УДК 004.021

ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВОГО АЛГОРИТМА ДЛЯ ОЦЕНИВАНИЯ КРИВИЗНЫ КРИВОЙ, ЗАДАННОЙ НАБОРОМ ДИСКРЕТНЫХ ТОЧЕК

Д.В. Бушмелев
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: dmitriibushmelev@mail.ru

А.В. Макаренко
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: avm.work@mail.ru

Ключевые слова: оценивание параметров, кривизна кривой, нейронные сети, метод наименьших квадратов.

Аннотация: Оценивание кривизны по набору дискретных точек, описывающих кривую, имеет существенное прикладное значение. В докладе проведено исследование нейросетевого алгоритма, функционирующего в контролируемых условиях (использован синтетический датасет). Эксперименты показали, что данный подход обеспечивает точные и надежные оценки кривизны кривой даже в случае существенного зашумления входных данных. Полученные результаты являются базисными для проведения дальнейших исследований по данной тематике.

1. Введение

Оценивание кривизны кривой, задаваемой набором дискретных точек, является важным прикладным инструментом и имеет применение в различных областях, таких как компьютерное зрение, робототехника, медицинская диагностика и т.д. Подходить к решению данной задачи можно несколькими способами. Один из них – стандартный способ – использует метод наименьших квадратов для нахождения параметров кривой, проходящей через данные точки. Другим возможным вариантом решения данной задачи является нейросетевой подход. Каждый из вариантов имеет свои преимущества и недостатки. Метод наименьших квадратов ищет решение, минимизирующее сумму квадратов разностей, в то время как нейросетевые алгоритмы, обладающие возможностью нелинейной аппроксимации (обобщающей способностью) [1] обучающей выборки, способны дать лучший результат при сильных зашумлениях данных.

Оба данных подхода были протестированы при решении задачи определения кривизны кривой на синтетических данных. Эксперименты показали, что нейросетевой подход обеспечивает более робастные оценки кривизны кривой по сравнению с методом наименьших квадратов. Продемонстрированный потенциал нейросетей для решения задачи регрессии может быть использован для дальнейших исследований и практических применений.

2. Основная часть

2.1. Описание датасета

Данные для эксперимента создаются искусственным способом. Размер датасета составляет 10 тыс. примеров, каждый из которых состоит из восьми точек, расположенных на дуге окружности некоторого радиуса. Радиус r окружности выбирается из непрерывного равномерного распределения на отрезке $[100; 10000]$. Координаты центра окружности выбираются из непрерывного равномерного распределения на отрезке $[-3000; 3000]$. Точки дуги окружности, чья градусная мера не превышает 30°, выбираются равномерно. Точки выбираются по дуге, а не по всей окружности для усложнения условий работы алгоритмов, а также в виду отсутствия прикладных задач, в которых точки распределены по всей окружности. К получившимся координатам точек добавлен гауссовский шум, среднеквадратическое отклонение которого равно 15. На обучающую часть отводится 80% данных, на тестовую – 20%.

Средняя кривизна любой дуги окружности равна обратной величие радиуса окружности и остается постоянной от точки к точке [2]. Соответственно, в качестве меток (ответов) будем использовать радиус окружности, а не кривизну кривой в явном виде. Данный подход позволит получать лучшую численную стабильность.

2.2. Метод наименьших квадратов

Уравнение окружности можно записать в векторном виде

(1) $$\begin{pmatrix} 2x & 2y & -1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ x_0^2 + y_0^2 - r^2 \end{pmatrix} = (x^2 + y^2).$$

где x, y – координаты точек окружности, x_0, y_0 – координаты ее центра.

Для точек дуги окружности с помощью уравнения (1) составим векторно-матричное уравнение вида $A\boldsymbol{x} = \boldsymbol{b}$, где $\boldsymbol{x} = (x_0 \quad y_0 \quad x_0^2 + y_0^2 - r^2)$. Вектор $\widehat{\boldsymbol{x}}$, доставляющий решение в смысле метода наименьших квадратов, должен удовлетворять

$$\min_{\boldsymbol{x}} \|A\boldsymbol{x} - \boldsymbol{b}\|^2 = \|A\widehat{\boldsymbol{x}} - \boldsymbol{b}\|^2.$$

Так как в каждом примере восемь точек, то система является переопределенной и, соответственно, имеет единственное решение. Получить радиус окружности r можно из третьей координаты вектора $\widehat{\boldsymbol{x}}$ с помощью первых двух.

