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

Наука и Образование. МГТУ им. Н.Э. Баумана.
Электрон. журн. 2014. № 12. С. 407–415.
DOI: 10.7463/1214.0742725
Представлена в редакцию: 01.12.2014

УДК 004.3+519.6

Устойчивые оценки параметра авторегрессионного уравнения со случайным коэффициентом

Горяинов В. Б.[1,*], Горяинова Е. Р.[2]

[*] vb-goryainov@mail.ru

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

Работа посвящена анализу одной из наиболее распространенных нелинейных моделей временных рядов — авторегрессионной модели со случайными коэффициентами. Рассмотрена задача оценивания параметра одномерного авторегрессионного уравнения для стационарного случая. Предполагается, что распределение вероятности обновляющего процесса и авторегрессионного коэффициента неизвестно. Предложен робастный метод оценивания, основанный на подходе Хьюбера. Доказаны состоятельность и асимптотическая нормальность построенной оценки. Получено выражение для ее асимптотической относительной эффективности по отношению к оценке наименьших квадратов.

Ключевые слова: оценка наименьших квадратов; асимптотическая относительная эффективность; авторегрессионное уравнение; случайные коэффициенты; робастная оценка

Введение

Во многих областях науки и техники (см., например, [1, 2, 3]) наблюдения описываются уравнением авторегрессии

$$X_t = \Phi_t X_{t-1} + \varepsilon_t, \quad t = 0, \pm 1, \pm 2, \ldots, \tag{1}$$

в котором коэффициент Φ_t является случайным. В наиболее распространенном случае

$$\Phi_t = a + \eta_t, \quad t = 0, \pm 1, \pm 2, \ldots, \tag{2}$$

где η_t, $t = 0, \pm 1, \pm 2, \ldots,$ — последовательность независимых одинаково распределенных случайных величин с нулевыми математическими ожиданиями $\mathsf{E}\,\eta_t = 0$. Основной задачей анализа уравнения (1) является оценивание авторегрессионного параметра a. Традиционным методом оценивания является метод наименьших квадратов [4], который при разумных ограничениях на вероятностные свойства процессов η_t, ε_t позволяет получить состоятельные и асимптотически нормальные оценки.

В частном случае, когда коэффициент Φ_t неслучаен ($\eta_t = 0$), существуют оценки, имеющие во многих случаях бо́льшую эффективность, чем оценка наименьших квадратов. Например, М-оценки являются предпочтительнее оценки наименьших квадратов, если ε_t имеет распределение Тьюки [5].

В данной работе построены М-оценки параметра $a = \mathsf{E}\,\Phi_t$, доказана их состоятельность и асимптотическая нормальность, что позволило вычислить асимптотическую относительную эффективность этих оценок по отношению к оценке наименьших квадратов.

1. Постановка задачи

Далее всюду предполагается, что $\mathsf{E}\,\varepsilon_t = 0$ для любого $t = 0, \pm 1, \pm 2, \ldots$ и у случайных величин η_t и ε_t существуют конечные дисперсии:

$$\mathsf{D}\,\eta_t = \omega^2 < \infty, \quad \mathsf{D}\,\varepsilon_t = \sigma^2 < \infty, \tag{3}$$

удовлетворяющие условию

$$\omega^2 + \sigma^2 < 1. \tag{4}$$

При выполнении этих условий существует стационарное решение уравнения (1), представимое в виде сходящегося с вероятностью 1 ряда

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

где $\delta_0 = 1$ и

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

Обозначим через $\mathfrak{F}_k$ σ-алгебру событий, порожденную множеством случайных величин $\eta_t, \varepsilon_t, t \le k$. Оценка наименьших квадратов a_n^* параметра a по наблюдениям $X_0, X_1, \ldots, X_n$ случайного процесса X_t, описываемого уравнением (1), определяется как точка минимума функции [6]

$$L_{LS}(a) = \sum_{t=1}^{n} (X_t - \mathsf{E}(X_t | \mathfrak{F}_{t-1}))^2,$$

где $\mathsf{E}(X_t|\mathfrak{F}_{t-1})$ — условное математическое ожидание $\mathsf{E}\,X_t$ относительно σ-алгебры $\mathfrak{F}_{t-1}$. Поскольку $\mathsf{E}(X_t|\mathfrak{F}_{t-1}) = aX_{t-1}$, то

