УДК 519.2

М.С. АБРАМОВИЧ

Оценка мощности спектрального критерия обнаружения момента «разладки» временных рядов

Для обнаружения моментов «разладки» временных рядов широко применяются методы, основанные на модели авторегрессии и скользящего среднего [1], и метод кумулятивных сумм [2]. В меньшей мере для решения этой проблемы исследовано применение методов спектрального анализа [3].

В [4] построен критерий для обнаружения момента «разладки» с использованием оценок спектральных плотностей, найдено асимптотическое распределение статистики критерия и построен тест для обнаружения момента «разладки». В настоящей работе исследуется мощность этого теста.

1. Математическая модель. Пусть ξ_t, $t \in Z$ - стационарный временной ряд с нулевым средним и некоторой спектральной плотностью $S_1^0(\lambda)$, $\lambda \in [-\pi, \pi]$, $\eta_t, t \in Z$ - стационарный временной ряд с нулевым средним и некоторой другой спектральной плотностью $S_2^0(\lambda)$, $\lambda \in [-\pi, \pi]$, отличной от $S_1^0(\cdot) : S_1^0(\cdot) \neq S_2^0(\cdot)$. Наблюдаемый временной ряд $X = \{x_t : t = 1,2,\ldots,T\}$ длительностью T имеет «разладку» в неизвестный момент времени $t_0 \in \{\tau_-, \tau_- + 1, \ldots, \tau_+, T+1\}$, которая порождается «скачкообразным изменением» спектральной плотности:

$$x_t = \begin{cases} \xi_\tau, t \in \{1,2,\ldots,t_0 - 1\}, \\ \eta_\tau, t \in \{t_0, t_0 + 1, \ldots, T\}; \end{cases}$$

$1 < \tau_- < \tau_+ < T$ - некоторые априорно заданные граничные значения. Если $t_0 = T+1$, то наблюдаемый временной ряд - однородный.

Задача обнаружения «разладки» заключается в оценивании момента «разладки» t_0. Для $\tau \in \{\tau_-, \tau_- + 1, \ldots, \tau_+\}$ осуществим разбиение наблюдаемого временного ряда X на два фрагмента $X_1 = (x_1, \ldots, x_{\tau-1}), X_2 = (x_\tau, \ldots, x_T)$ с длинами T_1 и T_2 соответственно:

$$X = (X_1 \vdots X_2),\ T_1 = \tau - 1,\ T_2 = T - \tau + 1, T_1 + T_2 = T.$$

Определим гипотезу $H_{1\tau} : t_0 = \tau$, означающую, что наблюдаемый временной ряд имеет «разладку» в момент τ, и гипотезу $H_{0\tau} : t_0 = T+1$, означающую, что наблюдаемый временной ряд - однородный.

Оценка спектральной плотности по j-му фрагменту запишется в виде [6]:

$$\hat{S}_j(f_i) = \frac{1}{2\pi} \sum_{k=-(T_j-1)}^{T_j-1} W^*_{kT_j} \hat{R}_j(k) \cos(f_i k),\ j = 1,2, \tag{1}$$

где $i \in \{1, \ldots, m\}$ — номер частоты, m- число частот $f_1, \ldots, f_m$, для которых вычисляются оценки спектральной плотности, $\hat{R}_j(k)$ - значение выборочной ковариационной функции для лага k:

$$\hat{R}_j(k) = \frac{1}{T_j} \sum_{t=T_{j-1}+1}^{T_{j-1}+T_j-k} x_t x_{t+k} \quad (T_0 ::= 0).$$

$$W^*_{k T_j} = \begin{cases} r\left(\frac{k}{K_{T_j}}\right), k = 0,1, \ldots, K_{T_j}, \\ 0, k = \left(K_{T_j} + 1\right), \ldots, T_j - 1, \end{cases} \tag{2}$$

$r(x)$ - весовая функция, $K_{T_j} \in N$ - параметр сглаживания, $K_{T_j} < T_j$. В соответствии с [6] введем параметр h:

$$h = \lim_{x \to 0} \frac{1 - r(x)}{|x|^q}, q > 0. \tag{3}$$

Обозначим:

$$U = (u_k) \in R^{2m}, u_k = \hat{S}_1(f_k), u_{m+k} = \hat{S}_2(f_k),$$

$$U^0 = (u_k^0) \in R^{2m}, u_k^0 = S_1^0(f_k), u_{m+k}^0 = S_2^0(f_k), k = \overline{1, m},$$

$$Q = \int_{-1}^{1} r^2(x) dx.$$

Определим статистику для обнаружения момента «разладки» как меру различия выборочных спектральных плотностей $\hat{S}_1(\cdot)$ и $\hat{S}_2(\cdot)$:

$$G = \frac{\sum_{k=1}^{m} \left(u_{m+k} - u_k\right)^2}{\sum_{k=1}^{m} \left(u_{m+k}^2 + u_k^2\right)}. \tag{4}$$

