УДК 519.226

В.Р. Крашенинников, Е.А. Гладких

ТЕСТ ДЛЯ ПРОВЕРКИ ГИПОТЕЗ О КОВАРИАЦИОННОЙ ФУНКЦИИ И СПЕКТРАЛЬНОЙ ПЛОТНОСТИ СЛУЧАЙНОГО ПРОЦЕССА[1]

Крашенинников Виктор Ростиславович, *доктор технических наук, профессор, окончил Казанский государственный университет, заведующий кафедрой «Прикладная математика и информатика» Ульяновского государственного технического университета. Имеет работы по статистическим методам обработки сигналов и изображений. [e-mail: kvr@ulstu.ru].*

Гладких Екатерина Анатольевна, *кандидат технических наук, системный аналитик Московского отделения Яндекс. Окончила экономико-математический факультет УлГТУ. Имеет работы в области по моделированию и анализу случайных процессов прогнозирования временных рядов. [e-mail: kate-glad@yandex.ru].*

Аннотация

Математической моделью самых разнообразных явлений, протекающих во времени, являются случайные процессы (СП). Например, морское волнение, ветер, помехи, шумы, ошибки траекторных измерений и т. д. В связи с этим возникает задача определения ковариационной функции (КФ) или спектральной плотности мощности (СПМ) анализируемого СП. Это может быть реальный процесс или же имитируемый процесс, используемый для тестирования алгоритма обработки, например, имитация морского волнения. По методам оценивания КФ и СПМ имеется очень большое количество публикаций, в которых предложен целый ряд эффективных алгоритмов. Однако остается недостаточно исследованным вопрос об идентификации характеристики СП «в целом». Например, имеет ли исследуемый процесс КФ некоторого предполагаемого (гипотетического) вида? На этот вопрос можно ответить, анализируя разницу между измерениями КФ и их предполагаемыми значениями. Однако остается неопределенность в допустимых расхождениях, которые должны оцениваться в их совокупности. Другими словами, нужно найти критерий для проверки статистической гипотезы о виде КФ или СПМ процесса, используя имеющуюся реализацию. В данной работе предлагается и исследуется такой критерий.

Ключевые слова: случайный процесс, ковариационная функция, спектральная плотность, проверка гипотез, критерий, уровень значимости, мощность.

HYPOTHESIS TEST FOR COVARIANCE FUNCTION AND SPECTRAL DENSITY OF RANDOM PROCESS

Victor Rostislavovich Krasheninnikov, *Doctor of Engineering, Professor; graduated from Kazan State University; a head of the Department of Applied Mathematics and Computer Sciences at Ulyanovsk State Technical University; an author of papers in the field of statistical methods of signal and image processing. e-mail: kvr@ulstu.ru.*

Ekaterina Anatolievna Gladkikh, *Candidate of Engineering, an antirobot system analyst at Moscow Yandex Marketing Department; graduated from the Faculty of Economics and Mathematics at Ulyanovsk State Technical University; an author of articles in the field of random process modeling. e-mail: kate-glad@yandex.ru.*

1 Результаты этой работы были получены в рамках госзадания между вузом и Минобрнауки РФ №2014/232 .

Abstract

Random processes (RP) are a mathematical model of a variety of phenomena occurring in time such as sea waves, wind, jams, noise, trajectory measurement errors, etc. Thereby, the problem of determining the covariance function (CF) or power spectral density (SD) of the analyzed process emerged. This can be a real process or simulated process used to test the algorithm processing for example the simulation of sea waves. There are a large number of publications on methods for estimating CF and PD in which a great number of efficient algorithms is proposed. However, it is not sufficient to investigate the question of identifying the characteristics of the process "as a whole". For example, does the monitoring CF process some alleged (hypothetical) species? This question can be answered by analyzing the difference between the measurements and their putative CF values. However, uncertainty remains within the allowable differences, which should be assessed in their entirety. In other words, you need to detect a criterion for significance test of a CF or SD process using the existing implementation. In this paper such a criterion is proposed and investigated.

Key words: random process, covariance function, spectral density, hypothesis tests, criterion, significance level, test power.

Введение

Для успешного решения многих задач обработки СП очень важно адекватное описание СП в виде некоторых его характеристик, например, СПМ или КФ. Этим объясняется большое количество публикаций по методам измерения характеристик СП, например, [1–16]. При этом непосредственные оценки СПМ являются несостоятельными, поэтому СПМ оценивается косвенно на базе оценок КФ, с использованием соотношения между этими характеристиками [1].

Относительно рассматриваемого СП может быть выдвинута некоторая гипотеза. Имеется много работ по критериям для проверки гипотез о СП [2, 3, 5, 9 и т. д.]. Например, проверка гипотез о гауссовости, среднем значении, однородности и т. д. Однако нам неизвестны работы по критериям проверки гипотез о виде СПМ или КФ. Можно, конечно, оценить значения КФ в нескольких точках и сделать вывод о соответствии КФ исследуемого СП некоторому ее гипотетическому виду, если эти оценки «достаточно близки» к предполагаемым. Но возникает вопрос, что значит «достаточно близки»? Для ответа на этот вопрос следует, как это и делается в математической статистике, построить соответствующий критерий с заданным уровнем значимости и с хорошей мощностью.

