УДК 519.216.3

ОБ ОЦЕНИВАНИИ ПАРАМЕТРОВ ТРИГОНОМЕТРИЧЕСКОГО СИГНАЛА С АДДИТИВНЫМИ ШУМАМИ

© Т. В. Емельянова, А. О. Шерстобитова*, А. А. Конищева

Национальный исследовательский Томский государственный университет
Россия, 634050 г. Томск, пр. Ленина, 36.

Тел.: +7 (953) 920 62 94.

*Email: annasherstobitova06@gmail.com

Рассматривается задача оценивания коэффициентов тригонометрической регрессионной модели с дискретным временем по зашумленным наблюдениям на фоне зависимых шумов авторегрессионного типа с неизвестными параметрами и распределением. Построена последовательная процедура оценивания коэффициентов сигнала с использованием специального правила остановки наблюдений. Получены результаты, характеризующие среднюю длительность последовательной процедуры и качество последовательных оценок параметров.

Ключевые слова: последовательное оценивание, момент остановки, матрица Фишера, тригонометрическая регрессия.

Введение

Оценивание параметров сигналов является актуальной задачей для построения моделей в задачах автоматического управления, фильтрации, прогнозирования и идентификации. К настоящему времени разработано несколько методов оценки параметров сигналов с дискретным и непрерывным временем на фоне аддитивного шума с заранее известной информацией о типе сигнала и форме шума [1; 4; 6]. Наиболее полно задача выделения сигналов исследована в случае дискретного времени при условии, что помехи являются последовательностью независимых случайных величин. В случае невыполнения условия, может возникнуть проблема, касающаяся наличия дополнительных неизвестных параметров шума, что в свою очередь значительно усложняет задачу расчета точности оценок параметров сигнала. Таким образом, задача идентификации параметров сигнала при шумах с неизвестными спектральными свойствами менее изучена.

В данной работе рассматривается задача оценивания коэффициентов тригонометрического сигнала с зависимыми шумами в случае неизвестной дисперсии. С помощью специального правила остановки наблюдений осуществляется построение последовательного плана $\{\tau(\tilde{h}), \alpha^*(\tilde{h})\}$. В работе представлены результаты, характеризующие среднюю длительность последовательного плана и точность последовательных оценок неизвестных параметров.

Постановка задачи. Построение последовательной процедуры

Рассматривается задача идентификации параметров $a_1, a_2, c_{j1}, c_{j2}, j = 1, \dots, r$ тригонометрического сигнала с дискретным временем

$$S_n = a_1 + (-1)^n a_2 + \sum_{j=1}^r c_{j1} \cos \omega_j n + c_{j2} \sin \omega_j n \quad (1)$$

по наблюдениям

$$x_n = S_n + \xi_n \quad (2)$$

где ξ_n – шум авторегрессионного типа порядка p с неизвестными параметрами

$$\xi_n = \lambda_1 \xi_{n-1} + \dots + \lambda_p \xi_{n-p} + \varepsilon_n, \quad (3)$$

где $\{\varepsilon_n\}$ – независимые одинаково распределенные величины с нулевым математическим ожиданием, конечной дисперсией, причем $\lambda_1, \dots, \lambda_p$ – неизвестные параметры такие, что все корни характеристического полинома $P(q) = q^p - \lambda_1 q^{p-1} - \dots - \lambda_p$ лежат внутри единичного круга комплексной плоскости. $\omega_j, j = 1, \dots, r$, предполагается, что $0 < \omega_j < \pi, \omega_i \neq \omega_j$ при $i \neq j$.

С учетом (1) и (3), наблюдаемый процесс (2) удовлетворяет уравнению

$$x_n = m_1 + (-1)^n m_2 + \sum_{j=1}^r (\gamma_{j1} \cos \omega_j n + \gamma_{j2} \sin \omega_j n) + \sum_{k=1}^p \lambda_k x_{n-k} + \sigma^2 \varepsilon_n, \quad (4)$$

$$n \geq p + 1$$

Здесь $m_1, m_2, \gamma_{ji}, \lambda_k, k = 1, \dots, p$ – неизвестные параметры, связанные с параметрами сигнала равенствами

$$m_1 = a_1 \left(1 - \sum_{l=1}^p \lambda_l \right), m_2 = a_2 \left(1 - \sum_{l=1}^p (-1)^l \lambda_l \right),$$

$$\gamma_{j1} = c_{j1} \left(1 - \sum_{l=1}^p \lambda_l \cos \omega_j l \right) + c_{j2} \sum_{l=1}^p \lambda_l \sin \omega_j l,$$

$$\gamma_{j2} = -c_{j1} \sum_{l=1}^p \lambda_l \sin \omega_j l + c_{j2} \left(1 - \sum_{l=1}^p \lambda_l \cos \omega_j l \right).$$

