УДК 303.723

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ОДНОМЕРНЫХ МАССИВОВ

студент гр. 914302 Воробей Д. А.

Научный руководитель - канд. техн. наук Ролич О. Ч.

Белорусский государственный университет информатики и радиоэлектроники

Минск, Беларусь

Изучение методов цифровой обработки сигналов является важным и перспективным направлением развития современной науки и техники [1]. В связи с этим стремительно прогрессируют различные подходы к генерированию, обработке и анализу цифровых сигналов.

В реальных задачах часто ставится вопрос о степени идентичности или независимости процессов. Иными словами, требуется определить взаимосвязь между соответствующими процессам сигналами, то есть найти их корреляцию. Для этого используется корреляционный анализ – статистическое исследование различных выборок данных, направленное на выявление взаимосвязи между ними.

Значения корреляционной функции, с учётом нормировки выборок, расположены в диапазоне [–1; 1]. Если взять две абсолютно независимые случайные последовательности, то их сумма произведений стремится к нулю. Это говорит о том, что подобные последовательности или цифровые сигналы обладают нулевой корреляцией. Причём, чем длиннее последовательности, тем сильнее стремление результата к нулевому значению.

Корреляция может быть положительной и отрицательной. При положительной корреляции взаимосвязь процессов представляется увеличением параметра сигнала, отражающим поведение первого процесса, с одновременным увеличением параметра другого сигнала, отражающим второй процесс. При отрицательной корреляции увеличение параметра одного сигнала связано с уменьшением параметра другого сигнала.

Одним из простых и эффективных способов нахождения корреляции одномерных массивов является расчёт коэффициента Пирсона, который характеризует существование линейной взаимосвязи между двумя выборками *X* и *Y*, и расчитывается по формуле [2]:

$$r_{XY} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(X-\bar{X})^2(Y-\bar{Y})^2},$$

где $\bar{X} = \frac{1}{n}\sum_{i=0}^{n-1} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=0}^{n-1} Y_i$ – средние значения выборок, *n* – длина выборок.

В данной работе корреляционный анализ используется для поиска подобных или эталонных фигур и сигналов в контейнере – одномерном детерминированном или случайном массиве данных с учётом его зашумлённости. Для этого эталонный объект (сигнал) перемещается слева направо вдоль контейнера, и на каждой итерации вычисляется коэффициент корреляции, значение которого присваивается элементу одномерного массива с индексом, равному номеру итерации.

Коэффициент корреляции Фехнера или коэффициент знаковой корреляции является наиболее простой мерой связи между двумя одномерными массивами данных [3]. Его расчёт основан на оценке степени согласованности направлений отклонений индивидуальных значений массивов от соответствующих средних значений. Для расчёта коэффициента Фехнера вычисляются средние значения для каждой из анализируемых последовательностей, а затем находятся знаки отклонений от средних для всех значений результативного и факторного признаков. Если сравниваемое значение больше среднего, то ставится знак «+» если меньше – знак «–». Совпадение знаков по отдельным значениям последовательностей определяет согласованную вариацию, а их несовпадение – нарушение согласованности. Затем подсчитывается количество совпадений и несовпадений, и знаковая корреляция вычисляется по следующей формуле:

$$R_F = \frac{n_a - n_b}{n_a + n_b},$$

где n_a – количество совпадений знаков, n_b – количество несовпадений знаков.

При решении поставленной задачи поиска подобных и эталонных фигур в контейнере разработана программа, в которой реализована возможность выбора вида контейнерного и эталонного сигналов как одномерных массивов данных, расчёта и построения корреляционных функций. Результатом работы программы являются визуализация контейнерного сигнала и отображение графиков корреляционной и знаковой корреляционной функций.

Так, на рисунке 1 представлен тестовый сигнал-контейнер синусоидальной формы с математической моделью стандартного вида:

$$f(t) = A \cdot sin\left(\frac{2\pi}{N} \cdot \omega \cdot t + \varphi\right),$$

где $A = 3$ – амплитуда синусоиды, $\omega = 4$ – частота колебаний, $\varphi = 0$ – начальная фаза, $N = 200$ – длина контейнера как цифрового массива.

Так как на практике крайне редко встречается чистый сигнал без шумов, для получения более значимых результатов к контейнеру добавляется случайная аддитивная составляющая с нормальным законом распределения. При задании случайной составляющей необходимо учитывать стандартное отклонение и форму имеющегося «чистого» сигнала.