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

В [4] показано, что при выполнении (3)–(4) оценка наименьших квадратов является состоятельной и асимптотически нормальной. Состоятельность означает, что последовательность a_n^* при $n \to \infty$ сходится по вероятности к истинному значению a_0 параметра a, т.е. для любого $\varepsilon > 0$

$$\mathsf{P}\{|a_n^* - a| > \varepsilon\} \to 0 \quad \text{при } n \to \infty.$$

Асимптотическая нормальность последовательности a_n^* означает, что последовательность $\sqrt{n}(a_n^* - a_0)$ стремится по распределению к нормальной случайной величине с нулевым математическим ожиданием и дисперсией, равной

$$\frac{(1 - a_0^2 - \omega^2)^2(\sigma^2 \mathsf{E}\, X_0^2 + \omega^2 \mathsf{E}\, X_0^4)}{\sigma^4}. \tag{5}$$

Обобщим оценку наименьших квадратов, определив оценку $\hat{a}_n$ параметра a как точку минимума функции

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

где ρ — некоторая функция. Эта оценка является аналогом М-оценок в авторегрессионных моделях с неслучайными коэффициентами. В частном случае, когда $\rho(x) = x^2$, получается оценка наименьших квадратов.

Для построения робастных оценок в качестве ρ обычно выбирается четная функция, которая либо ограничена, либо растет на бесконечности медленнее, чем x^2. Наиболее распространенными являются ρ-функция Хьюбера [7]

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

и бивес Тьюки

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

Эти функции зависят от параметра $k > 0$, изменение которого позволяет регулировать степень робастности оценок.

2. Основные результаты

Основным результатом работы является доказательство состоятельности и асимптотической нормальности М-оценок $\hat{a}_n$, приведенное в следующей теореме.

Теорема 1. Пусть выполнены условия (3)–(4), плотности f и g случайных величин ε_t и η_t являются четными функциями, функция $\rho(x)$ выпукла, а ее вторая производная $\rho''(x)$ непрерывна и ограничена. Тогда случайная величина $\sqrt{n}(\hat{a}_n - a_0)$ является асимптотически нормальной с нулевым математическим ожиданием и дисперсией

$$\frac{\mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2]}{(\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1)X_0^2])^2}. \tag{7}$$

Д о к а з а т е л ь с т в о. Идея доказательства заключается в том, чтобы приблизить $\mathcal{F}(a)$ квадратичной формой (параболой) $\mathcal{F}^*(a)$, доказать асимптотическую нормальность решения a_n^* уравнения $\mathcal{F}^*(a) = 0$ и затем доказать, что $\sqrt{n}(\hat{a}_n - a_n^*) \to 0$ по вероятности при $n \to \infty$.

Заметим, что $\hat{a}_n$ является точкой минимума (6) тогда и только тогда, когда $\hat{\beta}_n = \sqrt{n}(\hat{a}_n - a_0)$ — точка минимума функции

$$L(b) = \sum_{t=1}^{n} \rho\Big(X_t - a_0 X_{t-1} - \frac{\beta X_{t-1}}{\sqrt{n}}\Big) - \sum_{t=1}^{n} \rho(X_t - a_0 X_{t-1})$$

или, с учетом

$$X_t - a_0 X_{t-1} = \eta_t X_{t-1} + \varepsilon_t$$

— точка минимума функции

$$L(b) = \sum_{t=1}^{n} \rho\Big(\eta_t X_{t-1} + \varepsilon_t - \frac{\beta X_{t-1}}{\sqrt{n}}\Big) - \sum_{t=1}^{n} \rho(X_t - a_0 X_{t-1}).$$

Разложив левую часть $L(b)$ в окрестности нуля по степеням β, получим

$$L(b) = -A_n b + \frac{1}{2} B_n b^2 + \alpha_n(b),$$

где

$$A_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \rho'(\eta_t X_{t-1} + \varepsilon_t) X_{t-1}, \quad B_n = \frac{1}{n} \sum_{t=1}^{n} \rho''(\eta_t X_{t-1} + \varepsilon_t) X_{t-1}^2,$$
$$\alpha_n(b) = \frac{1}{2n} \sum_{t=1}^{n} \Big(\rho''\Big(\eta_t X_{t-1} + \varepsilon_t - \frac{\tau b X_{t-1}}{\sqrt{n}}\Big) - \rho''(\eta_t X_{t-1} + \varepsilon_t)\Big) X_{t-1}^2 b^2, \quad 0 < \tau < 1.$$

