ОБНАРУЖЕНИЕ РАЗЛАДОК СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ОСНОВЕ ВЕЙВЛЕТ-АНАЛИЗА

Лобач С. В.
Кафедра математического моделирования и анализа данных, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: serg88@mail.ru

Предлагается тест, основанный на максимуме вейвлет-периодограммы, для обнаружения разладки бинарной случайной последовательности.

Введение

Обнаружение разладок (скачкообразного изменения свойств) случайных последовательностей является важной прикладной задачей. Для решения таких задач в настоящее время широко используется вейвлет-анализ [1]. Оказывается, что абсолютные значения вейвлет-коэффициентов, подсчитанные для таких случайных последовательностей, имеют максимальные значения в моменты разладки. Решение задач, связанных с обнаружением изменения свойств объектов, применяется при исследовании различных технологических процессов и динамических систем.

I. Вейвлет-периодограмма случайной последовательности

Рассмотрим бинарную последовательность независимых случайных величин:

$$x(t) = \{x_t | t = \overline{0, T-1}\}, x_t \in \{1, -1\}, \quad (1)$$

где $T = 2^M$, M – известное натуральное число.

Последовательность (1) имеет разладку в неизвестный момент $\tau \in \{\tau_-, \tau_- + 1, \ldots, \tau_+\}$, где $0 < \tau_- < \tau_+ < T-1$ – известные моменты времени. Таким образом, она состоит из двух последовательностей:

$$x(t) = \begin{cases} x_1(t), & 0 \le t \le \tau - 1, \\ x_2(t), & 0 \le t \le T - 1, \end{cases}$$

причем $\mathrm{P}\{x_1(t) = 1\} = p_1$, $\mathrm{P}\{x_2(t) = 1\} = p_2$, где $p_1, p_2 \in [0, 1]$, – неизвестные вероятности. Математически решение задачи о разладке сводится к проверке гипотезы о распределении вероятностей последовательности (1) $H_0 : p_1 = p_2 = 0.5$; $H_1 : p_1 \neq p_2$.

Для проверки гипотезы о распределении вероятностей будем использовать статистику, предложенную в [2] и названную вейвлет-периодограммой.

Дискретное вейвлет-преобразование последовательности (1) задается путем вычисления коэффициентов:

$$J_{j,k} = \sum_{t=0}^{T-1} x_t \varphi_{j,k}(t); \quad d_{j,k} = \sum_{t=0}^{T-1} x_t \psi_{j,k}(t), \quad (2)$$

где $\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right), \varphi_{j,k}(t) = 2^{-j/2}\varphi\left(2^{-j}t - k\right)$, $j = \overline{1, M}$; $k = \overline{0, 2^{M-j} - 1}$; $\psi(t)$ – материнский вейвлет, $\varphi(t)$ – масштабирующая функция, j – параметр масштаба (уровень разрешения), k – параметр сдвига.

Для простоты будем использовать вейвлет Хаара [1], для которого

$$\psi(t) = \begin{cases} 1, & 0 \le t < 0.5, \\ -1, & 0.5 \le t \le 1, \\ 0, & t < 0, t > 1. \end{cases} \quad \varphi(t) = \begin{cases} 1, & t \in [0, 1], \\ 0, & t < 0, t > 1. \end{cases}$$

Теорема 1. [2] Коэффициенты вейвлет-преобразования последовательности (2) $\{x_t\}$, $t = \overline{1, T}$, $T = 2^M$, построенные с использованием вейвлета Хаара, на каждом уровне разрешения $j = \overline{1, M}$ независимы и имеют распределение вероятностей

$$\mathrm{P}\left\{2^{j/2} d_{j,k} = r\right\} = \frac{C_{2^i}^{\frac{2^j + r}{2}}}{2^{2^j}},$$

$$j = \overline{1,\ M};\ r \in \left\{-2^j, -2^i + 1, \ldots, 2^i - 1, 2^j\right\}.$$

Теорема 2. Если верна гипотеза H_0, то коэффициенты вевлет-преобразования при $T \to \infty$ на каждом уровне разрешения $j = \overline{1, M}$ имеют асимптотически нормальное распределение.