Используя обозначения

$$Y_n = \begin{pmatrix} \Phi_n \\ X_{n-1} \end{pmatrix}, X_n = \begin{pmatrix} x_n \\ \dots \\ x_{n-p+1} \end{pmatrix}, \Phi_n = \begin{pmatrix} \phi_1(n) \\ \dots \\ \phi_{2r+2}(n) \end{pmatrix},$$

$$\phi_1(n) = 1, \phi_2(n) = (-1)^n,$$

$$\phi_k(n) = \cos \omega_{k-2} n \text{ при } 3 \leq k \leq r+2,$$

$$\phi_k(n) = \sin \omega_{k-r-2} \text{ при } r+3 \leq k \leq 2r+2,$$

уравнение (4) записывается в векторной форме

$$X_n = \alpha' Y_n + \sigma^2 \varepsilon_n, n \geq p+1, \quad (5)$$

где $\alpha = (m_1, m_2, \gamma_{11}, \gamma_{r1}, \gamma_{12}, \dots, \gamma_{r2})$ – вектор оцениваемых параметров.

В работе [3] построена последовательная процедура оценивания параметров для (1)–(3) в случае известной дисперсии шума. Цель данной работы заключается в использовании метода оценивания параметров тригонометрического сигнала при условии, что дисперсия шума σ^2 неизвестна. В этом случае σ^2 рассматривается как неизвестный параметр, влияние которого на качество оценки α должно быть по возможности нейтрализовано. В работах, посвященных исследованию аналогичных задач [2; 10], для этой цели рассматривается оценивание неизвестной дисперсии по выборке фиксированного размера. Этой концепции будем придерживаться далее в работе. Пусть $(x_1, \dots, x_{m_0})$ – выборка фиксированного размера. Оценка по МНК вектора параметров α имеет следующий вид

$$\alpha(n) = M_n^{-1} \sum_{k=m_0+p+1}^n Y_{k-1} x_k,$$

причем $M_n = \sum_{k=m_0+p+1}^n Y_{k-1} Y_{k-1}'$ – выборочная информационная матрица Фишера размера $l \times l$, $l = 2r+2+p$.

Введем случайный порог

$$\tilde{h} = h \cdot R_{m_0}^2,$$

где h – положительное число, $R_{m_0}^2$ – некоторая выборочная борелевская функция, определяемая следующим образом

$$R_{m_0}^2 = \frac{1}{m_0} \sum_{i=1}^{m_0} x_i^2.$$

Определим момент остановки $\tau(\tilde{h})$ и последовательную оценку МНК $\alpha^*(\tilde{h})$ вектора параметров α

$$\tau(\tilde{h}) = \inf \{n \geq n_0 : \|M_n^{-2}\| \leq \frac{1}{\tilde{h}}, \quad (6)$$

$$\alpha^*(\tilde{h}) = \tilde{M}_{\tau(\tilde{h})}^{-1} \sum_{k=m_0+p+1}^{\tau(\tilde{h})} Y_{k-1} x_k, \quad (7)$$

где $n_0 = \inf \{n > m_0 + p : \det M_n > 0\}$, $\tilde{M}_{\tau(\tilde{h})} =$

$$\sum_{k=m_0+p+1}^{\tau(\tilde{h})} Y_{k-1} Y_{k-1}'.$$

Свойства последовательного плана оценивания

Теорема 1 (о средней длительности последовательного плана). Пусть в модели (2)–(3) последовательность независимых одинаково распределенных случайных величин $(\varepsilon_n)_{n \geq 1}$ удовлетворяет условиям

$$\sup_{\alpha \in \Lambda} E_{\alpha} \|X_0\|^8 < \infty, E \varepsilon_n^8 < +\infty.$$

Пусть Λ – множество допустимых значений вектора параметров α . Пусть (6)–(7) – последовательный план для оценивания параметров модели (2)–(3). Тогда для произвольного компактного множества $K \subset \Lambda$ и статистики $R_{m_0}^2$ такой, что выполняется условие

$$\sup_{\alpha \in \Lambda} E_{\alpha} R_{m_0}^{-4} < +\infty,$$

средняя длительность последовательной процедуры удовлетворяет соотношению

$$\lim_{h \rightarrow 0} \sup_{\alpha \in K} \left| E_{\alpha} \frac{\tau(\tilde{h})}{h} - \|F^{-2}\|^{1/2} E_{\alpha} R_{m_0}^2 \right| = 0.$$

