УДК 519.8

Вестник СибГАУ
2014. № 5(57). С. 57–61

МОДЕЛИРОВАНИЕ СЛОЖНЫХ ПРОЦЕССОВ ПРИ ПОМОЩИ АВТОРЕГРЕССИИ

А. А. Городов, Л. В. Городова, А. А. Кузнецов

Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, 660014, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: Glexx84@mail.ru

Сделан обзор по использованию числовых рядов для оценки параметров в моделях авторегрессии. Описан метод и основная процедура определения порядка модели авторегрессии и выбора числового ряда. Цель настоящей работы – обобщение полученных результатов, выявление новых характеристик авторегрессии при моделировании сложных процессов и упрощение процедуры подбора параметров, основанного на использовании числовых рядов. Основные теоретические и прикладные результаты работы получены на основе методологии системного анализа, теории случайных процессов, а также информационных технологий и методов фундаментальной и прикладной математики. Проведено сопоставление метода числовых рядов с известными способами оценки параметров в моделях авторегрессии. Приведена взаимосвязь между моделями авторегрессии и такими классическими математическими понятиями, как золотое сечение, треугольник Паскаля, а также числами Фибоначчи и трибоначчи. Помимо этого даны оценки для моделей авторегрессии при долгосрочном прогнозировании. Показана взаимосвязь между авторегрессией четвертого порядка с рядом чисел квадробоначчи. Обобщены известные рекомендации по применению метода числовых рядов. Данные рекомендации позволяют упростить процедуру подбора числового ряда и повысить качество модели. При краткосрочном и среднесрочном прогнозировании процессов с постоянным средним рекомендуется использовать коэффициенты, которые вычисляются при помощи рядов Фибоначчи и трибоначчи. Полученные результаты работы будут полезны при моделировании сложных процессов с временной составляющей.

Ключевые слова: авторегрессия, метод наименьших квадратов, метод числовых рядов, оценка параметров авторегрессии, свойства прогнозов.

Vestnik SibGAU
2014, No. 5(57), P. 57–61

MODELING OF COMPLEX PROCESSES USING AUTOREGRESSIVE METHODS

A. A. Gorodov, L. V. Gorodova, A. A. Kuznetsov

Siberian State Aerospace University named after academician M. F. Reshetnev
31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660014, Russian Federation
E-mail: Glexx84@mail.ru

We reviewed the use of numerical series to estimate the parameters in the autoregressive model in this paper. A method and the basic procedure for determining the order of autoregressive model and the choice of a numerical series are described. The purpose of this work is the generalization of the obtained results, identifying characteristics of the autoregressive when modeling complex processes and simplified procedures for the selection of parameters is based on the use of numerical series. The basic theoretical and practical results are obtained on the basis of methodology of system analysis, the theory of random processes, information technologies and methods of fundamental and applied mathematics. We compared the method of numerical series with known methods of parameter estimation in autoregressive models. It is shown the relationship between the autoregressive models and classical mathematical concepts such as the 'golden section', 'Pascal's triangle', the Fibonacci numbers and the Tribonacci numbers. Also we gave estimates for autoregressive models for long-term forecasting. It is shown the relationship between the fourth-order autoregressive with the series of the Kvadrobonacci numbers. We summarized the known recommendations on the application of numerical series. These recommendations help to simplify the procedure for selection of numerical series and improve the quality of the model. For short and medium-term forecasts with a constant average it is recommended to use coefficients which are calculated using the Fibonacci and the Tribonacci series. The obtained results will be useful in modeling of complex processes with a time component.

Keywords: the autoregressive, method of numerical series, parameter estimation of autoregressive, golden section, Fibonacci numbers.

Введение. С давних пор человек пытается предсказать поведение исследуемых объектов в будущем, используя прогноз как инструмент управления и регулирования. Поэтому в теории системного анализа и прикладной математике выделились целые разделы, подчас выступающие как отдельные направления науки. К сожалению, существующий математический аппарат не может дать точного ответа о возможном поведении объекта, поскольку всегда присутствует вероятностная составляющая данного прогноза. В большинстве случаев проблема прогнозирования решается при помощи моделирования случайных процессов.

Большинство работ, посвященных проблеме моделирования, основываются на принципах саморегулирования и самоподобия. В практических задачах прогнозирования процессов часто применяется модель авторегрессии *p*-го порядка. Данному вопросу посвящены труды С. А. Айвазяна, С. Алмон, Т. Андерсон, Дж. Бокса, Р. Г. Брауна, Вольда, Г. М. Дженкинса, Л. М. Койка, Я. Р. Магнуса, А. И. Орлова, П. Р. Уинтерса, С. Хольта и многих других [1–10].