В данной работе предлагается такой критерий для проверки гипотез о виде КФ или СПМ. В этом критерии проверяется гипотеза о соответствии измерений КФ в нескольких точках гипотетическим значениям КФ в этих точках [10]. Используемая статистика имеет распределение хи-квадрат. Приводится пример использования этого критерия для авторегрессионного СП. Исследуется мощность критерия.

Постановка задачи

Пусть имеется стационарный гауссовский дискретный процесс $x_1, x_2, \ldots$, который без потери общности можно считать центрированным. Его КФ

$$V_n = M[x_i, x_{i\pm n}]$$

неизвестна. Требуется по имеющейся выборке

$$\overline{x} = (x_1, x_3, \ldots, x_N)^T$$

с уровнем значимости α проверить простую гипотезу

$$H_0 : V_n = c_n,\ n \in N_r, \qquad (1)$$

против сложной гипотезы

$$H_1 : \exists n\left(V_n \neq c_n\right), \qquad (2)$$

где N_r – множество из r значений параметра n; а (2) означает, что $V_n \neq c_n$ хотя бы для одного значения n из N_r. Отметим, что в (2) фигурируют значения ковариаций c_n только для значений n из некоторого заданного множества N_r, а не для всех целочисленных значений. Сделано это по двум причинам. Во-первых, интерес обычно и представляют значения ковариаций для ограниченного множества значений аргумента. Во-вторых, по ограниченной реализации возможно оценивать ковариации процесса только на ограниченных расстояниях по времени. Таким образом, фактически проверяется гипотеза только о том, что выборочные значения ковариаций и их теоретические значения статистически не противоречат друг другу на множестве N_r.

Оценка значений КФ

Построим сначала оценку максимального правдоподобия (МП) значений V_n. Функция правдоподобия (ФП) в нашем случае имеет вид:

$$L\left(\overline{x} \mid \{V_n\}\right) = \frac{1}{(2\pi)^{n/2}|V|^{1/2}} exp\left(-\frac{1}{2}\overline{x}^T V^{-1}\overline{x}\right), \qquad (3)$$

где V – ковариационная матрица:

$$V = \begin{pmatrix} V_0 & V_1 & V_2 & \ldots & V_{N-1} \\ V_1 & V_0 & V_1 & \ldots & V_{N-2} \\ V_2 & V_1 & V_0 & \ldots & V_{N-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ V_{N-1} & V_{N-2} & V_{N-3} & \ldots & V_0 \end{pmatrix}.$$

Обратная матрица $U = V^{-1} = (u_{ij})$ в (3) также симметрична. Кроме того, в ней элементы, стоящие на равных расстояниях от концов (или от середины) главной и побочной диагоналей, равны между собой. Итак,

$$u_{ij} = u_{ji},\quad u_{ij} = u_{N-i,N+2-j},\quad u_{ij} = u_{N+2-i,N-j}.$$

Например, при $N = 4$ имеем:

$$V = \begin{pmatrix} V_0 & V_1 & V_2 & V_3 \\ V_1 & V_0 & V_1 & V_2 \\ V_2 & V_1 & V_0 & V_1 \\ V_3 & V_2 & V_1 & V_0 \end{pmatrix},$$

$$V^{-1} = U = \begin{pmatrix} u_1 & u_2 & u_3 & u_4 \\ u_2 & u_5 & u_6 & u_3 \\ u_3 & u_6 & u_5 & u_2 \\ u_4 & u_3 & u_2 & u_1 \end{pmatrix}. \quad (4)$$

В частных случаях обратная матрица может иметь более простую структуру. Например, для авторегрессионного процесса первого порядка

$$V = \sigma_x^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & \dots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{N-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \dots & 1 \end{pmatrix},$$

$$V^{-1} = \frac{1}{\sigma_x^2\left(1-\rho^2\right)} \begin{pmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & -\rho \\ 0 & 0 & 0 & -\rho & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} u_1 & u_2 & 0 & \dots & 0 \\ u_2 & u_3 & u_2 & \dots & 0 \\ 0 & u_2 & u_3 & \dots & 0 \\ \dots & \dots & \dots & u_3 & u_2 \\ 0 & 0 & 0 & u_2 & u_1 \end{pmatrix}. \quad (5)$$

Однако в общем случае структура матрицы V^{-1} имеет вид (4).

Из (4) следует, что $\bar{x}^T V^{-1} \bar{x}$ является взвешенной суммой статистик:

$$\begin{aligned} &S_{01} = x_1^2 + x_N^2, \quad S_{02} = x_2^2 + x_{N-1}^2, \dots, \\ &\qquad \dots, S_{0k} = x_k^2 + x_{N-k+1}^2, \dots \\ &S_{11} = x_1 x_2 + x_{N-1} x_N, \quad S_{12} = x_2 x_3 + x_{N-2} x_{N-1}, \dots \\ &S_{21} = x_1 x_3 + x_{N-2} x_N, \quad S_{22} = x_2 x_4 + x_{N-3} x_{N-1}, \dots \\ &\dots \\ &S_{N-2,1} = x_1 x_{N-1} + x_2 x_N, \\ &S_{N-1,1} = x_1 x_N. \end{aligned} \quad (6)$$

Последняя статистика в каждом ряду (6) имеет одно или два слагаемых в зависимости от нечетности или четности номера ряда, считая снизу.

