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

Всероссийская молодежная научная конференция студентов, аспирантов и молодых ученых «Все грани математики и механики»

(23–27 апреля 2019 г.)

Сборник статей

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

Томск
Издательский Дом Томского государственного университета
2019

О последовательном оценивании параметров авторегрессионной модели с непрерывным временем

Хабарова Е. Л., Емельянова Т. В.

Томский государственный университет, Томск

e-mail: flo-00@mail.ru

Аннотация. Рассматривается задача оценивания параметров авторегрессионной модели второго порядка с непрерывным временем. Предлагается последовательный план идентификации модели на основе оценок максимального правдоподобия. Для построения плана используется модифицированная выборочная информационная матрица Фишера. Доказывается, что полученные оценки имеют нормальное распределение.

Ключевые слова: авторегрессия с непрерывным временем, метод максимального правдоподобия, нормальное распределени, последовательное оценивание, момент остановки.

Актуальность и история вопроса. Многие процессы в физике, химии, биологии, экономике описываются стохастическими дифференциальными и стохастическими разностными уравнениями. Однако, прежде чем их использовать, необходимо оценить параметры этих уравнений. Поэтому оценивание параметров стохастических

дифференциальных и разностных уравнений — одна из наиболее актуальных задач современного анализа данных.

Для оценивания параметров имеется немало различных методов (метод моментов, метод максимального правдоподобия, метод наименьших квадратов и другие), однако, в случае зависимых наблюдений эти оценки являются нелинейными функциями, что создает сложности в исследовании их свойств.

Один из методов решения задачи оценивания параметров в неасимптотической постановке — это метод последовательного оценивания, предложенный Липцером и Ширяевым [1] для оценки параметра уравнения:

$$dX_t = \theta f_t dt + dW_t.$$

Ими было предложено заменить детерминированный промежуток наблюдений $[0, T]$ на промежуток случайной длины $[0, \tau(H)]$, где

$$\tau(H) = \inf\left\{T > 0: \int_0^T f_s^2 ds = H\right\}.$$

Как оказалось, предложенная в [1] оценка обладает очень хорошими свойствами. Она является нормальной, несмещенной и ее среднеквадратическая погрешность не превышает $\frac{1}{H}$.

Однако, в случае, когда количество параметров превышает размерность процесса, возникают сложности с построением такого же красивого аналога. Решению этой

проблемы посвящено немало работ. Например, в этом направлении были сделаны следующие продвижения:

- Коневу и Пергаменщикову удалось построить двухшаговую последовательную процедуру (без учета априорной информации об области параметров) для моделей авторегрессионного типа [2];
- Коневу и Емельяновой удалось построить более простую процедуру (одноэтапную), но с учетом априорной информации [3].

Цель нашей работы — построить последовательную оценку с нормальным распределением для двухпараметрической авторегрессионной модели с непрерывным временем.

Постановка задачи. Построение последовательной оценки. Рассмотрим задачу оценивания параметров авторегрессионной модели второго порядка по наблюдениям процесса X_t:

$$d\dot{X}_t = \left(\theta_1 \dot{X}_t + \theta_2 X_t\right)dt + dW_t, \qquad (1)$$

или в векторном виде:

$$d\dot{X}_t = (\theta_1 \quad \theta_2)\begin{pmatrix}\dot{X}_t \\ X_t\end{pmatrix} dt + dW_t,$$

где $\theta = (\theta_1, \theta_2)'$ — вектор-столбец неизвестных параметров, штрих обозначает транспонирование, W_t — винеровский процесс.

Оценка максимального правдоподобия вектора θ по наблюдениям процесса (1) на промежутке $[0, T]$ имеет вид

$$\theta_T^* = M_T^+ \int_0^T \left(\dot{X}_t, X_t\right)' dX_t, \qquad (2)$$

где

$$M_T = \begin{pmatrix} \int_0^T \dot{X}_s^2 ds & \int_0^T \dot{X}_s X_s ds \\ \int_0^T \dot{X}_s X_s ds & \int_0^T X_s^2 ds \end{pmatrix}$$

— выборочная информационная матрица Фишера, M_T^+ — обратная к M_T матрица, если она невырождена, и $M_T^+ = 0$ — в противном случае.