2.3. Нейросетевой подход

Для решения данной задачи используется полносвязная нейронная сеть с двумя скрытыми слоями. Входной слой сети состоит из 16 нейронов, каждый из которых получает на вход одну из координат точек дуги окружности. Каждый скрытый слой содержит 64 нейрона. В качестве выхода нейронной сети получаем радиус окружности, описывающей подаваемые на вход точки.

Вышеописанная нейронная сеть имеет 5 313 обучаемых параметров. В качестве функций активации в скрытых слоях используется функция ReLU, в выходном слое – линейная функция активации. Обучение происходит в течение 40 эпох, а в качестве оптимизатора используется алгоритм Adam [3] с темпом обучения 0,001. Для обучения использовалась библиотека Keras [4].

2.4. Результаты экспериментов

На рис. 1 представлены гистограммы распределения ответов исходных данных совместно с ответами метода наименьших квадратов и с помощью нейросети.

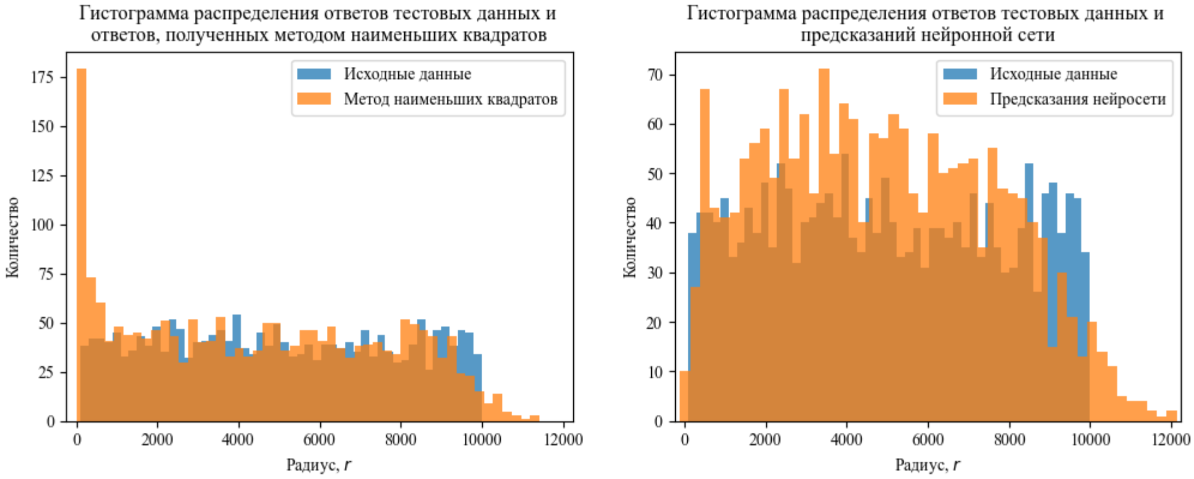

Рис. 1. Гистограммы распределения ответов тестовых данных и ответов, полученных нейросетевым алгоритмом и методом наименьших квадратов.

Как можно видеть из рис. 1, ответы, полученные с помощью метода наименьших квадратов, некорректны при маленьких радиусах окружностей. Данное явление можно объяснить существенным влиянием шума на данные с малым радиусом. Нейронная сеть с такими данными справляется достаточно хорошо.

На рис. 2 представлены диаграммы рассеивания истинных ответов и ответов, полученных с помощью метода наименьших квадратов и нейронной сети.