Случайные последовательности

$$\zeta_{1t} = \rho''(\eta_t X_{t-1} + \varepsilon_t) X_{t-1}^2$$

и

$$\zeta_{2t} = \Big(\rho''\Big(\eta_t X_{t-1} + \varepsilon_t - \frac{\tau b X_{t-1}}{\sqrt{n}}\Big) - \rho''(\eta_t X_{t-1} + \varepsilon_t)\Big) X_{t-1}^2$$

являются стационарными и эргодическими как преобразования процессов типа «белого шума» η_t и ε_t [8, p. 170, 182]. Из стационарности ζ_{2t} следует, что

$$|\alpha_n(b)| \leq \frac{1}{2} \mathsf{E}\left[\Big|\rho''\Big(\eta_1 X_0 + \varepsilon_1 - \frac{\tau b X_0}{\sqrt{n}}\Big) - \rho''(\eta_1 X_0 + \varepsilon_1)\Big| X_0^2 b^2\right].$$

Поэтому, а также в силу непрерывности и ограниченности ρ'' по теореме Лебега о мажорируемой сходимости $\mathsf{E}\,|\alpha_n(b)| \to 0$ при $n \to \infty$, откуда вытекает, что $\alpha_n(b) = o_p(1)$ при $n \to \infty$ (здесь и в дальнейшем $o_p(1)$ означает последовательность случайных величин, сходящихся к нулю по вероятности). Из стационарности и эргодичности ζ_{1t} следует [8, p. 181], что существует предел по вероятности

$$B = \lim_{n\to\infty} B_n = \mathsf{E}\Big[\rho''(\eta_1 X_0 + \varepsilon_1)|X_0^2\Big].$$

Поэтому

$$L(b) = -A_n b + \frac{1}{2}Bb^2 + \gamma_n(b),$$

где $\gamma_n(b) = o_p(1)$ при $n \to \infty$.

Обозначим через $\tilde{b}_n = \dfrac{A_n}{B}$ точку минимума функции $-A_n b + \dfrac{1}{2}Bb^2$. Покажем, что последовательность $\tilde{b}_n$ асимптотически нормальна и $\hat{b}_n - \tilde{b}_n = o_p(1)$, откуда будет следовать асимптотическая нормальность последовательности $\hat{b}_n$.

Так как плотности f и g случайных величин ε_t и η_t — четные функции, а ρ' — нечетная функция, то $\mathsf{E}[\rho'(\eta_t X_{t-1} + \varepsilon_t)_X t - 1|\mathfrak{F}_{t-1}] = 0$. Поэтому [9, теорема 23.1] последовательность A_n является асимптотически нормальной с нулевым математическим ожиданием и дисперсией $\mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1)X_0)^2]$. Следовательно, последовательность $\tilde{b}_n$ является асимптотически нормальной с нулевым математическим ожиданием и дисперсией (7).

Теперь докажем, что $\hat{b}_n - \tilde{b}_n = o_p(1)$ при $n \to \infty$. Так как $L(b)$ выпукла, то выпуклой будет и последовательность функций $L(b) + A_n b$, которая при $n \to \infty$ сходится по вероятности к выпуклой функции Bb^2. Поэтому [10] $\sup\limits_{b\in K} |\gamma_n(b)| = o_p(1)$ для любого компакта $K \subset \mathbb{R}$ при $n \to \infty$. Покажем, что для любого $\delta > 0$

$$\lim_{n\to\infty} \mathsf{P}\left\{|\hat{b}_n - \tilde{b}_n| > \delta\right\} = 0, \tag{8}$$