Очевидно, что набор (6) является достаточной статистикой для оценки всех N значений КФ $V_0, V_1, \dots, V_{N-1}$.

Множество (6) содержит $N(N+1)/4$ статистик. Это уже при $N \geq 4$ больше, чем в тривиальной достаточной статистике $\{x_1, x_2, \dots, x_N\}$. В частных случаях количество статистик может быть значительно меньше. Например, для (5) достаточным является набор:

$$T = x_1^2 + x_N^2, \; S = x_2^2 + x_3^2 + \dots + x_{N-1}^2,$$
$$P = x_1 x_2 + x_2 x_3 + \dots + x_{N-1} x_N.$$

Таким образом, в общем случае построение оценок МП с помощью достаточных статистик не приводит к упрощению оценивания. Следовательно, не приходится надеяться и на построение простых критериев различения гипотез на основе достаточных статистик.

Отметим также, что в общем случае элементы матрицы V^{-1} в (4) зависят сложным образом от всех значений $V_0, V_1, \dots, V_{T-1}$ матрицы V. Поэтому оценка даже части значений $\{V_n : n \in N_r\}$ явно или неявно повлечет за собой оценку всех $V_0, V_1, \dots, V_{N-1}$, так как уравнение правдоподобия приводит к нелинейной системе уравнений относительно $V_0, V_1, \dots, V_{N-1}$.

Будем использовать следующую оценку КФ:

$$\hat{V}_n = \frac{1}{N-n} \sum_{k=1}^{N-n} x_k x_{k+n}, \quad (7)$$

которая не смещена, состоятельна и асимптотически нормальна. Построим критерий различения гипотез о КФ на основе статистик:

$$S_n = \frac{1}{N-n} \sum_{k=1}^{N-n} x_k x_{k+n}, \; n \in N_r. \quad (8)$$

В наборе (8) содержится r статистик, независимо от длины выборки N. Так что в смысле количества мы имеем минимальный набор для требуемого оценивания r параметров V_n, $n \in N_r$.

Построение критерия

Построим критерий [10] проверки гипотезы (1) против (2), основываясь на статистиках (8), то есть на статистике

$$\bar{V} = \{S_n, n \in N_r\}. \quad (9)$$

Пусть гипотеза H_0 верна, тогда (9) будет несмещенной оценкой вектора параметров

$$\bar{c} = \{c_n, n \in N_r\}.$$

Известно [2, 9], что при $N \to \infty$ случайные величины

$$\varphi_n = \sqrt{N-n}\left(S_n - c_n\right) \quad (10)$$

асимптотически нормальны с нулевым средним и ковариациями

$$\Psi_{hg} = \sum_{k=-\infty}^{\infty} \left[c_{k+h} c_{k+g} + c_{k+h} c_{k-g}\right] = $$
$$= \sum_{k=-\infty}^{\infty} \left[c_k c_{k+h-g} + c_{k+g} c_{k-g}\right], \quad (11)$$

то есть вектор

$$\bar{\varphi} = \{\varphi_n, n \in N_r\}$$

имеет нулевое среднее и ковариационную матрицу

$$\Psi = \{\psi_{h,g}\}, \; h, g \in N_r.$$

Таким образом, ФП для гипотезы H_0 будет

$$L\left(\bar{S} | H_0\right) = L\left(\bar{\varphi} | H_0\right) \frac{1}{(2\pi)^{\frac{r}{2}} |\Psi|^{\frac{1}{2}}} exp\left(-\frac{1}{2} \bar{\varphi}^T \Psi^{-1} \bar{\varphi}\right). \quad (12)$$

Рассмотрим сначала простую конкурирующую гипотезу

$$\tilde{H}_1 : V_n = d_n, n \in N_r. \qquad (13)$$

Для нее ФП при аналогичных обозначениях будет

$$L(\overline{S} \mid \tilde{H}_1) = L(\overline{\mu} \mid \tilde{H}_1) \frac{1}{(2\pi)^{\frac{r}{2}} |G|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\overline{\mu}^T G^{-1}\overline{\mu}\right). \qquad (14)$$

Наилучшей критической областью (НКО) для гипотезы H_0 против $\tilde{H}_1$ является [5]

$$\frac{L(\overline{S} \mid H_0)}{L(\overline{S} \mid \tilde{H}_1)} = \frac{L(\overline{\varphi} \mid H_0)}{L(\overline{\mu} \mid \tilde{H}_1)} \le k_\alpha, \qquad (15)$$