Доказательство. В доказательстве теоремы 1 будут использованы свойства выборочной информационной матрицы Фишера, установленные в работе [5]

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=p+1}^N Y_n Y_n' = F,$$

F – положительно определенная матрица

$$F = \begin{pmatrix} M_0 & M_1 \\ M_1' & F_0 + D M_0 D' \end{pmatrix}, \quad (8)$$

где

$$M_0 = \text{diag} \left(1, 1; \frac{1}{2}, \dots, \frac{1}{2} \right),$$

$$V(k) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (-1)^k & 0 & 0 \\ 0 & 0 & V_1(k) & V_2(k) \\ 0 & 0 & -V_2(k) & V_1(k) \end{pmatrix},$$

причем $V_1(k) = \text{diag}(\cos \omega_1 k, \dots, \cos \omega_r k)$, $V_2(k) = \text{diag}(\sin \omega_1 k, \dots, \sin \omega_r k)$,

$$A = \begin{pmatrix} \lambda_1 & \dots & \lambda_p \\ I_{p-1} & \dots & 0 \end{pmatrix},$$

$$\Gamma = \begin{pmatrix} m_1 & m_2 & \gamma_{11} & \gamma_{12} & \dots & \gamma_{r1} & \gamma_{r2} \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix},$$

$$D = \sum_{k \geq 0} A^k \Gamma V(k), M_1 = M_0 V'(1) D',$$

$$F_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=p+1}^N \zeta_n \zeta_n' \text{ п.н.},$$

$$\zeta = A \zeta_{n-1} + \eta_n, \zeta_p = 0, \eta_n = (\varepsilon_n, 0, \dots, 0)'.$$

Оценим скорость сходимости матрицы $\frac{M_N}{N}$ к предельной матрице F , определенной в (8).

Лемма 1. Пусть в модели (2)–(3) для последовательности случайных величин $(\varepsilon_n)_{n \geq 1}$ выполняются условия $E\varepsilon_n^8 < \infty$ и $\sup_{\alpha \in \Lambda} E_\alpha \|X_0\|^8 \leq \infty$. Тогда для произвольного компактного множества $K \subset \Lambda$

$$\sup_{\alpha \in K} E_\alpha \left(\left\| \frac{M_N}{N} - F \right\|^4 | \Phi_{m_0} \right) \leq \frac{L}{N^2} (\|X_{m_0}\|^8 + 1).$$

Пользуясь леммой 1 и схемой доказательства теоремы 3.1 из [8], получим

$$\lim_{h \rightarrow +\infty} \sup_{\alpha \in K} E_\alpha \frac{\tau(h)}{h} < +\infty.$$

Таким образом, имеем оценку

$$\begin{aligned} & \left| E_\alpha \frac{\tau(h)}{h} - \|F^{-2}\|^{\frac{1}{2}} \right| = \\ & = \left| E_\alpha \left\| \left(\frac{M_{\tau(h)}}{\tau(h)} \right)^{-2} \right\|^{\frac{1}{2}} - \|F^{-2}\|^{\frac{1}{2}} \right| \leq \\ & \leq \Delta + E_\alpha \left[\frac{\tau(h)}{h} \chi \left(\left\| \frac{M_{\tau(h)}}{\tau(h)} - F \right\| \geq \eta \right) \right] + \\ & + \|F^{-2}\|^{\frac{1}{2}} P_\alpha \left(\left\| \frac{M_{\tau(h)}}{\tau(h)} - F \right\| \geq \eta \right) \end{aligned} \quad (27)$$

Здесь Δ – некоторая положительная постоянная.

Воспользуемся результатом из работы [8, фл. 3.25].

Лемма 2. В условиях леммы 1 для любого компакта $K \subset \Lambda$ и произвольного η среднеквадратическая точность, определяющая план последовательной оценки (6)–(7) удовлетворяет неравенству

$$\lim_{h \rightarrow +\infty} \sup_{\alpha \in K} P_\alpha \left(\left\| \frac{M_{\tau(h)}}{\tau(h)} - F \right\| \geq \eta \right) = 0.$$

Используя эту оценку и устремляя $h \rightarrow +\infty$, правая часть неравенства (27) стремится к нулю.

Теорема 1 доказана.

Следующая теорема относится к точности последовательных оценок неизвестных параметров.