что равносильно $\hat{b}_n - \tilde{b}_n = o_p(1)$. Обозначим через U_n отрезок $[\tilde{b}_n, \tilde{b}_n + \delta]$. Так как последовательность $\tilde{b}_n$ сходится по распределению, то существует такой компакт $K \subset \mathbb{R}$, что $\mathsf{P}\{U_n \subset K\}$ сколь угодно близка к единице. Поэтому при $n \to \infty$

$$\Delta_n = \sup_{b\in U_n} |\gamma_n(b)| = o_p(1).$$

Так как $L(b)$ выпукла, то для $b_1 = \tilde{b}_n + \delta$ и для любого $b = \tilde{b}_n + s$, где $s > \delta$,

$$\left(1 - \frac{\delta}{s}\right)L(\tilde{b}_n) + \frac{\delta}{s}L(b) \geq L(b_1).$$

Поэтому

$$L(b) \geq \frac{s}{\delta}(L(b_1) - L(\tilde{b}_n)) + L(\tilde{b}_n).$$

Так как

$$L(b_1) = L(\tilde{b}_n) + \frac{1}{2}\delta^2 B + \gamma_n(b_1) - \gamma_n(\tilde{b}_n),$$

то

$$\inf_{b\in U_n} L(b) = L(\tilde{b}_n) + \inf_{s>\delta} \frac{s}{\delta}(L(b_1) - L(\tilde{b}_n)) \geq L(\tilde{b}_n) + \frac{1}{2}\delta^2 B - 2\Delta_n.$$

Поэтому с вероятностью, стремящейся к единице, минимум $L(b)$ не может быть больше, чем $\tilde{b}_n + \delta$.

Аналогично показывается, что минимум $L(b)$ не может быть меньше, чем $\tilde{b}_n - \delta$. Таким образом, для любого $\delta > 0$ выполняется (8) и, следовательно, асимптотическое распределение $\hat{b}_n$ совпадает с асимптотическим распределением $b = \tilde{b}_n$. Теорема доказана.

Полученная теорема позволяет найти асимптотическую относительную эффективность e М-оценки по отношению к оценке наименьших квадратов. Асимптотическая относительная эффективность оценок определяется как обратное отношение их асимптотических дисперсий. Поэтому из (5) и (7) следует, что

$$e = \frac{(1 - a_0^2 - \omega^2)^2(\sigma^2 \,\mathsf{E}\, X_0^2 + \omega^2 \,\mathsf{E}\, X_0^4)(\mathsf{E}[\rho''(\eta_1 X_0 + \varepsilon_1) X_0^2])^2}{\mathsf{E}[(\rho'(\eta_1 X_0 + \varepsilon_1) X_0)^2]\sigma^4}.$$

Вычислить e для различных вероятностных распределений процессов η_t и ε_t можно при помощи компьютерного моделирования.

Заключение

В работе рассмотрена нелинейная модель временных рядов, описываемая уравнением авторегрессии со случайным коэффициентом. Для параметра авторегрессионного уравнения построены оценки, занимающие промежуточное положение между оценками наименьших квадратов и наименьших модулей. При достаточно умеренных предположениях о вероятностном распределении обновляющего процесса и авторегрессионного коэффициента доказаны состоятельность и асимптотическая нормальность предложенных оценок. Получено явное выражение для асимптотической относительной эффективности этих оценок по отношению к оценке наименьших квадратов.

Список литературы

1. Aknouche A. Two-stage weighted least squares estimation of nonstationary random coefficient autoregressions // J. Time Ser. Econom. 2013. Vol. 5, no. 1. P. 25–46.
2. Liang Y., Thavaneswaran A., Ravishanker N. RCA models: joint prediction of mean and volatility // Statist. Probab. Lett. 2013. Vol. 83, no. 2. P. 527–533.
3. Prášková Z., Vaněček P. On a class of estimators in a multivariate RCA(1) model // Kybernetika. 2011. Vol. 47, no. 4. P. 501–518.
4. Nicholls D.F., Quinn B.G. Random coefficient autoregressive models: an introduction. New York: Springer, 1982. 154 p.
5. Maronna R.A., Martin D., Yohai V. Robust Statistics: Theory and Methods. Chichester: Wiley, 2006. 403 p.
6. Tjøstheim D. Estimation in nonlinear time series models // Stochastic Process. Appl. 1986. Vol. 21, no. 2. P. 251–273. DOI: 10.1016/0304-4149(86)90099-2

7. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В.А. Робастность в статистике. Подход на основе функций влияния: пер. с англ. М.: Мир, 1989. 512 с. [Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. Robust statistics. The approach based on influence functions. New York: Wiley, 1986. 502 p.]
8. Stout W.F. Almost Sure Convergence. New York: Academic Press, 1974. 381 p.
9. Биллингсли П. Сходимость вероятностных мер: пер. с англ. М.: Наука, 1977. 352 с. [Billingsley P. Convergence of probability measures. New York: Wiley, 1968. 253 p.]
10. 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.