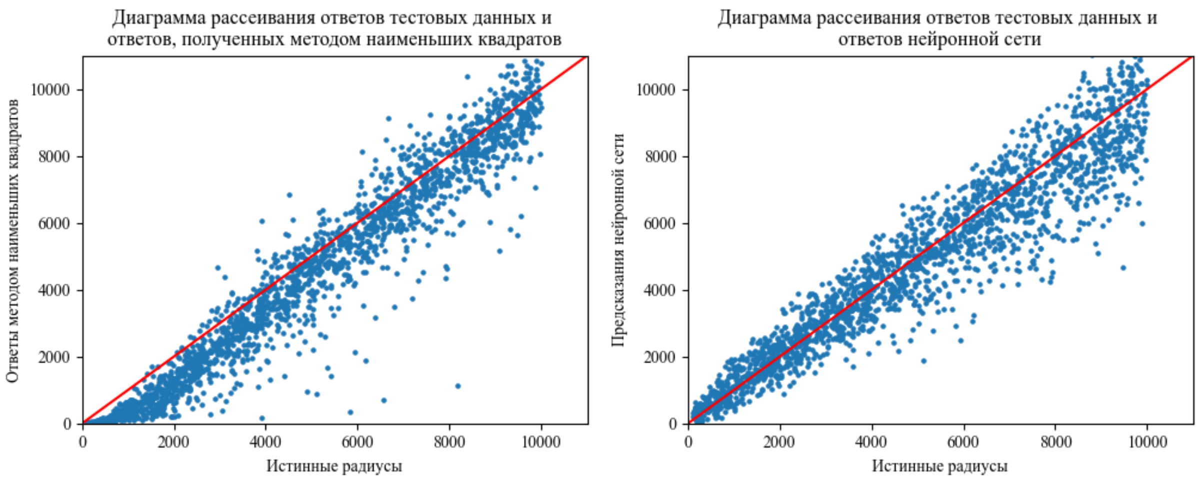

Рис. 2. Диаграмма рассеивания ответов тестовых данных и ответов, полученных методом наименьших квадратов и нейронной сети.

Как можно видеть из рисунка 3, в случае метода наименьших квадратов можно наблюдать отклонение ответов от теоретической прямой при малых радиусах окружностей. Это объясняется влиянием шумовых эффектов на данные, соответствующие малым радиусам окружностей. Для случая нейронной сети имеет место гетероскедастичность – дисперсия ответов сети увеличивается с ростом радиуса окружности. Это можно объяснить малым количеством обучающих данных, которых оказывается недостаточно для хорошей нелинейной аппроксимации.

Если предположить координаты центра окружности точно известными, то можно построить гистограммы ошибок оценок радиуса для рассматриваемых алгоритмов. Данные гистограммы представлены на рис. 3, где красным цветом обозначена медианное значение. Как можно видеть, медиана ошибок ответов нейронной сети меньше медианы ошибок ответов, получаемых методом наименьших квадратов.

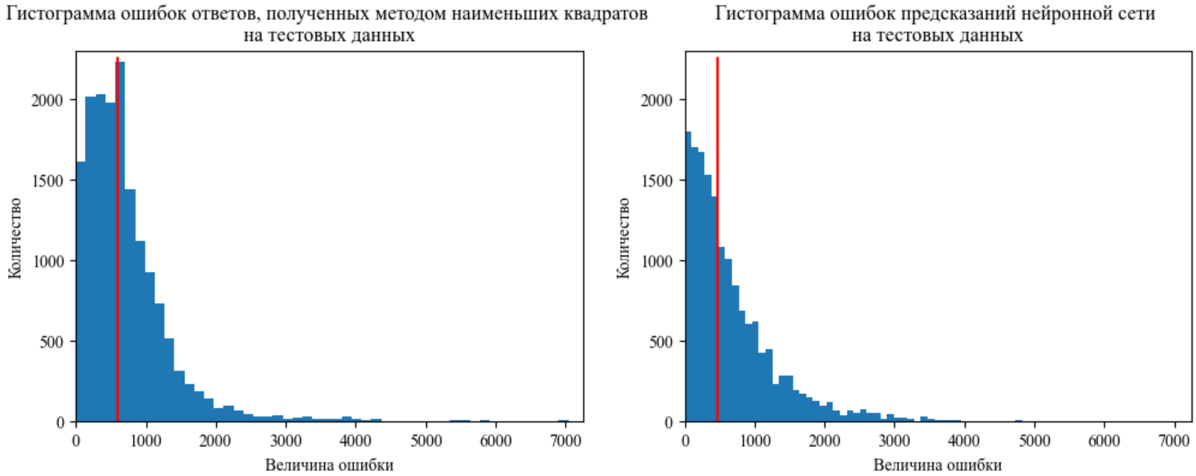

Рис. 3. Гистограммы ошибок оценки радиуса окружности методом наименьших квадратов и нейронной сетью.

