УДК 519.6

**ИССЛЕДОВАНИЕ НЕПАРАМЕТРИЧЕСКОГО МЕТОДА ВОССТАНОВЛЕНИЯ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ С АДАПТИВНО-ОПТИМАЛЬНОЙ НАСТРОЙКОЙ**

Д.  В. Вятскин

Научный руководитель – М. К. Семёнов

Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева

Российская Федерация, 660037, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31

Е-mail: dimavyatskin@gmail.com

*В работе представлена модификация метода локальной регрессии с адаптивно-оптимальной настройкой. Её эффективность подтверждена сравнением с исходным и другими регрессионными алгоритмами.*

*Ключевые слова:* *непараметрическая регрессия, уменьшение шумов, локальный метод наименьших квадратов.*

**STUDY OF A NONPARAMETRIC METHOD FOR FUNCTIONAL DEPENDENCY RECONSTRUCTION WITH ADAPTIVE-OPTIMAL TUNING**

D. V. Vyatskin

Scientific supervisor – M. K. Semenov

ReshetnevSiberian State University of Science and Technology

31, Krasnoyarskii rabochii prospekt, Krasnoyarsk, 660037, Russian Federation

Е-mail: dimavyatskin@gmail.com

*The report presents a modification of the local regression method with adaptive-optimal tuning. Its efficiency is confirmed by comparison with the original and other regression algorithms.*

*Keywords: nonparametric regression, noise reduction, local least squares method.*

**Введение.** Все собираемые данные в той или иной степени подвержены действию случайных факторов, вызывающих шумы. Для уменьшения шумов в данных может применяться дорогостоящее высокоточное оборудование, или повторное получение экспериментальных данных, однако особый интерес представляют математические методы, для своей реализации, не требующие никаких ресурсов кроме вычислительных. В данной работе был исследован один из существующих алгоритмов устранения шумов в данных [1] с целью внесения в него модификаций, для реализации возможности его адаптивно-оптимальной настройки [2], с целью превзойти применяемые на текущий момент регрессионные методы [3] в точности и робастности [4].

**Разработанная модификация.** В данной работе используется термин плечо - половина интервала аппроксимации за вычетом точки амортизации, находящейся в центре этого интервала т.е. точки для которой мы находим оценку. От величины плеча зависит сила сглаживания, поэтому важно выбирать оптимальное плечо, не вызывающее недо- или пересглаживания. При этом, чем больший шум применён к данным, тем оптимальное плечо будет больше. Для повышения точности работы алгоритма было предложено подбирать плечо для восстановления каждой точки отдельно, т.е. сглаживать данные вектором плеч. Сглаживанием с помощью динамического плеча в данном контексте мы будем называть сглаживание с помощью вектора плеч , где -й точке соответствует одно значение плеча . Оптимальный вектор будет состоять из оптимальных плеч , которое можно определить как

*.*

Критерий минимизации при этом примет вид

где – реальное значение функции в точке , – найденная оценка для точки .

Преимущество такого подхода заключается в возможности эффективной обработки наборов данных, имеющих как сегменты, требующие сглаживания большим так и сегменты, требующие сглаживания малым плечом.

Несмотря на то, что при работе с простыми функциями, увеличение точности при использовании динамического плеча визуально не очень заметно, реальное значение отклонения может отличаться в разы, а при аппроксимации нетривиальных зависимостей, разница становится различимой даже на глаз, и зависит от скорости осцилляции, и неоднородности восстанавливаемой функции.