где k_α зависит от заданного уровня значимости α. Подставляя (12) и (14) в (15) и логарифмируя, получаем

$$\overline{\mu}^T G^{-1}\overline{\mu} - \overline{\varphi}^T \Psi^{-1}\overline{\varphi} \le l_\alpha = ln\left(\frac{k_\alpha |G|^{1/2}}{|\Psi|^{1/2}}\right). \qquad (16)$$

Напомним, что

$$\overline{\varphi} = \left\{\sqrt{N-n}\,(S_n - c_n),\, n \in N_r\right\},$$
$$\overline{\mu} = \left\{\sqrt{N-n}\,(S_n - d_n),\, n \in N_r\right\}, \qquad (17)$$

поэтому относительно $\overline{\mu}$ и $\overline{\varphi}$ выражения в левой части (16) – обычные, то есть приведенные к началу координат квадратичные формы, а относительно $\overline{S}$ – смещенные квадратичные формы с центрами в точках $\overline{c}$ и $\overline{d}$ r-мерного пространства. За исключением вырожденных случаев, когда $|\Psi| = 0$ и/или $|G| = 0$, эти квадратичные формы положительно определены. Им соответствуют гиперэллипсоиды равных значений. Однако разность двух квадратичных форм может быть квадратичной формой с любым рангом и любой сигнатурой, то есть с любым количеством отрицательных, положительных и нулевых квадратов в каноническом виде. Поэтому поверхность, определяющая НКО в (16), может быть самой разнообразной. Например, в двумерном случае ($r = 2$) это может быть кривая второго порядка любого типа (гиперболического, эллиптического или параболического), включая вырожденные случаи. Анализ НКО существенно усложняется еще и тем, что матрица формы зависит от значений КФ d_n.

Зависимость НКО от параметров простой гипотезы $\tilde{H}_1$ означает, что не существует равномерно наиболее мощного (РНМ) критерия для сложной гипотезы (2). Поэтому можно только принять какой-нибудь компромиссный критерий. С этой целью преобразуем выражение (16). Из (17) следует:

$$\mu_n = \sqrt{N-n}\,(S_n - d_n) =$$
$$= \sqrt{N-n}\,\big((S_n - c_n) - (d_n - c_n)\big) =$$
$$= \sqrt{N-n}\,(S_n - c_n) - \sqrt{N-n}\,(d_n - c_n) =$$
$$= \varphi_n - \sqrt{N-n}\,\delta_n = \varphi_n - \Delta_n,$$

где

$$\delta_n = d_n - c_n,\ \Delta_n = \sqrt{N-n}\,\delta_n.$$

Далее

$$\overline{\mu}^T G^{-1}\overline{\mu} = (\overline{\varphi} - \overline{\Delta})^T G^{-1} (\overline{\varphi} - \overline{\Delta}) =$$
$$= \overline{\varphi}^T G^{-1}\overline{\varphi} - 2\overline{\Delta}^T G^{-1}\overline{\varphi} + \overline{\Delta}^T G^{-1}\overline{\Delta}. \qquad (18)$$

Подставляя (18) в (16), получим

$$\overline{\varphi}^T G^{-1}\overline{\varphi} - \overline{\varphi}^T \Psi^{-1}\overline{\varphi} - 2\overline{\Delta} G^{-1}\overline{\varphi} \le l_\alpha - \overline{\Delta}^T G^{-1}\overline{\Delta}. \qquad (19)$$

Пусть верна гипотеза H_0. Тогда $M[\overline{\varphi}] = 0$, $M[\overline{\Delta}^T G^{-1}\overline{\varphi}] = 0$, $M[\overline{\varphi}^T G^{-1}\overline{\varphi}] = 0$ и $M[\overline{\varphi}^T \Psi^{-1}\overline{\varphi}] = 0$ ограничены и стремятся к своим предельным значениям при $N \to \infty$. Величина $\overline{\Delta}^T G^{-1}\overline{\Delta}$ не случайна, является положительно определенной квадратичной формой от переменных $\Delta_n = \sqrt{N-n}\,(d_n - c_n)$, которые при $d_n \ne c_n$ увеличиваются с ростом N:

$$\lim_{N\to\infty} |\Delta_n| = \infty, \text{ при } d_n \ne c_n.$$

Таким образом, при истинности гипотезы H_0 имеем:

$$\lim_{N\to\infty} \overline{\Delta}^T G^{-1}\overline{\Delta} = +\infty. \qquad (20)$$

Из (19) и (20) следует, что при любой простой альтернативе $\tilde{H}_1$ НКО для H_0 бесконечно сокращается (при непрерывных распределениях $\overline{\varphi}$).

Перепишем (19) в виде:

$$\overline{\varphi}^T \Psi^{-1}\overline{\varphi} \ge \overline{\Delta}^T G^{-1}\overline{\Delta} - l_\alpha - 2\overline{\Delta} G^{-1}\overline{\varphi} + \overline{\varphi}^T G^{-1}\overline{\varphi}. \qquad (21)$$

Учитывая, что с ростом N правая часть (21) бесконечно возрастает, можно сделать вывод, что при любой простой гипотезе $\tilde{H}_1$ и достаточно большом объеме выборки НКО для H_0 содержит точки, в которых $\overline{\varphi}^T \Psi^{-1}\overline{\varphi}$ имеет большие значения. Поэтому целесообразно только такие точки включить в критическую область гипотезы H_0 при сложной гипотезе (2):

$$\overline{\varphi}^T \Psi^{-1}\overline{\varphi} \ge b_\alpha,$$

где пороговое значение b_α соответствует уровню значимости α.