Теорема 2. Для произвольного компакта $K \subset \Lambda$ среднеквадратическая точность последовательной оценки (6)–(7) удовлетворяет неравенству

$$\sup_{\alpha \in K} E_\alpha (\|\alpha^*(h) - \alpha\|^2) \leq \frac{a_k}{h} (1 + o(1)),$$

где $a_k = \sup_{\alpha \in K} U(\alpha) \|F^{-2}\|^{\frac{1}{2}}$, $o(1) \rightarrow 0$ при $h \rightarrow \infty$, K – произвольный компакт.

Доказательство. Оценка $\alpha^*(h)$, определенная в (7), для любого $h > 0$ удовлетворяет неравенству

$$E_\alpha \|\alpha^*(h) - \alpha\|^2 \leq \frac{E_\alpha \text{tr} M_{\tau(h)}}{h^2}. \quad (28)$$

Для проверки этого неравенства оценку уклонения оценки МНК в модели типа (5) [8, с. 338].

$$\|\alpha(n) - \alpha\|^2 \leq \|M_n^{-2}\| \cdot \|m_n\|^2, m_n = \sum_{k=1}^n Y_k \varepsilon_k,$$

применим для уклонения оценки $\alpha^*(h)$.

Лемма 3. Пусть $(\varepsilon_n)_{n \geq 1}$ – н.о.р. случайные величины, $E\varepsilon_n = 0$, $E\varepsilon_n^2 = \sigma^2 < +\infty$. Тогда для любого $h > 0$

$$E_\alpha \text{tr} \tilde{M}_\tau \leq U(\alpha) E_\alpha \tau + m \sqrt{E_\alpha \tau} + Z,$$

где

$$\begin{aligned} U(\alpha) &= \text{tr} F_0 + (2r + 2) + (2r + 2)^2 \left(\|D\|^2 + \frac{c}{1-q} \right), \\ m &= \frac{c\sigma}{1-q} (2r + 2) \cdot (\|D\| + \|\Gamma\|), Z = \\ &= \frac{c}{1-q} ((2r + 2) \cdot \|D\| + 2) E_\alpha \|X_p\|. \end{aligned}$$

Подставляя результат леммы 3 в (28) и применяя теорему 1, получаем утверждение теоремы 2.

Таким образом, теорема 2 доказана.

Результаты численного моделирования

Для тригонометрического сигнала

$$S_n = a_1 \cos \frac{\pi n}{6} + a_2 \sin \frac{\pi n}{3}$$

рассматривает задача оценивания параметров a_1, a_2 по наблюдениям

$$X_n = S_n + \xi_n,$$

в котором ξ_n – аддитивный шум, описываемый уравнением

$$\xi_n = \lambda \xi_{n-1} + \sigma^2 \varepsilon_n.$$

Здесь $\{\varepsilon_n\}$ – последовательность независимых одинаково распределенных случайных величин, такая что $E\varepsilon_n = 0$, $E\varepsilon_n^2 = \sigma^2$, λ – неизвестный параметр, $|\lambda| < 1$.

Предельная матрица F имеет вид

$$F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{a_1 \sqrt{3}}{4} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & -\frac{a_2 \sqrt{3}}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{a_1}{4} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{a_2}{4} \\ 0 & 0 & \frac{a_1 \sqrt{3}}{4} & -\frac{a_2 \sqrt{3}}{4} & \frac{a_1}{4} & \frac{a_2}{4} & \frac{a_1^2 + a_2^2}{2} + \frac{1}{1-\lambda^2} \end{pmatrix}$$

В табл. 1 приведены результаты вычислений, полученные при значении условия моделирования процесса: $a_1 = 0.2, a_2 = 0.4, \lambda = 0.5$. Здесь h – величина порога, $\tau(h)$ – выборочная средняя длительность процедуры, $\tau(h)/h$ – удельная средняя длительность процедуры, $\|F^{-2}\|^{\frac{1}{2}}$ – асимптотическое значение.

Таблица 1

Результаты вычислений при использовании последовательной процедуры

h	25	50	100	250	500	1000
$\tau(h)$	142	191	326	622	1105	1733
$\tau(h)/h$	5.68	3.82	3.26	2.49	2.21	1.73
$\ F^{-2}\ ^{\frac{1}{2}}$	2.94	2.94	2.94	2.94	2.94	2.94

Рис. 1 иллюстрирует построенный процесс X_n , а также модель, построенную с помощью полученных оценок параметров. Здесь сплошная линия отражает оригинальные данные (построенный процесс), пунктирная линия отражает модельные данные.

Рис. 1. Сравнение данных процесса и полученных модельных данных.