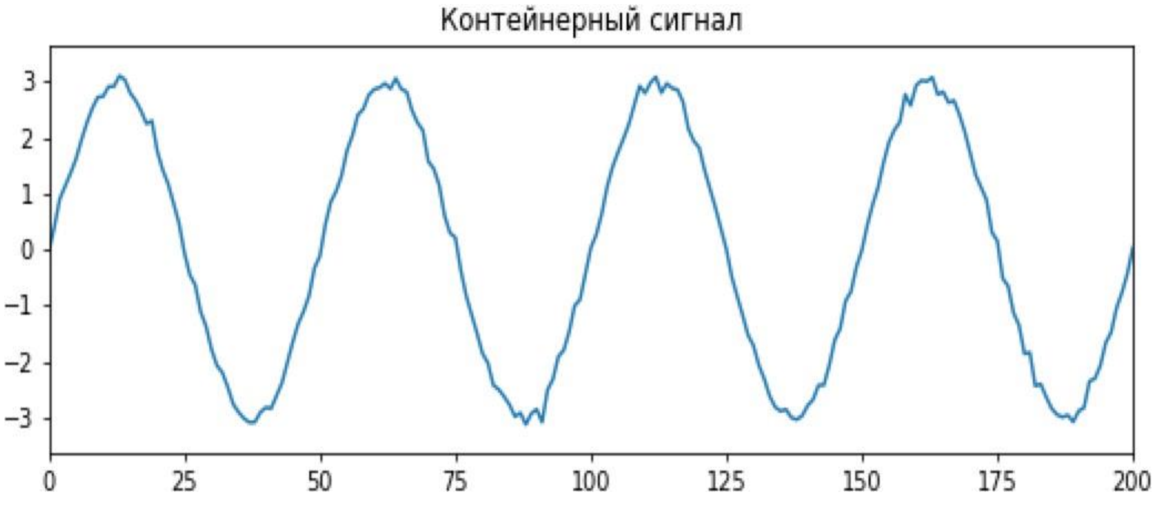

Рис 1. Визуализация контейнера с учётом зашумлённости.

Для задачи поиска подобных и эталонных фигур в контейнере разработана программа, в которой реализована возможность выбора вида контейнерного и эталонного сигналов как одномерных массивов данных, расчёта и построения корреляционных функций. Результатом работы программы являются визуализация контейнерного сигнала и отображение графиков корреляционной и знаковой корреляционной функций (см. рисунок 2).

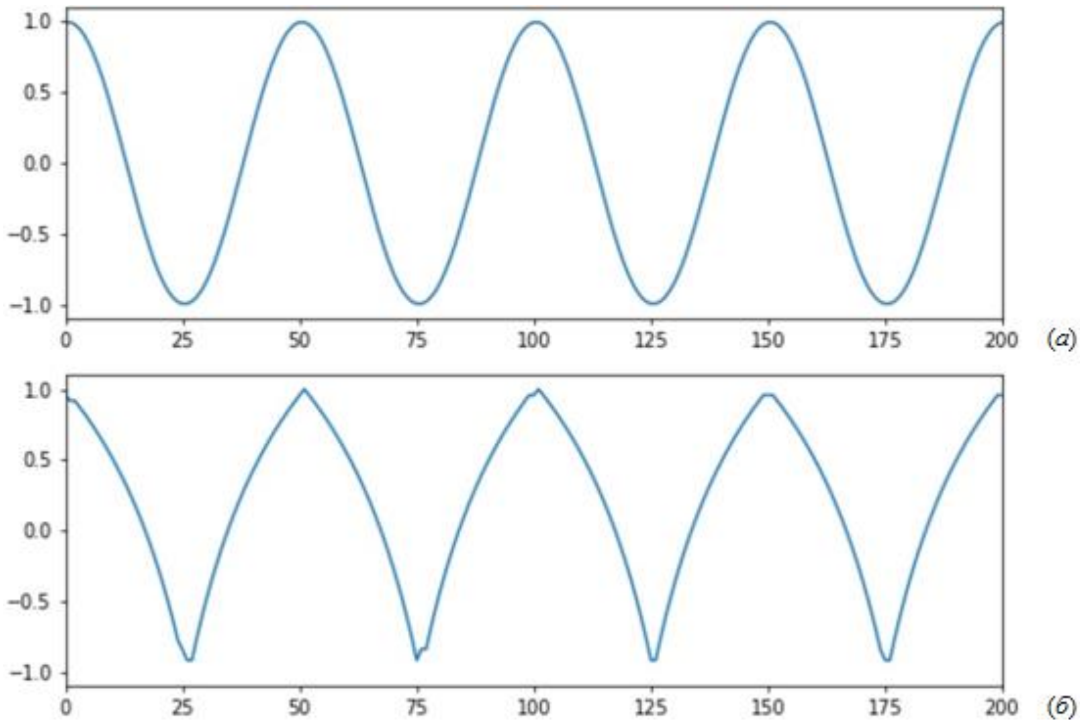

Рис 2. Результат работы программы с отображением корреляционной функции (*а*) и функции знаковой корреляции (*б*).

На рисунке 2 корреляционные функции принимают максимальные, равные единице значения при совпадении фрагментов контейнерного сигнала с эталоном-гармоникой, и близкие к –1 величины в случаях полной противоположности их формы.

По результатам статистических испытаний программы и работы в целом можно сделать вывод об эффективности и актуальности корреляционного анализа в практических задачах обработки цифровых сигналов. Коэффициент корреляции Пирсона и коэффициент знаковой корреляции имеют весомое значение в разработке информационных систем и систем анализа больших данных.

Корреляционный анализ цифровых сигналов находит широкое применение в следующих сферах деятельности:

– определение расстояния до источника звука;
– выделение полезных сигналов в фоновом шуме;
– определение импульсных характеристик электрических систем;
– реализация систем распознавания речи;
– обработка цифровых изображений;
– анализ и прогнозирование экономических процессов.

Литература

1. Фаерман, В.А. Корреляционный анализ в методах цифровой обработки сигналов / В. А. Фаерман, В. С. Аврамчук. – [Электронный ресурс]. – 2020. – Режим доступа: https://www.lib.tpu.ru/fulltext/c/2012/C04/033.pdf.
2. Коэффициент корреляции Пирсона [Электронный ресурс]. – 2020. – Режим доступа: http://www.machinelearning.ru/wiki/index.php?title=Коэффициент_корреляции_Пирсона.
3. Коэффициент корреляции Фехнера [Электронный ресурс]. – 2020. – Режим доступа: https://studfile.net/preview/5855743/page:27/.