Таким образом, получаем следующий критерий проверки гипотезы H_0 против сложной альтернативы H_1:

$$\overline{\varphi}^T \Psi^{-1}\overline{\varphi} < b_\alpha \Rightarrow \text{принимается } H_0,$$
$$\overline{\varphi}^T \Psi^{-1}\overline{\varphi} \ge b_\alpha \Rightarrow \text{принимается } H_1. \qquad (22)$$

Смысл этого критерия прост: гипотеза H_0 принимается, если отклонения выборочных значений $\overline{S}_n$ КФ от ее точных значений c_n в совокупности невелики.

Мощность критерия (22), естественно, зависит от конкретной альтернативы $\tilde{H}_1$ и может быть существенно ниже мощности наиболее мощного критерия (НМК) для этой альтернативы при небольших значениях N.

Рассмотрим теперь случай, когда верна некоторая простая альтернатива $\tilde{H}_1$, то есть когда точными значениями КФ в (16) и (17) будут d_n. На этот раз представим статистики $\overline{\varphi}$ в виде:

$$\varphi_n = \sqrt{N-n}\,\big((S_n - d_n) + (d_n - c_n)\big) = \mu_n + \Delta_n,$$

то есть $\overline{\varphi} = \overline{\mu} + \overline{\Delta}$,

приведем (16) к виду:

$$\overline{\mu}^T G^{-1}\overline{\mu} \le \overline{\Delta}^T \Psi^{-1}\overline{\Delta} + l_\alpha + 2\overline{\Delta}^T \Psi^{-1}\overline{\mu} + \overline{\mu}^T \Psi^{-1}\overline{\mu}. \quad (23)$$

При возрастании N значение $\overline{\Delta}^T \Psi^{-1}\overline{\Delta}$ бесконечно возрастает, а величины $\overline{\mu}$ в среднем убывают, поэтому при $N \to \infty$ вероятность выполнения неравенства (23), а также и (22), стремится к единице. Это означает, что при любой простой альтернативе $\tilde{H}_1$ из сложной гипотезы H_1 и любом уровне значимости α альтернатива $\tilde{H}_1$ принимается с вероятностью, сколь угодно близкой к единице при $N \ge N(\tilde{H}_1)$.

Отметим, однако, что нет оснований утверждать, что $N(\tilde{H}_1) \le N_0$ для всех $\tilde{H}_1 \in H_1$. Поэтому, вообще говоря, нельзя указать объем выборки N_0, достаточный для обеспечения мощности критерия (22), не меньшей заданной, при всех альтернативах H_1. Существование такого равномерного N_0 возможно, если H_1 является классом альтернатив определенного вида.

Найдем теперь порог b_α в (22), соответствующий заданному уровню значимости. Для этого будем рассматривать случайные величины как элементы векторного пространства со скалярным произведением, равным ковариации. Тогда систему линейно независимых величин $\overline{\varphi}$ можно ортонормировать с помощью подходящего линейного преобразования

$$\overline{\xi} = R\,\overline{\varphi}.$$

Матрица R этого преобразования должна удовлетворять соотношению

$$R\,\Psi\,R^T = E.$$

Например, если применить процедуру Грама-Шмидта, то R – нижняя треугольная матрица, а R^T – верхняя треугольная.

Тогда $\overline{\varphi} = R^{-1}\xi$,

$$\overline{\varphi}^T \Psi^{-1}\overline{\varphi} = \overline{\xi}^T \left(R^{-1}\right)^T \Psi^{-1} R^{-1}\overline{\xi} = \overline{\xi}^T \left(R\Psi R^T\right)^{-1}\overline{\xi} = \overline{\xi}^T\overline{\xi},$$

и (22) принимает вид:

$$\begin{aligned}&\overline{\xi}^T\overline{\xi} < b_\alpha \Rightarrow \text{принимается } H_0,\\ &\overline{\xi}^T\overline{\xi} \ge b_\alpha \Rightarrow \text{принимается } H_1.\end{aligned} \quad (24)$$

При этом

$$\overline{\xi}^T\overline{\xi} = \xi_1^2 + \xi_2^2 + \ldots + \xi_r^2, \quad (25)$$

где $\xi_1, \xi_2, \ldots, \xi_r$ – независимые стандартные случайные величины с распределениями, асимптотически приближающимися к гауссовским при $N \to \infty$. Следовательно распределение статистики в (22) имеет распределение, близкое к распределению хи-квадрат с r степенями свободы. Таким образом, порог в (24) или (22) определяется как $(1-\alpha)$–квантиль хи-квадрат распределения с r степенями свободы. В другой принятой терминологии $b_\alpha = \chi^2_{кр}$ – критическое значение. Следует учитывать, что при малых значениях N найденный здесь порог может давать неточное значение уровня значимости из-за существенных отклонений распределения статистики в (25) от хи-квадрат.