В AR(*p*)-моделях одними из самых актуальных вопросов являются определение *p*-длины предыстории процесса, а также выбор метода подбора параметров авторегрессии. В статье [11] был предложен метод подбора параметров в моделях авторегрессии на основе числовых рядов. В работах [12–15] доказана взаимосвязь прогнозов по авторегрессии с золотым сечением, треугольником Паскаля и числами Фибоначчи. Помимо этого, указанные работы определили отдельное направление исследований в области моделирования сложных процессов авторегрессионными методами, основанными на числовых рядах. В данных исследованиях остались неразрешенными вопросы о свойствах прогнозов в моделях авторегрессии высоких порядков, а также возможности расширения области применения авторегрессионных моделей с методом подбора параметров на основе числовых рядов.

Вышесказанное определило цель работы – обобщение полученных результатов, выявление новых характеристик авторегрессии при моделировании сложных процессов и упрощение процедуры подбора параметров, основанного на использовании числовых рядов.

Основные теоретические и прикладные результаты получены на основе методологии системного анализа, теории случайных процессов, а также информационных технологий и методов фундаментальной и прикладной математики.

Метод числовых рядов (МЧР). В работе [11] был предложен метод подбора параметров в AR(*p*)-моделях, основанный на использовании нормированных числовых рядов (сумма которых равна 1). Сделан сравнительный анализ результатов моделирования временных рядов на основе МЧР с другими известными методами. В дальнейшем были рассмотрены не только нормированные, но и другие сходящиеся числовые ряды [14; 16].

Прогнозное значение y_{t+1} вычисляется следующим образом:

$$y_{t+1}^{(m;p)} = \sum_{i=0}^{p-1} b_i^{(m;p)} x_{t-i+1}\,, \qquad (1)$$

где m – номер сходящегося числового ряда из некоторой базы рядов; p – порядок модели, верхний индекс $(m;p)$ – указывает на номер ряда и на порядок модели.

Для определения оптимального порядка модели *p*, а также вида $\sum_{i=0}^{\infty} b_i$ будем использовать функцию

$$\Theta\left(m\left\{b_0, b_1, \ldots, b_{p-1}\right\}, p\right) = \min_{1\le m\le M}\left\{\min_{1\le p\le p_0}\left\{\Delta^{(m,p)}\right\}\right\}, \qquad (2)$$

где p_0 – граница порядка модели; $\Delta^{(m,p)} = \frac{1}{t-1}\sum\left(\frac{x_i - y_{t+1}^{(m;p)}}{x_i}\right)^2 100\ \%$ – среднеквадратическая относительная погрешность модели.

Если $x_i = 0$ для некоторых i, то необходимо использовать новую систему координат $x_i' = x_i + x_0$, в которой все $x_i' \ne 0$, где $x_0 = \text{const}$.

Последующие теоретические исследования позволили выявить ряд замечаний.

Замечание 1. Формула (1) представляет собой авторегрессию, аналогичную разложению по полной предыстории, с коэффициентами, подобранными по модифицированному методу Койка [10].

Замечание 2. Использование нормированных числовых рядов удовлетворяет условию стационарности процесса, поэтому полученные оценки авторегрессии состоятельные, эффективные и несмещенные. Но для моделирования нестационарных процессов необязательно использовать условие нормированности ряда.

Замечание 3. Если *p* = *T*, то формула (1) представляет собой модель прогнозирования по полной предыстории (разложение Вольда [1; 2; 16]) с условием, что прогнозное значение $y_{t+1}^{(m;p)}$ будет представлять собой сумму предыстории процесса с весовыми коэффициентами, являющимися элементами числового ряда (аналогично методам Альмон [9] и Койка [10]).

Замечание 4. Функция (2) аналогична оптимизирующей функции МНК и позволяет не только подобрать оптимальные коэффициенты, но и определить порядок авторегрессии.

Свойства прогнозов в моделях авторегрессии. В большинстве задач прогнозирования необходимо делать прогноз на k значений вперед из x_t. Как указано выше, прогноз на одно значение вперед согласно методу числовых рядов будет рассчитываться по формуле (1).

В МЧР используются в качестве коэффициентов числовые ряды, и это позволило выделить ряд свойств прогнозов в моделях авторегрессии, характерных не только для МЧР, но и в целом для авторегрессионных моделей.