В [4] построен спектральный критерий обнаружения момента «разладки» в момент τ:

$$\text{принимается:} \begin{cases} H_{0\tau}, \text{если} \quad G < \Delta, \\ H_{1\tau}, \text{если} \quad G \geq \Delta, \end{cases} \tag{5}$$

где $\Delta = \Delta(\varepsilon)$ - пороговое значение, соответствующее заданному уровню значимости $\varepsilon \in (0,1)$.

2. Асимптотический анализ мощности критерия. Для исследования мощности спектрального критерия обнаружения «разладки» (4), (5) найдем асимптотическое при $T_1, T_2 \to \infty$ распределение вспомогательной статистики

$$\tilde{G} = \frac{\sum_{k=1}^{m}\left(u_{m+k} - u_k\right)^2}{\sum_{k=1}^{m}\left(\left(u_{m+k}^0\right)^2 + \left(u_k^0\right)^2\right)}. \tag{6}$$

Для оценки мощности критерия воспользуемся следующей леммой, доказательство которой приведено в [5].

ЛЕММА. Пусть случайный k-вектор $X_n = \left(X_{n1}, \cdots, X_{nk}\right)$ при $n \to \infty$ асимптотически распределен по нормальному закону $N_k\left(\mu, b_n^2 \Sigma\right)$, где μ-математическое ожидание, Σ-ковариационная матрица и $b_n^2 \to 0$. Пусть $g(x), x = \left(x_1, \ldots, x_k\right)$ — функция, имеющая ненулевой k-вектор частных производных в точке $x = \mu$: $D=\left.\left(\frac{\partial g}{\partial x_i}\right)\right|_{x=\mu}, \left(i = 1, k\right)$. Тогда асимптотическое распределение статистики $Y_n = g\left(X_n\right)$ есть $N_1\left(g(\mu), b_n^2 D^T \Sigma D\right)$.

ТЕОРЕМА. Пусть выполняются следующие условия (j=1,2):

1) $\sum_{k=-\infty}^{\infty}\left|R_j(k)\right| < \infty$, где $R_j(k)$-ковариационная функция временного ряда;

2) $\sum_{k,s,t=-\infty}^{\infty}\left|\kappa_j\left(k,s,t\right)\right| < \infty$, где $\kappa_j\left(k,s,t\right)$-семинвариант четвертого порядка;

3) временные ряды $\left\{\xi_t\right\}$и$\left\{\eta_t\right\}$ представляются в виде:

$$\xi_t = \sum_{k=-\infty}^{\infty}\gamma_k v_{t-k}, \eta_t = \sum_{k=-\infty}^{\infty}\beta_k \tilde{v}_{t-k},$$

где $\sum_{k=-\infty}^{\infty}\left|\gamma_k\right| < \infty, \sum_{k=-\infty}^{\infty}\left|\beta_k\right| < \infty$, а $\left\{v_t\right\}, \left\{\tilde{v}_t\right\}$-последовательности независимых и одинаково распределенных случайных величин с

$E\left\{v_t\right\} = E\left\{\tilde{v}_t\right\} = 0; E\left\{v_t^2\right\} = E\left\{\tilde{v}_t^2\right\} = \sigma^2, E\left\{v_t^4\right\} = E\left\{\tilde{v}_t^4\right\} < \infty$;

4) имеет место асимптотика:

$$T_1, T_2 \to \infty, K_{T_1}, K_{T_2} \to \infty, \text{и} \frac{K_{T_1}}{T_1} \to 0, \frac{K_{T_2}}{T_2} \to 0,$$

$$m = O\left(T^{1-\mu}\right), 0 < \mu < 1, m \to \infty.$$

Тогда при выполнении гипотезы $H_{1\tau}$ статистика $\tilde{G}$ распределена асимптотически нормально:

$$L\left\{\frac{\tilde{G} - a\left(T_1, T_2\right)}{\sqrt{B\left(T_1, T_2\right)}}\right\} \to N_1(0,1),$$