Заметим, что описанное здесь ортонормирование выполнять необязательно. Можно использовать критерий в форме (22). Однако тогда нужно обращать матрицу Ψ. Вычисление статистики в (24) требует примерно в два раза меньше операций, чем в (22).

Критерий для проверки гипотезы о СПМ $S(\sigma)$ стационарного СП может быть построен подобным же образом, поскольку выборочные оценки этой характеристики также нормализуются при $N \to \infty$ [6]. Однако более качественные оценки спектральной плотности получаются из оценок КФ [1, 11, 14–16]. Поэтому целесообразно проверку гипотезы о СПМ $S(\sigma)$ свести к проверке гипотезы о соответствующей ей КФ V_n.

Пример применения критерия для авторегрессионного процесса

Приведем в качестве примера применение представленного критерия для проверки гипотезы об экспоненциальной КФ [10]:

$$V_n = \sigma^2 exp\left(-\lambda\,|\,n\,|\right) = \sigma^2\rho^{|n|} \quad (26)$$

при большой длине реализации N, где $\rho = exp(-\lambda)$. Отметим, что эта ковариационная функция соответствует спектральной плотности

$$S(\sigma) = \frac{4\sigma\lambda}{\lambda^2 + \sigma^2}, \quad (27)$$

поэтому проверка гипотезы о спектре (27) сводится к проверке гипотезы о ковариационной функции (26). В рассматриваемом случае (10) и (11) принимают вид:

$$\varphi_n = \sqrt{N-n}\left[\hat{V}_n - \sigma^2\rho^n\right],$$

$$\begin{aligned}\psi_{hg} &= cov\left(\varphi_h, \varphi_g\right) = \\ &= \sigma^2\left(\sum_{k=-\infty}^{\infty}\rho^{|k+h|+|k+g|} + \sum_{k=-\infty}^{\infty}\rho^{|k+h|+|k-g|}\right) = \\ &= \sigma^2\left[\left(\frac{2}{1-\rho^2} + h + g - 1\right)\rho^{h+g} + \right.\\ &\qquad \left. + \left(\frac{2}{1-\rho^2} + |\,h-g\,| - 1\right)\rho^{|h-g|}\right].\end{aligned} \quad (28)$$

Для статистического испытания разработанного критерия была составлена программа, с помощью которой были произведены следующие эксперименты.

Процесс с нулевым средним и КФ (28) имитировался с помощью авторегрессионной модели:

$$x_1 = \sigma\varsigma_1,\; x_n = \rho x_{n-1} + \sigma\sqrt{1-\rho^2}\,\varsigma_n, \quad (29)$$

где ς_n – независимые стандартные гауссовские случайные величины. Были взяты значения $\sigma = 1,\ \rho = 0{,}9$ и множество $N_r = N_2 = \{0,\ 1\}$. Таким образом, гипотеза H_0 состоит в том, что процесс (29) имеет ковариационную функцию $V_n = 0{,}9^{|n|}$ на указанном множестве N_2.

В данном примере ξ_k вычислялись по формулам процедуры Грамма-Шмидта:

$$\xi_1 = \frac{\varphi_1}{\sqrt{\left|cov(\varphi_1, \varphi_1)\right|}},$$

$$\xi_2 = \frac{cov(\varphi_1, \varphi_1)\cdot \varphi_2 - cov(\varphi_1, \varphi_2)\cdot \varphi_1}{\sqrt{\left|cov(\varphi_1, \varphi_1)\left[cov(\varphi_1, \varphi_1)cov(\varphi_2, \varphi_2) - cov(\varphi_1, \varphi_2)^2\right]\right|}}.$$