Предложение 1. Допустим $\sum_{i=0}^{\infty} b_i = 1$ – нормированный знакоположительный степенной ряд, где $0 \le b_i < 1$ и $b_n = b_0^{n+1}$, т. е. $\sum_{i=0}^{\infty} b_i = b_0 + b_0^2 + b_0^3 + b_0^4 + b_0^5 + \cdots$

Согласно введенному предложению были сформулированы и доказаны следующие теоремы [12–15].

Теорема 1. Пусть $y_{t+k} = \alpha_1(k)x_t + \alpha_2(k)x_{t-1}$ – прогнозная модель AR(2) по МЧР. Тогда справедливы следующие утверждения:

1. $\alpha_1(k) = F_2(k+1)b_0^k$,

где $F_2(k)$ – ряд Фибоначчи с начальными условиями $F_2(1) = F_2(2) = 1$.

2. $\alpha_2(k) = F_2(k)b_0^{k+1}$.

3. $\alpha_2(k) = b_1^2\alpha_1(k-1)$.

4. $\lim_{k\to\infty} y_{t+k} = 0$ при $b_0 < \frac{2}{1+\sqrt{5}}$.

5. $\lim_{k\to\infty} y_{t+k} = \frac{1}{\sqrt{5}}\frac{1+\sqrt{5}}{2}x_t + \frac{1}{\sqrt{5}}\frac{2}{1+\sqrt{5}}x_{t-1}$

при $b_0 = \frac{2}{1+\sqrt{5}}$.

6. $\lim_{k\to\infty} y_{t+k} = \infty$ при $b_0 > \frac{2}{1+\sqrt{5}}$.

Данная теорема позволяет сделать вывод о том, что любой прогноз в модели AR(2) – это распределение предыстории в будущем через золотое сечение.

Теорема 2. Пусть $y_{t+k} = \alpha_1(k)x_t + \alpha_2(k)x_{t-1} + \alpha_3(k)x_{t-2}$ – прогнозная модель AR(3) по МЧР. Тогда справедливы следующие утверждения:

1. $\alpha_1(k) = F_3(k+1)b_0^k$,

где $F_3(k) = \frac{3\beta\left(\frac{1}{3}(\lambda_1+\lambda_2+1)\right)^k}{\beta^2-2\beta+4}$ – ряд чисел трибоначчи, $\beta = \sqrt[3]{586+102\sqrt{33}}$, $\lambda_1 = \sqrt[3]{19+3\sqrt{33}}$, $\lambda_2 = \sqrt[3]{19-3\sqrt{33}}$.

2. $\alpha_2(k) = F_3'(k+1)b_0^{k+1}$.

3. $\alpha_3(k) = F_3(k)b_0^{k+2}$.

4. $\alpha_3(k) = b_0^3\alpha_1(k-1)$.

5. $\lim_{k\to\infty} y_{t+k} = 0$ при $b_0 < \frac{3}{(\lambda_1+\lambda_2+1)}$.

6. $\lim_{k\to\infty} y_{t+k} = \frac{\beta[\lambda_1+\lambda_2+1]}{\beta^2-2\beta+4}x_t + \frac{3\beta}{\beta^2-2\beta+4}x_{t-1} + \frac{27\beta}{(\beta^2-2\beta+4)(\lambda_1+\lambda_2+1)^2}x_{t-2}$ при $b_0 = \frac{3}{(\lambda_1+\lambda_2+1)}$.

7. $\lim_{k\to\infty} y_{t+k} = \infty$ при $b_0 > \frac{3}{(\lambda_1+\lambda_2+1)}$.

Полученная теорема дает возможность делать прогноз в условиях стационарности рассматриваемых процессов без построения самой модели. Доказанные утверждения свидетельствуют о том, что прогноз авторегрессионных моделей при определенных условиях имеет распределение золотого сечения.

Для авторегрессионного процесса четвертого порядка удалось выявить ряд свойств, характерных для AR(2) и AR(3) [17].