Из теоремы 2 вытекает, что гипотезы H_0 и H_1 о распределении вероятностей бинарной последовательности можно заменить семейством гипотез $\{H_{0j}, H_{1j}\}$, $j = \overline{1, M}$, о распределении коэффициентов вейвлет-преобразования на каждом уровне разрешения $j = \overline{1, M}$:

$$H_{0J} : d_{j,k} \sim N(0, 1),$$
$$H_{1j} : \text{в противном случае.}$$

Определим вейвлет-периодограмму случайной последовательности (1), следуя [2]:

$$\mathrm{I}_{j,k} = \left(d_{j,k}\right)^2, \ k = \overline{0, 2^{M-j} - 1},\ j = 1, M. \quad (3)$$

II. Критерий, основанный на максимуме вейвлет-периодограммы

Из теоремы 2 очевидно вытекает следующее свойство вейвлет-периодограммы: при $T \to \infty$ и при выполнении нулевой гипотезы статистика (3) имеет асимптотически распределение хи-квадрат с одной степенью свободы:

$$\mathrm{I}_{j,k} \xrightarrow{\mathrm{P}} \chi_1^2.$$

определим максимум вейвлет-периодограммы на каждом уровне разрешения $j = \overline{1, M}$:

$$\mathrm{I}_{j,\max} = \max_k \mathrm{I}_{j,k}. \qquad (4)$$

найдем асимптотическое распределение вероятностей статистики (4). Пусть $F_{\chi_1^2}(x)$ – функция распределения хи-квадрат с одной степенью свободы, $M_j = 2^{M-j}$, тогда функция распределения вероятностей статистики $\mathrm{I}_{j,\max}$ имеет вид:

$$F_{\mathrm{I}_{j,\max}}(x) = \left(F_{\chi_1^2}(x)\right)^{Mj}.$$

Критерий для обнаружения разладки последовательности (1), основанный на максимуме вейвлет-периодограммы, строится следующим образом. На каждом уровне разрешения проверяются гипотезы H_{0j} и H_{1j}, $j = \overline{1, M}$ решающее правило состоит в следующем: принимается гипотеза $\begin{cases} H_{0j} & \text{если } \mathrm{P}_j > \varepsilon^*, \\ H_{1j} & \text{если } \mathrm{P}_j \le \varepsilon^*, \end{cases}$ где P – значение для каждого уровня значимости P_j вычисляется как $\mathrm{P}_j = 1 - F_{\chi_1^2}^{M_j}(\mathrm{I}_{j,\max})$, $\varepsilon^* = \dfrac{\varepsilon}{M}$, ε – уровень значимости критерия.

Гипотеза H_0 принимается в том случае, если на всех уровнях разрешения $j = \overline{1, M}$ принимаются гипотезы H_{0j}, а альтернативная гипотеза H_1 принимается, если хотя бы на одном уровне разрешения принимается гипотеза H_1.

Приведем некоторые результаты численного моделирования. Моделируемые последовательности состояли из двух однородных фрагментов длины T_1 и T_2, $T = T_1 + T_2$, с разладкой в момент $\tau = 2^7$. Первый фрагмент представляет случайную последовательность с $P - 1 = 0.5$; а для второго фрагмента $p_2 = 0.2$. Оценка вероятности ошибки второго рода критерия при $T = 2^8$ была равна $\hat{\beta} \approx 0.06$, причем оценки вероятностей ошибок второго рода $\hat{\beta}$ уменьшались с ростом T. Этот факт свидетельствует об эффективности критерия максимума вейвлет-периодограммы для обнаружения разладок случайных последовательностей.

III. Список литературы

1. Дремин, И. М. Вейвлеты и их использование / И. М. Дремин, О. В. Иванов, В. А. Нечитайло // Успехи физической науки. – 2001. – Т. 171, № 5. – С. 5–79.
2. Chiann, C. Wavelet Analysis for Stationary Processes / C. Chiann, P. Morettin // Journal of Nonparametrical Statistics – 1998. – Vol. 36, № 10. – P. 1129–1142.