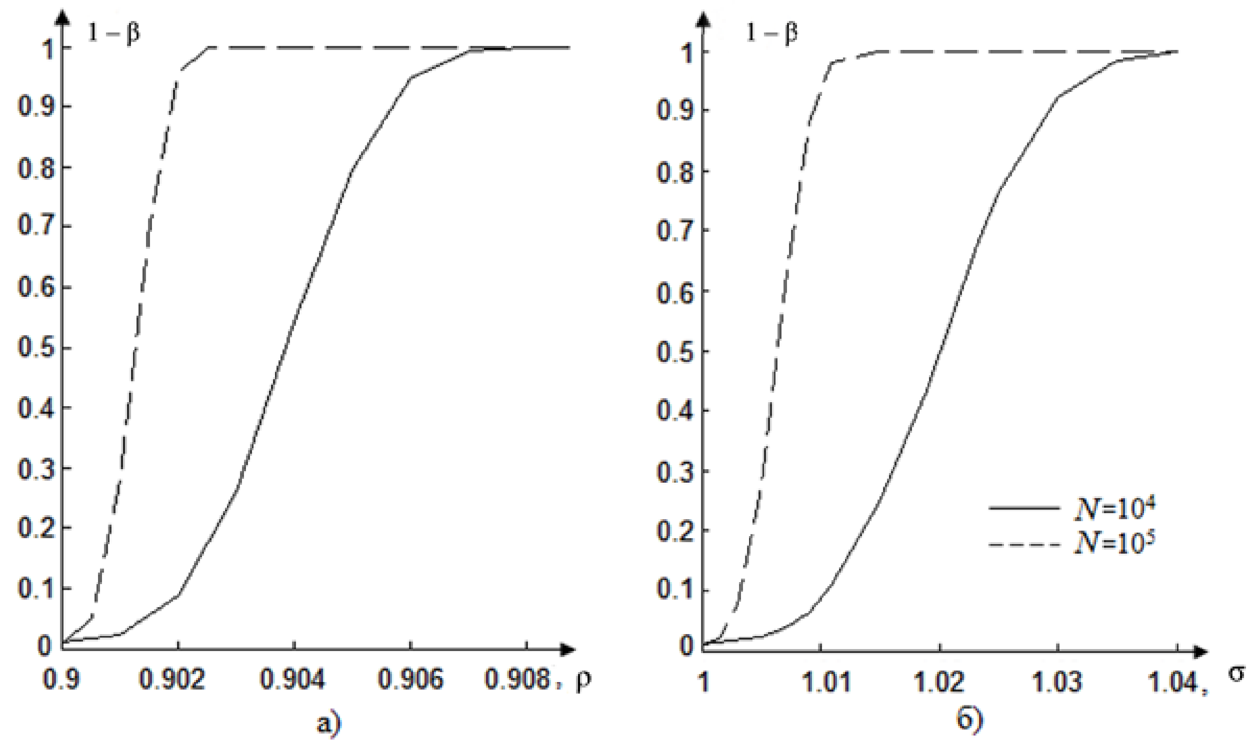

Рис. 1. Зависимость мощности критерия от параметров тестируемого процесса:
а) от коэффициента корреляции, б) от СКО

Были исследованы выборочные характеристики величин φ_n и ξ_k, которые оказались статистически соответствующими их теоретическому виду, то есть нормальны.

Оценка КФ процесса (29) находилась по формуле (8) при $\mu = 0$ и $N = 10\,000$, то есть использовались реализации из 10000 отсчетов этого процесса. По этой реализации вычислялось значение χ_2^2, которое сравнивалось в решающем правиле (22) с порогом $\chi_{кр}^2 = \chi_{0,01;2}^2 = 9{,}21$, соответствующим уровню значимости $\beta = 0{,}01$ при числе степеней свободы $r = 2$ [3]. Было произведено 10000 применений этого критерия. Оказалось, что верная гипотеза H_0 была отвергнута 102 раза, то есть относительная частота отвержения равна 0,0102. Это значение укладывается в 99% доверительный интервал $I_{0,99} = (0{,}0067;\ 0{,}0132)$. Таким образом, в этом эксперименте заданный уровень значимости можно считать выдержанным.

В другом эксперименте была исследована мощность критерия при проверке той же самой гипотезы H_0 против простых гипотез $\tilde{H}_1$ с другими значениями параметров σ и ρ. Критерий применялся также 10000 раз для каждого из этих наборов значений. На рисунке 1 приведены графики оцененной по результатам этих испытаний мощности, то есть относительной частоты отвержения неверной гипотезы. Из этих данных можно сделать вывод о довольно высокой чувствительности критерия к отклонениям параметров процесса от их предполагаемых значений.

Таким образом, предложенный критерий требует небольших вычислительных затрат, обладает достаточно высокой чувствительностью и может применяться для проверки гипотез о КФ и спектре СП.

Заключение

Предложен критерий для проверки гипотез о виде КФ наблюдаемого СП процесса. Этот критерий может быть использован и для проверки гипотез о виде СПМ, поскольку одна из этих характеристик выражается через другую. В критерии проверяется соответствие нескольких измерений характеристики их гипотетическим значениям. Статистика критерия имеет распределение хи-квадрат. Примеры применения предложенного критерия показали его высокую мощность, то есть чувствительность к отклонениям гипотетической характеристики от фактической. Критерий требует небольших вычислительных затрат.