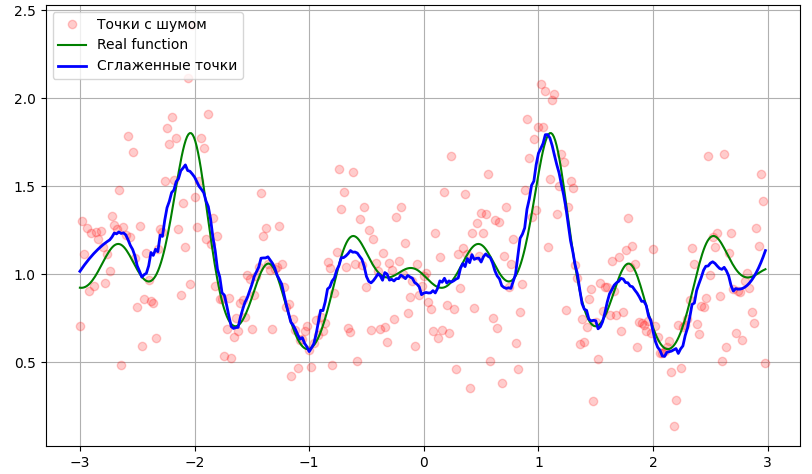
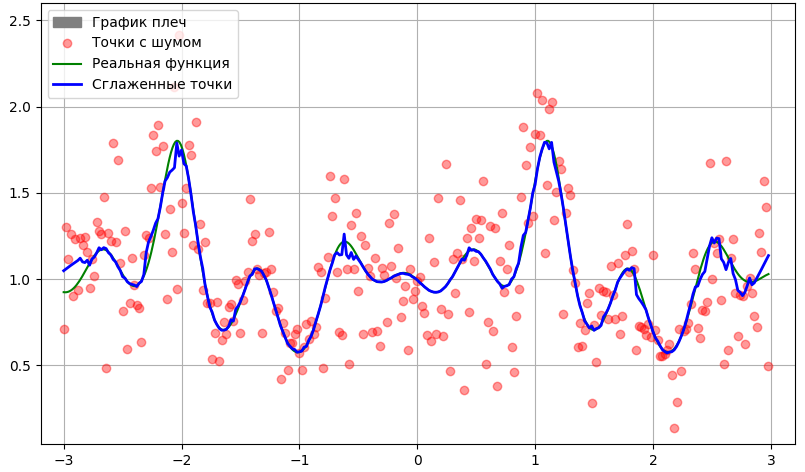
 

Рис. 1. Сравнение работы модификации с исходным алгоритмом

Точки, находящиеся на краях набора данных, не всегда могли быть хорошо сглажены оптимальным плечом так как максимальная длина плеча для них ограничивалась расстоянием до края. Для решения этой проблемы была реализована возможность смещения точки амортизации. Это решение позволило достичь более точного сглаживания точек на краях графика как для динамического, так и для фиксированного вариантов.

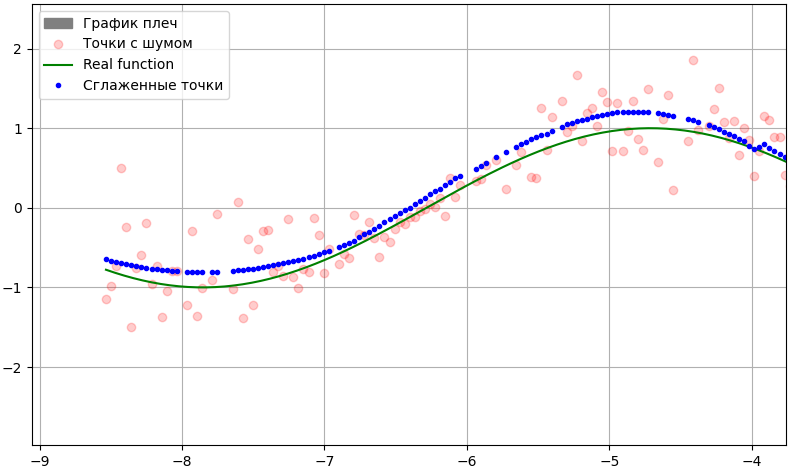
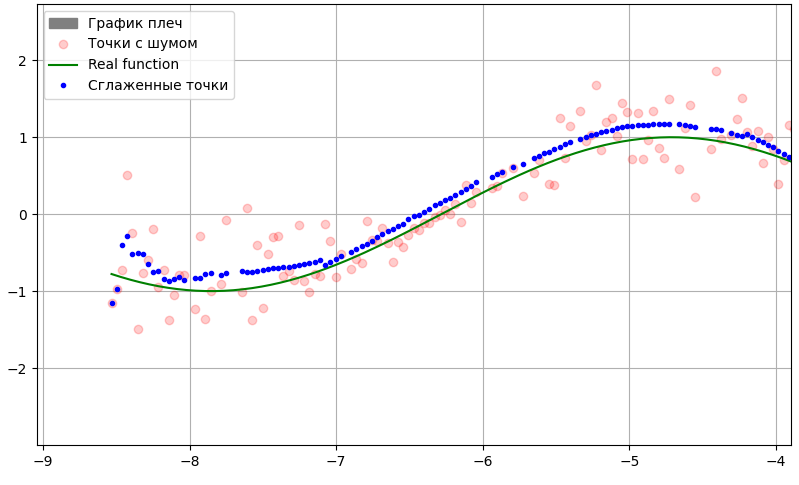