где

$$a\left(T_1,T_2\right)=\frac{\sum\limits_{k=1}^{m}\left(u^0_{m+k}-u^0_k+\frac{1}{K^q_{T_1}}h\left(u^0_k\right)^{[q]}-\frac{1}{K^q_{T_2}}h\left(u^0_{m+k}\right)^{[q]}\right)^2}{\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}{}^2\right)+\left(u^0_k\right)^2\right)}, \tag{7}$$

$$B\left(T_1,T_2\right)=\frac{4Q\left(\sum\limits_{k=1}^{m}\left(u^0_{m+k}-u^0_k\right)^2\left(\left(u^0_k\right)^2\frac{K_{T_1}}{T_1}+\left(u^0_{m+k}\right)^2\frac{K_{T_2}}{T_2}\right)\right)}{\left(\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2+\left(u^0_k\right)^2\right)\right)^2}, \tag{8}$$

ДОКАЗАТЕЛЬСТВО. В соответствии с [6] определим асимптотическое выражение для математического ожидания оценки спектральной плотности

$$E\left\{\hat{S}_j(\lambda)\right\}=S^0_j(\lambda)-\frac{1}{K^q_{T_j}}h\left(S^0_j\right)^{[q]}(\lambda), \tag{9}$$

где

$$\left(S^0_j\right)^{[q]}(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}|k|^q\cos(\lambda k)\,R_j(k),$$

и для дисперсии оценки спектральной плотности

$$D\left\{\hat{S}_j(\lambda)\right\}=\frac{K_{T_j}}{T_j}\left(S^0_j(\lambda)\right)^2 Q. \tag{10}$$

Учитывая, что в силу условий 1), 2) и 4) данной теоремы спектральные оценки $\hat{S}_1(f_1),\ldots,\hat{S}_1(f_m),\hat{S}_2(f_1),\ldots,\hat{S}_2(f_m)$ асимптотически независимы, и принимая во внимание (10), приходим к выводу, что асимптотическая $(2m\times 2m)$- ковариационная матрица оценок спектральных оценок будет иметь диагональный вид:

$$\Sigma=Q\cdot\mathrm{diag}\left\{\frac{K_{T_1}}{T_1}\left(u^0_1\right)^2,\ldots,\frac{K_{T_1}}{T_1}\left(u^0_m\right)^2,\frac{K_{T_2}}{T_2}\left(u^0_{m+1}\right)^2,\ldots,\frac{K_{T_2}}{T_2}\left(u^0_{2m}\right)^2\right\}. \tag{11}$$

Согласно (6), $\tilde{G}$ является функцией от $U=\{u_k\}$. Воспользуемся леммой для вычисления математического ожидания и дисперсии этой статистики.

В силу условий 3) и 4) теоремы, оценки спектральных плотностей асимптотически нормальны. Учитывая (9), значение математического ожидания $a\{T_1,T_2\}$ этого асимптотического распределения равно:

$$a\left(T_1,T_2\right)=\frac{\sum\limits_{k=1}^{m}\left(u^0_{m+k}-u^0_k+\frac{1}{K^q_{T_1}}h\left(u^0_k\right)^{[q]}-\frac{1}{K^q_{T_2}}h\left(u^0_{m+k}\right)^{[q]}\right)^2}{\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2+\left(u^0_k\right)^2\right)}, \tag{12}$$

что совпадает с (7).

Вектор-столбец производных D запишется в виде:

$$D = \left.\frac{\partial \tilde{G}}{\partial u_s}\right|_{U=U^0} = \begin{cases} \dfrac{-2\left(u^0_{m+s} - u^0_s\right)}{\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2 + \left(u^0_k\right)^2\right)}, s = \overline{1,m} \\ \dfrac{2\left(u^0_s - u^0_{s-m}\right)}{\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2 + \left(u^0_k\right)^2\right)}, s = \overline{m+1,2m} \end{cases} \tag{13}$$