Задача — построить последовательную оценку максимального правдоподобия, имеющую неасимптотическое нормальное распределение для любых возможных параметров.

Введем систему моментов остановки и произведем некоторую модификацию выборочной информационной матрицы Фишера.

Для произвольного порога $h > 0$ введем моменты остановки $\tau_i(h), i = 0,1,2$ по следующим формулам:

$$\tau_0(h) = 0,$$

$$\tau_1 = \inf\left\{T > 0: \int_0^T \dot{X}_s^2 ds = h\right\}, \qquad (3)$$

$$\tau_2(h) = \inf\left\{T > \tau_1(h): \int\limits_{\tau_1(h)}^{T} X_s^2 ds = h\right\}.$$

Пусть

$$m(h) = \begin{pmatrix} \int\limits_0^{\tau_1(h)} \dot{X}_s^2 ds & \int\limits_0^{\tau_1(h)} \dot{X}_s X_s ds \\ \int\limits_{\tau_1(h)}^{\tau_2(h)} \dot{X}_s X_s ds & \int\limits_{\tau_1(h)}^{\tau_2(h)} X_s^2 ds \end{pmatrix} \qquad (4)$$

— модифицированная выборочная информационная матрица Фишера.

Определим вектор $V(h) = \left(V_1(h), V_2(h)\right)'$ по формулам

$$V_1(h) = \int\limits_0^{\tau_1(h)} \dot{X}_s dX_s\,,$$

$$V_2(h) = \int\limits_{\tau_1(h)}^{\tau_2(h)} X_s dX_s\,. \qquad (5)$$

Определим последовательную оценку вектора неизвестных параметров θ равенством:

$$\theta^* = m^{-1}(h)V(h) \qquad (6)$$

или подробнее

$$\begin{pmatrix}\theta_1^*\\ \theta_2^*\end{pmatrix} = \begin{pmatrix} \int\limits_0^{\tau_1(h)} \dot{X}_s^2 ds & \int\limits_0^{\tau_1(h)} \dot{X}_s X_s ds \\ \int\limits_{\tau_1(h)}^{\tau_2(h)} \dot{X}_s X_s ds & \int\limits_{\tau_1(h)}^{\tau_2(h)} X_s^2 ds \end{pmatrix}^{-1} \begin{pmatrix} \int\limits_0^{\tau_1(h)} \dot{X}_s dX_s \\ \int\limits_{\tau_1(h)}^{\tau_2(h)} X_s dX_s \end{pmatrix}.$$

Основной результат

Теорема 1. *Пусть задан процесс*

$$d\dot{X}_t = \left(\theta_1 \dot{X}_t + \theta_2 X_t\right)dt + dW_t.$$

Пусть оценки вектора неизвестных параметров $\theta = (\theta_1, \theta_2)'$ *определяются формулами* (3)— (6). *Тогда для любого* $\theta \in \mathbb{R}^2$ *и* $h > 0$

$$\frac{m(h)}{\sqrt{h}}(\theta^*(h) - \theta) \in N(0, I_2),$$

где I_2 *— единичная матрица размера* 2×2.

Доказательство: Для доказательства нормальности вектора θ^* достаточно показать нормальность любой линейной комбинации его компонент. [4]

Представим элементы вектора V в следующем виде:

$$V_1(h) = \int\limits_0^{\tau_1(h)} \dot{X}_s dX_s = \int\limits_0^{\tau_1(h)} \dot{X}_s \left(\left(\theta_1 \dot{X}_s + \theta_1 X_s\right)ds + dW_s\right) =$$

$$= \int\limits_0^{\tau_1(h)} \left(\theta_1 \dot{X}_s{}^2 + \theta_1 \dot{X}_s X_s\right) ds + \int\limits_0^{\tau_1(h)} \dot{X}_s dW_s = \langle m(h)\theta\rangle_i + \eta_i(h),$$

аналогично

$$V_2(h) = \langle m(h)\theta\rangle_2 + \eta_2(h),$$

где $\langle m(h)\theta\rangle_i$ — i-ая строка матрицы $m(h)$, умноженная на вектор θ,

$$\eta_1(h) = \int\limits_0^{\tau_1(h)} \dot{X}_s dW_s\,, \qquad \eta_2(h) = \int\limits_{\tau_1(h)}^{\tau_2(h)} X_s dW_s.$$