Так, в общем AR(4) можно представить как

$$y_{t+k} = \alpha_1(k)x_t + \alpha_2(k)x_{t-1} + \alpha_3(k)x_{t-2} + \alpha_4(k)x_{t-3}.$$

Доказано следующее:

Лемма. Пусть $\sum_{i=0}^{\infty} b_0^i$ – нормированный знакоположительный степенной ряд, где $0 \le b_i < 1$. Тогда,

$$\alpha_1(k) = F_4(k+1)b_0^k,$$

где $F_4(k+1)$ – ряд чисел квадробоначчи, $F_4(k+1) = F_4(k) + F_4(k-1) + F_4(k-2) + F_4(k-3)$, с условием $F_4(0) = 0$, $F_4(1) = 1$, $F_4(2) = 1$, $F_4(3) = 2$, $F_4(4) = 4$;

$$\alpha_2(k) = F_4'(k+1)b_0^{k+1},$$

где $F_4'(k+1)$ – первый модифицированный ряд чисел квадробоначчи, причем $F_4'(k+1) = F_4(k) - F_4(k-4)$;

$$\alpha_3(k) = F_4''(k+1)b_0^{k+2},$$

где $F_4''(k+1)$ – второй модифицированный ряд чисел квадробоначчи, причем $F_4''(k+1) = F_4'(k) - F_4(k-3) = F_4(k) - F_4(k-4) - F_4(k-3)$;

$$\alpha_4(k) = F_4(k)b_0^{k+3}.$$

Лемма доказана и вводит взаимосвязь AR(4) с рядом чисел квадробоначчи.

Рекомендации по применению МЧР. На основе доказанных теорем и практических расчетов можно предложить рекомендации по применению числовых рядов при подборе параметров в авторегрессии малых порядков:

1. Для моделирования стационарных динамических процессов необходимо использовать нормированные числовые ряды.

2. Для процессов, имеющих убывающую тенденцию, необходимо использовать нормированные числовые ряды, сумма первых элементов которых меньше 1.

3. Если динамический процесс содержит возрастающий тренд, то эффективнее использовать при моделировании ненормированные сходящиеся числовые ряды.

4. При моделировании периодических процессов лучшими являются знакопеременные и тригонометрические ряды.

5. При краткосрочном и среднесрочном прогнозировании процессов с постоянным средним рекомендуется использовать для AR(2) в качестве весовых коэффициентов $\frac{1}{\sqrt{5}}\frac{1+\sqrt{5}}{2}$ и $\frac{1}{\sqrt{5}}\frac{2}{1+\sqrt{5}}$, для AR(3) – $\frac{\beta\left[\lambda_1+\lambda_2+1\right]}{\beta^2-2\beta+4}$, $\frac{3\beta}{\beta^2-2\beta+4}$ и $\frac{27\beta}{\left(\beta^2-2\beta+4\right)\left(\lambda_1+\lambda_2+1\right)^2}$.

Заключение. Теория применения авторегрессионных моделей при анализе сложных процессов довольно обширна и многогранна, но в то же время имеет большое количество ограничений (например, при использовании МНК, МНП и др.). Помимо этого не до конца изучен весь спектр направлений использования данных методов. В свою очередь, рассмотренный метод подбора параметров в моделях авторегрессии на основе числовых рядов органично дополняет классические, что позволяет расширить область их применения. Также МЧР дал возможность получить зависимости между моделями авторегрессии и такими классическими математическими понятиями, как золотое сечение и треугольник Паскаля. Кроме того, были обобщены рекомендации по применению МЧР, что позволяет сократить время обработки данных, а также повысить качество модели.

Библиографические ссылки

1. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. М. : Юнити-Дана. 1998. 1022 с.
2. Эконометрия / А. А. Цыплаков [и др.]. Новосибирск : Изд-во СО РАН. 2005. 744 с.
3. Brown R. G. Smoothing forecasting and prediction of discrete time series. N. Y., 1963.
4. Green W. H. Econometric analysis. Macmillan Publishing Company, N. Y., 1993.
5. Holt C. C. Forecasting trends and seasonals by exponentially weighted moving averages // O. N. R. Memorandum, Carnegie Inst. of Technology. 1957. № 2.
6. Winters P. R. Forecasting Sales by Exponentially Weighted Moving Averages. Mgmt. Sci. No. 6. P. 324.
7. Yule G. U. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis // Phil. Trans. Royal Soc. London B, 1925. P. 21–87.
8. Канторович Г. Г. Анализ временных рядов // Экономический журнал Высшей школы экономики. 2002. Том. 6. № 1–3.
9. Almon S. The Distributed Lag between Capital Appropriations and Expen-ditures. Econometrica, 1965.
10. Koyck L. M. Distributed Lags and Investment Analysis. Amsterdam : North- Holland, 1954.
11. Городов А. А. Моделирование временных рядов на основе нормированных числовых рядов // Системы управления и информационные технологии. 2010. Т. 39, № 1. С. 4–7.
12. Городов А. А., Кузнецов А. А. Свойства прогнозов в моделях авторегрессии по методу нормированных числовых рядов / Системы управления и информационные технологии. 2011. Т. 45, № 3. С. 12–16.
13. Городов А. А. Метод подбора параметров в моделях авторегрессии на основе числовых рядов : дис. … канд. физ.-мат. наук / СибГАУ. Красноярск. 2011. 114 с.
14. Городов А. А., Кузнецов А. А., Демьяненко О. В. Золотое сечение и прогнозирование по авторегрессии // Вестник КрасГАУ. 2012. № 2 (42). С. 47–64.
15. Городов А. А. Метод подбора параметров в моделях авторегрессии на основе числовых рядов : автореф. дис. … канд. физ.-мат. наук / СибГАУ. Красноярск. 2011. 19 с.
16. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып. 1–2. М. : Мир, 1974. 604 с.
17. Городов А. А., Демьяненко О. В. Свойства прогнозов в моделях авторегрессии четвертого порядка // Вестник КрасГАУ. 2014. № 8. С. 28–33.