СПИСОК ЛИТЕРАТУРЫ

1. Андерсон Т. Статистический анализ временных рядов. – М. : Мир, 1976. – 755 с.
2. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып. 1. – М. : Мир, 1974. – 406 с.
3. Ван-Трис Г. Теория обнаружения, оценок и модуляции. – М. : Советское радио, 1971.
4. Гладких Е.А., Крашенинников В.Р. Критерий согласия для ковариационной функции и спектра стационарного гауссовского случайного процесса // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем : сб. науч. тр. Седьмой вып. – Ульяновск : УлГТУ, 2010. – С. 69–73.
5. Грибанов Ю.И., Мальков В.Л. Выборочные оценки спектральных характеристик случайных процессов.– М. : Энергия, 1978. – 152 с.
6. Журбенко И.Г. Спектральный анализ временных рядов. – М. : Изд-во Моск. ун-та, 1982. – 168 с.
7. Кендалл М. Стьюарт А. Статистические выводы и связи.– М. : Главная редакция физико-математической литературы изд-ва «Наука», 1973. – 899 с.
8. Крамер Г. Математические методы статистики.– М. : Мир, 1975. – 648 с.
9. Куликов Е.И. Методы измерения случайных процессов. – М. : Радио и связь, 1986. – 272 с.
10. Лайонс Р. Цифровая обработка сигналов: 2-е изд. Пер. с англ. – М. : Бином–пресс, 2006. – 656 с.
11. Марпл С.Л. Цифровой спектральный анализ и его приложения. – М. : Мир, 1990. – 584 с.
12. Мирский Г.Я. Аппаратурное определение характеристик случайных процессов. – М. : Энергия, 1972. – 456 с.
13. Рабинер П., Гоулд Б. Теория и применение цифровой обработки сигналов. – М. : Мир, 1978. – 848 с.
14. Романенко А. Ф., Сергеев Г.А. Вопросы прикладного анализа случайных процессов. – М. : Советское радио, 1968. – 256 с.
15. Цветков Э.И. Основы теории статистических измерений. – Л. : Энергоатомиздат. Ленинградское отделение, 1986 – 256 с.
16. Hennan E.J. Time Series Analysis. New York, London. John Willey and Sons Inc., 1970.

REFERENCES

1. Anderson T. *Statisticheskiy analiz vremennykh ryadov* [The Statistical Time-Series Analysis]. Moscow, Mir Publ., 1976. 755 p.
2. Box G., Jenkins G. *Analiz vremennykh ryadov. Prognoz i upravleniye, Vyp.1* [Time-Series Analysis. Forecasting and Control, First Edition]. Moscow, Mir Publ., 1974. 406 p.
3. Harry L. Van Trees. *Teoriya obnaruzheniya, otsenok i modulyatsii* [Detection, Estimation, and Modulation Theory]. Moscow, Sovetskoye radio Publ., 1971.
4. Gladkikh Ye.A., Krasheninnikov V.R. Kriteriy soglasiya dlya kovariatsionnoy funktsii i spektra statsionarnogo gaussovskogo sluchaynogo protsessa [The Fitting Criterion for Covariance Function and Spectrum of Stationary Gaussian Random Process]. *Sovremennyye problemy proyektirovaniya, proizvodstva i ekspluatatsii radiotekhnicheskikh sistem : sbornik nauchnykh trudov, Sedmoy vypusk* [Actual Problems of Radioelectronic Systems Development, Production, and Operation: Collection of Scientific Papers], Ulyanovsk, UlSTU Publ., 2010, vol. 7, pp. 69–73.
5. Gribanov Yu.I., Malkov V.L. *Vyborochnyye otsenki spektralnykh kharakteristik sluchaynykh protsessov* [Sample Estimate of Random-Processes Spectral Characteristics], Moscow, Energiya Publ., 1978. 152 p.
6. Zhurbenko I.G. *Spektralnyy analiz vremennykh ryadov* [The Spectral Time-Series Analysis]. Moscow, Izdatelstvo Moskovskogo universiteta Publ., 1982. 168 p.
7. Kendall M., Stuart A. *Statisticheskiye vyvody i svyazi* [The Advanced Theory of Statistics. Inference and Relationship]. Moscow, Glavnaya redaktsiya fiziko-matematicheskoy literatury izd-va Nauka Publ, 1973. 899 p.
8. Cramer H. *Matematicheskiye metody statistiki* [Mathematical Methods of Statistics]. Moscow, Mir Publ.,1975. 648 p.
9. Kulikov Ye.I. *Metody izmereniya sluchaynykh protsessov* [Random-Processes Measurement Methods]. Moscow, Radio i svyaz Publ., 1986. 272 p.
10. Lyons R. *Tsifrovaya obrabotka signalov* [Digital Signal Processing. Second Edition]. Moscow, Binom–press Publ., 2006. 656 p.
11. Marple S.L. *Tsifrovoy spektralnyy analiz i yego prilozheniya* [Digital Spectral Analysis with Applications]. Moscow, Mir Publ., 1990. 584 p.
12. Mirskiy G.Ya. *Apparaturnoye opredeleniye kharakteristik sluchaynykh protsessov* [Hardware-based Random Processes Characterization]. Moscow, Energiya Publ., 1972. 456 p.
13. Rabiner L., Gold B. *Teoriya i primeneniye tsifrovoy obrabotki signalov* [Theory and Application of Digital Signal Processing]. Moscow, Mir Publ., 1978. 848 p.
14. Romanenko A. F., Sergeyev G.A. *Voprosy prikladnogo analiza sluchaynykh protsessov* [Application Analysis of Random Processes]. Moscow, Sovetskoye radio Publ., 1968. 256 p.
15. Tsvetkov E.I. *Osnovy teorii statisticheskikh izmereniy* [The Statistic Measurements Theory]. Leningrad, Energoatomizdat Leningradskoye otdeleniye Publ., 1986. 256 p.
16. Hannan E.J. *Time Series Analysis.* New York, London, John Willey and Sons Inc., 1970.