Baran, S., Pap, G. (2011). Parameter estimation in a spatial unit root autoregressive model. J. Multivariate Anal. 107. -Pp. 282–305.
6. Baran, S., Pap, G. and Zuijlen, M. v. Asymptotic inference for an unstable spatial AR model. Statistics 38, 2004. -Pp.465–482.
7. Baran, S., Pap, G. and Zuijlen, M. v. Asymptotic inference for unit roots in spatial triangular autoregression. Department of Mathematics, Radboud University Nijmegen, The Netherlands, Report No. 0506 (April 2005). Url: www.inf.unideb.hu/˜barans/prepr.html.
8. Baran, S., Pap, G. and Zuijlen, M. v. (2007). Asymptotic inference for unit roots in spatial triangular autoregression. Acta Appl. Math. 96, -Pp. 17–42.
9. Baran. S., Pap. G., Martien C.A. Van Zuijlen (2004). Asymptotic inference for a nearly unstable sequence of stationary spatial AR models. Statist. Probab. Lett, 2004. -V.69. -Pp. 53-61.
10. Справочник по прикладной статистике. Т.2. -М.: Финансы и статистика, 1990. -526 с.
11. Mann H., Wald A. On the statistical treatment of linear stochastic difference equations. Econometrics, 1943. -V.11. -Pp. 173-220.
12. Anderson T.V. On asymptotic distributions of estimates of parameters of stochastic difference Equations. Ann. Math. Statist, 1959. -V.30. -Pp. 676-687.
13. Chan N.H., Wei C.Z. Asymptotic inference for nearly nonstationary AR(1) processes. Annals of Statistics, 1987. -V.15. -Pp. 1050-1063.
14. Chan N.H., Wei C.Z. Limiting distributions of least squares estimates of unstable autoregression processes. Annals of Statistics, 1988. -V.16. -№1. -Pp. 367-401.
15. White, J.S. The limiting distribution of the serial correlation coefficient in the explosive case. Ann. Math. Statist, 1958. -V. 29. -Pp. 1188-1197.
16. Startsev A.N. A new approach to estimation of an autoregressive parameter. Proc. of Sixth USSR-Japan Symp. World Scientific, 1991. -Pp.377-381.

17. Старцев А.Н., Мирзаев Т.С. О нестандартных методах оценивания в моделях авторегрессии в неустойчивых случаях. Журнал средневолжского математического общества, 2011. -Т.13. -№2. -С. 25-35.
18. Basu. S., Reinsel. G. C. Properties of the spatial unilateral first-order ARMA model. Adv. Appl. Probab, 1993. -V. 25. -Pp. 631-648.