В силу условия 4) теоремы, значение параметра $b_n^2 = \dfrac{K_{T_j}}{T_j} \to 0.$ Тогда, учитывая (11) и (13), дисперсия асимптотического распределения $B(T_1,T_2)$ вычисляется следующим образом:

$$B(T_1,T_2) = D\Sigma D^T = \frac{4Q}{\left(\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2+\left(u^0_k\right)^2\right)\right)^2}\left(\sum_{k=1}^{m}\left(u^0_{m+k}-u^0_k\right)^2\left(u^0_k\right)^2\frac{K_{T_1}}{T_1}+\right.$$

$$\left.+\sum_{k=m+1}^{2m}\left(u^0_k - u^0_{k-m}\right)^2\left(u^0_k\right)^2\frac{K_{T_2}}{T_2}\right) =$$

$$= \frac{4Q\left(\sum\limits_{k=1}^{m}\left(u^0_{m+k}-u^0_k\right)^2\left(\left(u^0_k\right)^2\dfrac{K_{T_1}}{T_1}+\left(u^0_{m+k}\right)^2\dfrac{K_{T_2}}{T_2}\right)\right)}{\left(\sum\limits_{k=1}^{m}\left(\left(u^0_{m+k}\right)^2+\left(u^0_k\right)^2\right)\right)^2}, \tag{14}$$

что совпадает с (8).

Теорема доказана.

СЛЕДСТВИЕ. Справедлива следующая аппроксимация мощности теста

$$W = P_{H_{1\tau}}\{G \ge \Delta\} \approx 1 - \Phi\left(\frac{\Delta - a(T_1,T_2)}{\sqrt{B(T_1,T_2)}}\right), \tag{15}$$

где $\Phi(\cdot)$ - функция стандартного нормального распределения.

ДОКАЗАТЕЛЬСТВО. Мощность теста, основанного на статистике $\tilde{G}$, определяется следующим образом:

$$W_1 = P_{H_{1\tau}}\{\tilde{G} \ge \Delta\} \approx 1 - \Phi\left(\frac{\Delta - a(T_1,T_2)}{\sqrt{B(T_1,T_2)}}\right), \tag{16}$$

Статистика $\tilde{G}$ - непрерывное функциональное преобразование от оценок спектральных плотностей (1), которые являются состоятельными[6]. Поэтому для статистики $\tilde{G}$ имеет место сходимость по вероятности $\tilde{G} - G \xrightarrow{P} 0$. Следовательно, (16) можно переписать в виде (15), что и доказывает следствие.

СПИСОК ЛИТЕРАТУРЫ

1.Клигене Н., Телькснис Л. Методы обнаружения моментов из менения свойств случайных процессов.// Автоматика и телемеханика. 1983. N 10. с.5.

2.Никифоров И.В. Последовательное обнаружение изменения свойств временных рядов. М.,1983.

3.Giraitis L., Leipus R. Testing and estimating in change-point problem of the spectral function.// Liet Matem. Rink. 1992. vol. 32, N 1, p.20.

4.Абрамович М.С. Обнаружение "разладок" с использованием спектральных статистик.// "Компьютерный анализ данных и моделирование". Минск, 1995.

5.Serfling R.J. Approximation Theorems of Mathematical Statistics. New York, 1980.

6.Андерсон Т. Статистический анализ временных рядов. М.,1971.

ВЫПИСКА

из протокола № от апреля 1997 г.
заседания кафедры математического моделирования и анализа данных

СЛУШАЛИ: Информацию о статье М.С. Абрамовича **«Оценка мощности спектрального критерия обнаружения момента «разладки» временных рядов».**

ПОСТАНОВИЛИ: Рекомендовать статью М.С. Абрамовича **«Оценка мощности спектрального критерия обнаружения момента «разладки» временных рядов»** к опубликованию в журнале «Вестник Белорусского университета» (Серия 1: Физика. Математика. Информатика).

Зав. кафедрой
математического моделирования и анализа данных,
профессор Ю.С. Харин

УДК 519.2

Абрамович М.С. **Оценка мощности спектрального критерия обнаружения момента «разладки» временных рядов** // Вестн. Белорус. ун-та. Сер.1. 199 . №

Решена задача асимптотического анализа мощности критерия обнаружения момента «разладки» временных рядов, основанного на оценках спектральных плотностей.

Библиограф. 6 назв.

Сведения об авторе

АБРАМОВИЧ МИХАИЛ СЕМЕНОВИЧ, старший научный сотрудник.
Место работы: БГУ, ФПМИ, НИЛ статистического анализа и моделирования.
Адрес:220020, г. Минск, ул. Л. Украинки, д. 12, корп. 2, кв. 13.
Раб. тел. :268-70-58
Дом. тел. :250-52-35

Abramovitch M.S. The estimator of power of spectral criterion for disorder detection in time series.

The problem of asymptotic analysis of power of criterion, which is based on spectral density estimators , is investigated for disorder detection test in time series.