Рис. 2. Сравнение обработки края с помощью увеличения плеча и сдвига точки амортизации

Сравнение реализованного алгоритма с альтернативными методами сглаживания выявило значительные преимущества, заключающиеся в возможности адаптивно-оптимальной настройки параметров для выполнения корректного сглаживания, а также в высокой устойчивости к выбросам, с которыми большинство других методов не справляется.

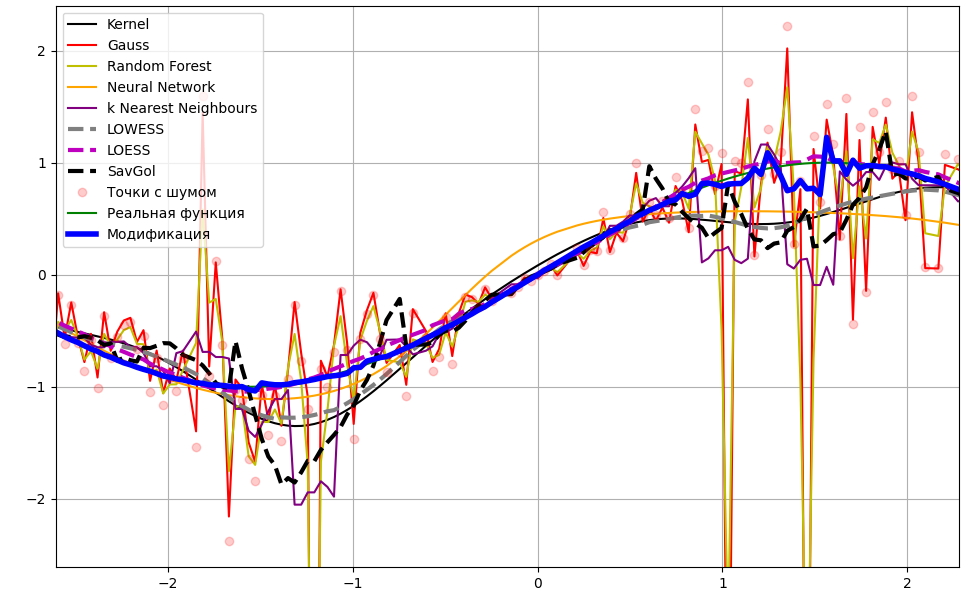


Рис. 3. Сравнение с другими алгоритмами на выборке с выбросами

**Заключение.** В процессе выполнения данной работы был разработан комплекс программного обеспечения, необходимый для качественного проведения исследований по восстановлению функциональной зависимости. Был программно реализован приведённый в статье алгоритм а так же программно реализованны все модификации непараметрической регрессии. Были исследованы факторы, влияющие на точность регрессии, важнейшим из которых является длина интервала аппроксимации. Полученные результаты позволили предложить модификации для повышения эффективности существующего алгоритма, а также заложили фундамент для более глубоких исследований.

**Библиографические ссылки**

1. Семёнов М. К. Восстановление функциональной зависимости с минимальной вычислительной структурой // Всероссийская научная конференция молодых учёных «Наука. Технологии. Инновации», НГТУ // Новосибирск, 2003.
2. Медведев А. В. «Основы теории адаптивных систем» // Монография // Красноярск, 2015.
3. Хардле В. «Прикладная непараметрическая регрессия» // Монография, перевод Назина А. В. // Москва, 1993.
4. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. «Робастность в статистике» // Монография, перевод Золотарев В. М. // Москва, 1989.

© Вятскин Д. В., 2023

**Подпись руководителя!**