

АНАЛИЗ МЕТОДОВ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ДАННЫХ, ИССЛЕДОВАНИЕ ВЛИЯНИЯ ШУМОВОГО ВОЗМУЩАЮЩЕГО ВОЗДЕЙСТВИЯ НА КОРРЕКТНОСТЬ ОБРАБОТКИ

Московцова А.В.

Научный руководитель – ас. Пономарева А.В.

Харьковский национальный университет радиоэлектроники
(61166, Харьков, пр. Ленина, 14, каф. ТАПР, тел (057) 702-14-86)

The analysis of the data pre-processing. The study of different methods of smoothing time series. The study of the effect of disturbances (noise) on the correct pre-treatment.

В современных информационных системах временные ряды являются широко распространенным типом данных. Временной ряд содержит информацию об особенностях и закономерностях протекания процессов.

Предварительный анализ временных рядов дает возможность определить и использовать данные закономерности для оценивания характеристик процесса в будущем (прогнозирования).

В данной работе рассматривается процедура предварительной обработки данных, представленных в виде временного ряда. Главной задачей данного этапа является получение полных, однородных, корректных данных.

Проведено исследование различных алгоритмов сглаживания временных рядов, выбор наиболее подходящих алгоритмов для различных типов входной информации, анализ влияния внешних возмущений на корректность проводимой предварительной обработки.

При исследовании использовались следующие методы сглаживания временных рядов: скользящее среднее; взвешенное скользящее среднее; экспоненциальное сглаживание.

При исследовании влияния внешнего возмущающего воздействия (белый шум) на входной синусоидальный сигнал использовались средства USB осциллографа (рис. 1). К исходному сигналу добавлялась шумовая составляющая с различными уровнями сигнал/шум от 5 до 30 дБ. Влияние шумовой составляющей на сигнал показано на Рисунке 2.

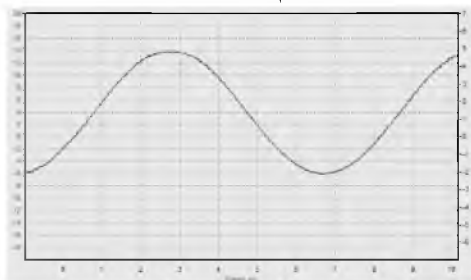


Рис. 1 – Исходный сигнал

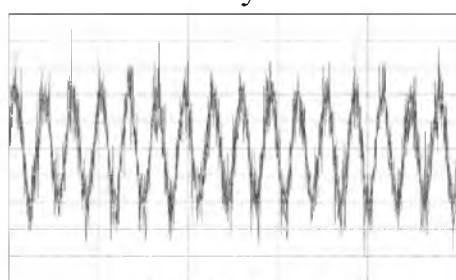


Рис. 2 – Зашумленный сигнал

Сглаживание проводилось по описанным ранее методам. Так, для методов скользящего среднего, выбрана более широкая апертура ($m=7$),

для экспоненциального – $\alpha=0,5$. Для оценки правильности сглаживания предложено рассчитывать среднюю ошибку аппроксимации (количество отклонений по каждому наблюдению, в процентном выражении). Полученные данные показаны на рисунке 3.

Предварительная обработка данных является также неотъемлемой составляющей процесса непараметрической идентификации линейных детерминированных объектов. Проведено исследование идентификации апериодического звена 1-го порядка по переходной характеристике. К переходной характеристике, которая являлась исходным сигналом, также была добавлена шумовой составляющей с различными уровнями сигнал/шум от 5 до 35 дБ.

Сглаживание проведено со следующими параметрами: для методов скользящего среднего среднее значение апертуры $m=5$, для экспоненциального сглаживания – $\alpha=0,5$. Рассчитанная средняя ошибка аппроксимации в зависимости от выбранного метода сглаживания приведена на графике (рис. 4).

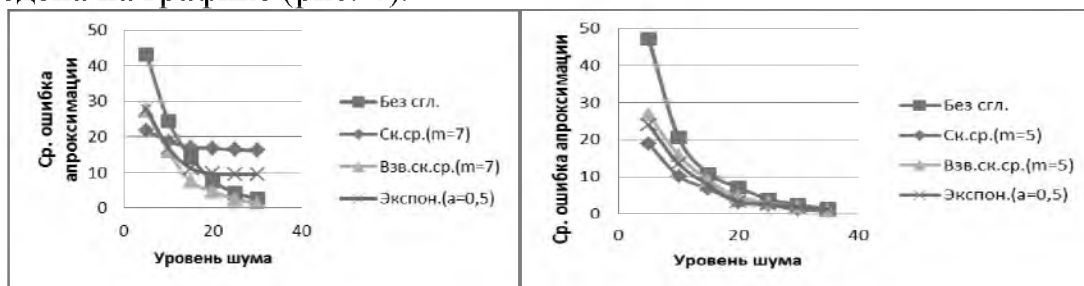


Рис.3 – Оценка сглаживания

Рис. 4 – Оценка сглаживания

По графической зависимости можно сделать выводы:

- для синусоидального сигнала наиболее подходящим методом является метод взвешенной скользящей средней.

- для переходной характеристики все использованные методы дают практически равновесные результаты, в данном случае наиболее предпочтительным является метод скользящей средней.

Стоит отметить, что при уровнях сигнал/шум менее 20дБ непараметрическая идентификация не обеспечивает заданную точность.

Для обеспечения заданной точности и корректности предварительной обработки выборка уровней временного ряда должна быть большего объема.

Литература

1. Бокс Дж., Дженкинс Г. Анализ временных рядов: прогноз и управление. Выпуск 1. М.: Мир, 1974. – 390 с.

2. Орлова И.В. «Экономико-математическое моделирование. Практическое пособие по решению задач» / И. В. Орлова; ВЗФЭИ. - М.: Вузовский учебник, 2004. - 144с.