Заключение

В работе предложен одноэтапный последовательный план оценивания параметров тригонометрической регрессии по зашумленным наблюдениям на фоне зависимых шумов авторегрессионного типа с неизвестными параметрами и распределением. Проведено численное моделирование методом Монте-Карло для оценивания параметров рассматриваемой модели. С помощью выборочной информационной матрицы Фишера определялось специальное правило прекращения наблюдений, что позволило контролировать среднеквадратическую точность оцениваемых параметров.

ЛИТЕРАТУРА

1. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976. 756 с.
2. Васильев В. А. Последовательное оценивание параметров динамических систем при неполном наблюдении // Известия АН СССР. Техническая кибернетика. 1982. №6. С. 145–154.
3. Емельянова Т. В. О последовательном оценивании периодического сигнала на фоне авторегрессионного шума // Вестник ТГУ. Математика и механика. 2015. №2(34). С. 18–29.
4. Ибрагимов И. А. Асимптотическая теория оценивания. М.: Наука, 1979. 527 с.
5. Конев В. В. Гарантированное оценивание периодического сигнала на фоне авторегрессионных помех с неизвестными параметрами // Проблемы передачи информации. 1997. Т. 33. Вып. 4. С. 26–44.
6. Липцер Р. Ш. Статистика случайных процессов. М.: Наука, 1974. 696 с.
7. Ширяев А. Н. Вероятность. М.: Наука, 1989. 576 с.
8. Galtchouk L. On sequential least squares estimates of autoregressive parameters // Sequential Analysis. 2005. Vol. 24, is. 4. P. 335–364.
9. Konev V. On guaranteed estimation of the mean of an autoregressive process // The Annals of Statistics. 1997. Vol. 20, is. 5. P. 2127–2163.
10. Stein C. A two-sample test for a linear hypothesis whose power is independent of the variance // The Annals of Mathematical Statistics. 1945. Vol. 16, is. 3. P. 243–258.

Поступила в редакцию 12.03.2019 г.

ON THE SEQUENTIAL PARAMETER ESTIMATION OF THE TRIGONOMETRIC SIGNAL WITH ADDITIVE NOISE

© T. V. Emelyanova, A. O. Sherstobitova*, A. A. Konischeva

*Tomsk State University
36 Lenin Avenue, 634050 Tomsk, Russia.*

Phone: +7 (953) 920 62 94.

**Email: annasherstobitova06@gmail.com*

A problem of estimation of signal parameters is an important task for building models in such spheres as automatic control, filtering, forecasting and identification. The problem of signal selection is studied in more detail in the case of discrete time, where the noise is a sequence of independent random variables. In case of non-fulfillment of the condition, there can be a problem concerning the presence of additional unknown noise parameters, which in turn significantly complicates the task of calculating the accuracy of the signal parameters estimates. The authors of the article studied the problem of trigonometric signal with estimation of discrete time coefficients. The signal is founded on observations with additive noise described by the stationary process of autoregression with unknown parameters and distribution. There were constructed a sequential procedure for estimating the coefficients of the signal using special rules of stop observations. The article presents the results of the study characterizing the average duration of the sequential plan and the accuracy of sequential estimates of unknown parameters.

Keywords: sequential parameter estimation, stopping moment, Fisher matrix, trigonometric regression model.

Published in Russian. Do not hesitate to contact us at bulletin_bsu@mail.ru if you need translation of the article.

REFERENCES

1. Anderson T. Statisticheskii analiz vremennykh ryadov [Statistical time series analysis]. Moscow: Mir, 1976.
2. Vasil'ev V. A. Izvestiya AN SSSR. Tekhnicheskaya kibernetika. 1982. No. 6. Pp. 145–154.
3. Emel'yanova T. V. Vestnik TGU. Matematika i mekhanika. 2015. No. 2(34). Pp. 18–29.
4. Ibragimov I. A. Asimptoticheskaya teoriya otsenivaniya [Asymptotic estimation theory]. Moscow: Nauka, 1979.
5. Konev V. V. Problemy peredachi informatsii. 1997. Vol. 33. No. 4. Pp. 26–44.
6. Liptser R. Sh. Statistika sluchainykh protsessov [Statistics of random processes]. Moscow: Nauka, 1974.
7. Shiryaev A. N. Veroyatnost' [Probability]. Moscow: Nauka, 1989.
8. Galtchouk L. Sequential Analysis. 2005. Vol. 24, is. 4. Pp. 335–364.
9. Konev V. The Annals of Statistics. 1997. Vol. 20, is. 5. Pp. 2127–2163.
10. Stein C. The Annals of Mathematical Statistics. 1945. Vol. 16, is. 3. Pp. 243–258.

Received 12.03.2019.