На рис. 4 представлены графики зависимости квантилей ошибок ответов алгоритмов от среднеквадратического отклонения (СКО) вносимых шумов. Для каждого значения СКО, варьируемого от 0 до 250, генерируется 100 примеров. Для ответов алгоритмов на данные 100 примеров при фиксированном СКО вычисляются соответствующие квантили.

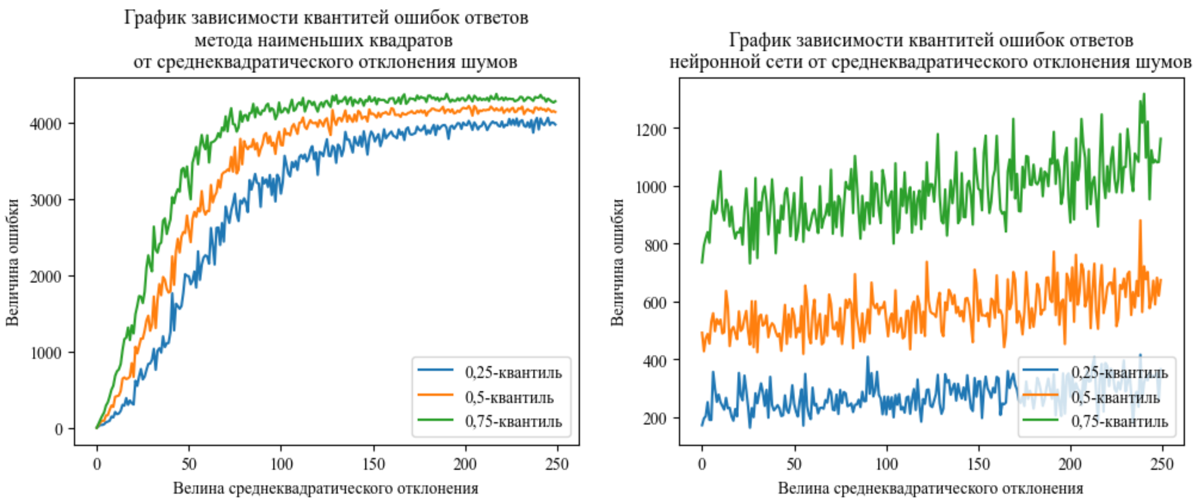

Рис. 4. Графики зависимости квантилей ошибок от среднеквадратического отклонения вносимого шума.

Как можно видеть из графиков, изображенных на рис. 4, величина ошибки метода наименьших квадратов становится постоянной при среднеквадратическом отклонении равном примерно 140. В то же время величина ошибок для нейросетевого алгоритма остается примерно постоянной вне зависимости от величины среднеквадратического отклонения вносимых шумов. Это свидетельствует о том, что нейронная сеть практически индифферентна к вносимому шуму, а также о том, что ошибка определяется иными факторами (возможно, нейросеть недообучена).

3. Заключение

Использование метода наименьших квадратов в оценивании кривизны кривой обладает своими преимуществами. Однако этот метод чувствителен к шумам в данных, что может приводить к неправильным оценкам кривизны.

Нейросетевой подход обладает способностью аппроксимировать сложные нелинейные зависимости в данных. Это позволяет получать более точные и устойчивые оценки кривизны в ситуациях с сильными зашумлениями данных. Однако нейросетевой подход требует большей вычислительной мощности и большего количества обучающих данных для достижения хороших результатов.

Тестирование на синтетических данных показало, что нейросетевой подход обеспечивает более робастные оценки кривизны кривой по сравнению с методом наименьших квадратов. Это подтверждает потенциал нейросетей в решении задачи оценивания кривизны и подталкивает к проведению дальнейших исследований и практического применения этого подхода.

Список литературы

1. DeVore R., Hanin B., Petrova G. Neural network approximation // Acta Numerica. 2021. Vol. 30. P. 327-444.
2. Рашевский П.К. Курс дифференциальной геометрии. М.: Гостехиздат, 1950. 428 с.
3. Kingma D., Ba J. Adam. A Method for Stochastic Optimization // Computing Research Repository. 2014.
4. Keras. https://keras.io (дата обращения 15.01.2024).