References

1. Ayvazian S. A., *Mkhitaryan C. C. Prikladnaya statistika i osnovi ekonometriki* [Applied statistics and econometrics]. Moscow, UNITY-DANA Publ., 1998, 1022 p.
2. Tsyplakov A. A., V. I. Suslov, N. M. Ibragimov, L. P. Malysheva, A. A. Tsyplakov. Ekonometriya [Econometrics] Novosibirsk, SO RAN Publ., 2005, 744 p.
3. Brown R. G. Smoothing forecasting and prediction of discrete time series. N. Y., 1963.
4. Green W. H. Econometric analysis. Macmillan Publishing Company, New York, 1993.
5. Holt C. C. Forecasting trends and seasonals by exponentially weighted moving averages. *O. N. R. Memorandum*, Carnegie Inst. of Technology, 1957, no. 2.
6. Winters P. R. Forecasting Sales by Exponentially Weighted Moving Averages. Mgmt. Sci., no. 6, p. 324.
7. Yule G. U. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis. Phil. Trans. Royal Soc. London B, 1925. P. 21–87.
8. Kantorovich, G. G. [Time series analysis]. *Ekonomicheskiy zhurnal Vysshey shkoly ekonomiki*, 2002, vol. 6, no. 1–3 (In Russ.).
9. Almon S. The Distributed Lag between Capital Appropriations and Expen-ditures. Econometrica, 1965.
10. Koyck L. M. Distributed Lags and Investment Analysis. Amsterdam : North- Holland, 1954.
11. Gorodov A. A. [Modeling time series based on the normalized number series]. *Sistemy upravleniya i informatsionnye tekhnologii.* 2010, vol. 39, no. 1, p. 4–7 (In Russ.).

12. Gorodov A. A., Kuznetsov A. A. [Property predictions in models of autoregressive method normalized numerical series]. *Sistemy upravleniya i informatsionnye tekhnologii*, 2011, vol. 45, no. 3, p. 12–16 (In Russ.).

13. Gorodov A. A. *Metod podbora parametrov v modeliakh avtoregressii na osnove chislovykh ryadov. Dis. kand. fiz.-mat. nauk.* [Method of selection of parameters in autoregressive models based on the numerical series. Candidate physical-mat. sci. diss.]. Krasnoyarsk, SibGAU Publ., 2011, 114 p.

14. Gorodov A. A., Kuznetsov A. A., Demyanenko O. V. [The Golden section and the prediction of autoregressive]. *Vestnik KrasGAU*. 2012, no. 2 (42), p. 47–64 (In Russ.).

15. Gorodov A. A. *Metod podbora parametrov v modeliakh avtoregressii na osnove chislovykh ryadov. Avtoref. dis. kand. fiz.-mat. nauk.* [Method of selection of parameters in autoregressive models based on the numerical series. Candidate physical-mat. sci. author. diss.]. Krasnoyarsk, SibGAU Publ., 2011, 19 p.

16. Box J., Jenkins G., *Analiz vremennykh riadov. Prognoz i upravlenia.* [Time series analysis. Prognosis and management] Moscow, Mir Publ., Vol. 1–2, 1974. 604 p.

17. Gorodov A. A., Demyanenko O. V. [Property predictions in models of fourth order autoregressive]. *Vestnik KrasGAU*. 2014, no. 8, p. 28–33 (In Russ.).