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

Science and Education of the Bauman MSTU, 2014, no. 12, pp. 407–415.
DOI: **10.7463/1214.0742725**
Received: 01.12.2014

Stable Parameter Estimation for Autoregressive Equations with Random Coefficients

Goryainov V. B.[1,*], Goryainova E. R.[2]

[*] vb-goryainov@mail.ru

[1] Bauman Moscow State Technical University, Moscow, Russia
[2] Higher School of Economics, Moscow, Russia

Keywords: least squares estimate, asymptotic relative efficiency, autoregressive equation, random coefficients, robust estimate

In recent years there has been a growing interest in non-linear time series models. They are more flexible than traditional linear models and allow more adequate description of real data. Among these models a autoregressive model with random coefficients plays an important role. It is widely used in various fields of science and technology, for example, in physics, biology, economics and finance. The model parameters are the mean values of autoregressive coefficients. Their evaluation is the main task of model identification. The basic method of estimation is still the least squares method, which gives good results for Gaussian time series, but it is quite sensitive to even small disturbances in the assumption of Gaussian observations. In this paper we propose estimates, which generalize the least squares estimate in the sense that the quadratic objective function is replaced by an arbitrary convex and even function. Reasonable choice of objective function allows you to keep the benefits of the least squares estimate and eliminate its shortcomings. In particular, you can make it so that they will be almost as effective as the least squares estimate in the Gaussian case, but almost never loose in accuracy with small deviations of the probability distribution of the observations from the Gaussian distribution.

The main result is the proof of consistency and asymptotic normality of the proposed estimates in the particular case of the one-parameter model describing the stationary process with finite variance. Another important result is the finding of the asymptotic relative efficiency of the proposed estimates in relation to the least squares estimate. This allows you to compare the two estimates, depending on the probability distribution of innovation process and of autoregressive coefficients. The results can be used to identify an autoregressive process, especially with non-Gaussian nature, and/or of autoregressive processes observed with gross errors.

References

1. Aknouche A. Two-stage weighted least squares estimation of nonstationary random coefficient autoregressions. *J. Time Ser. Econom.*, 2013, vol. 5, no. 1, pp. 25–46.
2. Liang Y., Thavaneswaran A., Ravishanker N. RCA models: joint prediction of mean and volatility. *Statist. Probab. Lett.*, 2013, vol. 83, no. 2, pp. 527–533.
3. Prášková Z., Vaněček P. On a class of estimators in a multivariate RCA(1) model. *Kybernetika*, 2011, vol. 47, no. 4, pp. 501–518.
4. Nicholls D.F., Quinn B.G. *Random coefficient autoregressive models: an introduction*. New York, Springer, 1982. 154 p.
5. Maronna R.A., Martin D., Yohai V. *Robust Statistics: Theory and Methods*. Chichester, Wiley, 2006. 403 p.
6. Tjøstheim D. Estimation in nonlinear time series models. *Stochastic Process. Appl.*, 1986, vol. 21, no. 2, pp. 251–273. DOI: 10.1016/0304-4149(86)90099-2
7. Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. *Robust statistics. The approach based on influence functions*. New York, Wiley, 1986. 502 p. (Russ. ed.: Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. *Robastnost' v statistike. Podkhod na osnove funktsiy vliyaniya*. Moscow, Mir Publ., 1989. 512 p.).
8. Stout W.F. *Almost Sure Convergence*. New York, Academic Press, 1974. 381 p.
9. Billingsley P. *Convergence of probability measures*. New York: Wiley, 1968. 253 p. (Russ. ed.: Billingsley P. *Skhodimost' veroyatnostnykh mer*. Moscow, Nauka Publ., 1977. 352 p.).
10. 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.