Обозначим $\eta(h) = \big(\eta_1(h), \eta_2(h)\big)'$, тогда

$$V(h) = m(h)\theta + \eta(h),$$

$$\eta(h) = m(h)(\theta^*(h) - \theta).$$

Докажем, что вектор $\frac{\eta(h)}{\sqrt{h}}$ имеет двумерное нормальное распределение с нулевым средним и единичной ковариационной матрицей $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Для этого достаточно показать, что для любого $\lambda \in \mathbb{R}^2$ характеристическая функция случайной величины

$$\zeta(h) = \frac{\lambda'\eta(h)}{\sqrt{h}} = (\lambda_1 \quad \lambda_2)\begin{pmatrix} \dfrac{\eta_1(h)}{\sqrt{h}} \\ \dfrac{\eta_2(h)}{\sqrt{h}} \end{pmatrix} = \frac{\lambda_1\eta_1(h)}{\sqrt{h}} + \frac{\lambda_2\eta_2(h)}{\sqrt{h}}$$

имеет вид

$$\varphi_{\zeta(h)}(u) = \mathrm{E}e^{iu\zeta(h)} = e^{-\frac{\lambda'\lambda u^2}{2}}.$$

Вычисляя повторные условные математические ожидания и учитывая, что

$$\mathrm{E}\left(e^{iu\frac{\lambda_j\eta_j(h)}{\sqrt{h}}}/\mathrm{F}_{\tau_{j-1}}\right)=e^{-\frac{\lambda_j^2u^2}{2}}, \qquad [1]$$

получим

$$\mathrm{E}e^{iu\zeta(h)}=\mathrm{E}e^{iu\left(\frac{\lambda_1\eta_1(h)}{\sqrt{h}}+\frac{\lambda_2\eta_2(h)}{\sqrt{h}}\right)}=\mathrm{E}\left(e^{iu\frac{\lambda_1\eta_1(h)}{\sqrt{h}}}\cdot e^{iu\frac{\lambda_2\eta_2(h)}{\sqrt{h}}}\right)=$$

$$=\mathrm{E}\left(\mathrm{E}\left(e^{iu\frac{\lambda_1\eta_1(h)}{\sqrt{h}}}/\,\mathrm{F}_0\right)\right)\cdot\mathrm{E}\left(\mathrm{E}\left(e^{iu\frac{\lambda_2\eta_2(h)}{\sqrt{h}}}/\,\mathrm{F}_{\tau_1}\right)\right)=$$

$$=\mathrm{E}e^{-\frac{\lambda_1^2u^2}{2}}\cdot\mathrm{E}e^{-\frac{\lambda_2^2u^2}{2}}=e^{-\left(\frac{\lambda_1^2u^2}{2}+\frac{\lambda_2^2u^2}{2}\right)}=e^{-(\lambda_1\ \ \lambda_2)\binom{\lambda_1}{\lambda_2}\frac{u^2}{2}}=e^{-\frac{\lambda'\lambda u^2}{2}}.$$

Итак, в работе построена последовательная оценка максимального правдоподобия с нормальным распределением для любых возможных параметров двухпараметрической авторегрессионной модели с непрерывным временем.

Список литературы

[1] Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. – М.: Наука, Главная редакция физико-математической литературы. - 1974. – С.696.

[2] Конев, В.В. Гарантированное оценивание периодического сигнала на фоне авторегрессионных помех с неизвестными параметрами / В.В. Конев, С. М. Пергаменщиков

// Проблемы передачи информации. – 1997. – Т. 33, вып. 4. – С.26—44.

[3] Емельянова Т.В., Конев В.В. О последовательном оценивании параметров непрерывной авторегрессии // Вестник Томского госуниверситета. Математика и механика. - 2013. № 5(25). - С.12—25.

[4] Ширяев А.Н. Вероятность. – М.: Наука, Главная редакция физико-математической литературы. - 1980. – С.576.

[5] Булинский А.В., Ширяев А.Н. Теория случайных процессов. – М.: ФИЗМАЛИТ; Лаборатория Базовых Знаний. - 2003. – С.400.