СОСТАВНЫЕ ВЕКТОРНЫЕ МОДЕЛИ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Тихонов В.А., Филь И.О., Кудрявцева Н.В.
Харьковский национальный университет радиоэлектроники
61166, Харьков, пр. Ленина,14, каф. РЭС тел. (057) 702-15-87,
E-mail: dragula@online.ua; (093) 477-01-97

The article considers a new class of complex stationary stochastic processes. Expressions are obtained to evaluate mean, variance, and correlation functions. The advantages of composite vector stochastic processes over conventional methods of statistical analysis are shown.

В рамках корреляционной теории для описания случайных последовательностей применяется гильбертово пространство [1]. Классический подход к построению гильбертовых пространств не позволяет в удобной форме строить модели некоторых классов сложных гауссовых и негауссовых сигналов и процессов. К таким классам относятся процессы, которые можно разбить на последовательность процессов меньшей длины, обладающих некоторыми общими свойствами. Например, ДНК последовательности, состоящие из кодонов, т.е. триплетов четырех возможных нуклеотидов. Предложенный в работе метод анализа подобных сигналов позволяет извлекать информацию, которую нельзя выявить обычными методами статистического анализа. Для получения и анализа, предложенных классов векторных коррелированных случайных процессов, использовались модели линейного предсказания [2, 3].

Был рассмотрен стационарный случайный процесс $X[t]$ в виде вектора $x[t]$ в линейном пространстве, который определяется своими координатами $x[1], x[2],..., x[N]$. Представим такой случайный процесс в виде последовательности подвекторов x_i одинаковой длины n с однородными статистическими свойствами. Таким образом, введено понятие «подвектора» x_i вектора $\vec{x}^n[t]$, по аналогии с отсчетом $x[t]$ выборки $X[t]$. Такой стационарный случайный процесс назван «составным векторным случайным процессом» (СВСП) $\vec{x}^n[t]$. СВСП является обобщением понятия случайного процесса, в котором его отсчет заменяется подвектором x_i длины n. При $n = 1$ СВСП становится обычным стационарным процессом в виде вектора $x[t]$. Такое естественное обобщение позволяет определить выражения для оценок статистических характеристик СВСП и синтезировать статистические модели. Коррелированным СВСП $x^n[t]$ является процесс, в котором существуют статистические связи второго порядка между подвекторами x_i. Процесс $x^n[t]$ представлен в виде их последовательности

$$x^n[t] = \{x_1, x_2,..., x_{N/n}\}.$$

Каждый подвектор определяется n координатами вектора $x^n[t]$:

$$x_1 = \{x[1], x[2],..., x[n]\},\ x_2 = \{x[n+1], x[n+2],..., x[2n]\},\ldots$$

$$\ldots, x_i = \{x[(i-1)n+1],..., x[in]\},\ldots, x_{N/n} = \{x[N-n+1],..., x[N]\},$$

где i – номер подвектора, N – номер последнего отсчета последнего подвектора. Если количество отсчетов вектора некратно длине подвектора n, то в качестве N/n берется целая часть этого числа, т.е. $N/n \sim \lfloor N/n \rfloor$. Для получения статистик второго порядка СВСП необходимо произведения координат вектора заменить на скалярное произведение подвекторов

$$(\vec{x}_i, \vec{x}_{i+k}) = \sum_{j=1}^{n} (x[(i-1)n+j]x[(i-1)n+j+kn]),$$

где k – сдвиг векторов, равный $1, 2, ..., \frac{N}{n} - 1$. Применяя усреднение скалярного произведения, получена формула оценки корреляционной функции СВСП

$$R^n[k] = \frac{1}{\frac{N}{n} - k} \sum_{i=1}^{\frac{N}{n} - k} \sum_{j=1}^{n} (x[(i-1)n + j]x[(i-1)n + j + kn]) .$$

Отсюда видно, что корреляционная функция СВСП описывает статистическую связь первого порядка между подвекторами. Также для СВСП найдены выражения для оценки математического ожидания, дисперсии, спектральной плотности мощности коррелограммным методом.

Очевидны существенные отличия между корреляционными функциями СВСП и стационарного процесса. Т.к. спектры Фурье и параметрические оценки спектров СВСП определяются их корреляционными функциями, они будут также существенно отличаться от классических спектров случайных процессов. Различия корреляционных функций и спектров были исследованы на имитационных процессах СВСП, полученных по моделям линейного предсказания. Классы моделей линейного предсказания включают модель авторегрессии (АР), модель скользящего среднего и модель авторегрессии-скользящего среднего. Спектры второго порядка этих процессов полностью описываются параметрами этих моделей. Разностное уравнение АР СВСП имеет вид

$$\vec{x}_i = \sum_{s=1}^{p} \Phi^n[s]\vec{x}_{i-s} + \vec{a}_i$$

где $\Phi^n[s]$ – коэффициенты АР СВСП, p – порядок модели АР СВСП, a_i – векторы длиной n отсчетов белого шума. Условие оптимальности модели АР СВСП состоит в статистической независимости подвекторов a_i. Для модели АР СВСП ошибки a_i должны быть некоррелированными, т.е. $E\{a_i a_{i-k}\} = 0,$ при $k \neq 0$. Это условие эквивалентно минимуму дисперсии ошибки предсказания СВСП D_a^n. Теоретические положения модели АР СВСП проверялись методами статистического моделирования.

В работе рассмотрен новый класс сложных стационарных случайных процессов, названный СВСП. К этому классу можно отнести разные виды векторных случайных процессов, представимых в виде последовательности подвекторов. Для СВСП, состоящего из непересекающихся смежных подвекторов найдены выражения для оценки математического ожидания, корреляционной функции, спектральной плотности мощности коррелограммным методом. Показано, что при длине подвектора, равного единице, найденные оценки совпадают с известными выборочными оценками статистик первого и второго порядков стационарных процессов.

Предложены выражения для расчета параметров АР СВСП и параметрической спектральной оценки. Получены оценки корреляционной функции и СПМ АР СВСП. Показано, что использование предложенных статистик и модели АР СВСП позволяет исследовать статистические характеристики подвекторов СВСП. Применение классических статистических методов исследования СВСП не позволяет найти такие характеристики. Работа может быть полезна для исследования случайных процессов и последовательностей, в частности для решения ряда задач анализа ДНК.

Литература**.** 1. Френкс Л. Теория сигналов. – М.: Сов. радио, 1969. – 344 с. 2. Бокс Дж., Дженкинс Г. Анализ временных рядов: Пер. с. англ. – М.: Мир, 1974. – Вып.1. – 406 с. 3. Марпл.–мл. С. Л. Цифровой спектральный анализ и его приложения – М.: Мир, 1990. – 584 с.