**УДК 519.237.5**

ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

*Е. С. Дубейко, О. Н. Канева, Д. Ч. Абдильдинова*

Омский государственный технический университет, г. Омск, Россия

*В работе был изучен и реализован обобщенный метод наименьших квадратов, проведены численные эксперименты.*

*Ключевые слова: аппроксимация, интерполяция, регрессионный анализ, регрессия, метод наименьших квадратов, матрица ковариаций.*

Метод наименьших квадратов (МНК) является одним из базовых методов [регрессионного анализа](https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D0%BE%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7) для оценки неизвестных параметров регрессионных моделей по выборочным данным[1].

(1)

Если нарушается одно из условий применимости МНК - условие отсутствия автокорреляции, то есть случайные величины ε*i*(ошибки)зависимы друг от друга, то следует применять обобщенный метод наименьших квадратов.

Обобщенный метод наименьших квадратов метод оценки параметров регрессионных моделей, являющийся обобщением классического метода наименьших квадратов. Обобщенный МНК сводится к минимизации «обобщенной суммы квадратов» остатков регрессии , где вектор остатков, симметрическая положительно определенная весовая матрица. Обычно обобщённым методом наименьших квадратов называют частный случай, когда в качестве весовой матрицы используется матрица, обратная ковариационной матрице случайных ошибок модели.

Проблема применения обобщённого МНК заключается в неизвестности ковариационной матрицы случайных ошибок. Поэтому на практике используют доступный вариант обобщенногоМНК, когда вместо *V* используется её некоторая оценка. Рассмотрено три процедуры [2]:

1. Процедура Кохрейна-Оркатта.
2. Процедура Хилдрета-Лу.
3. Процедура Дарбина.

Если случайные ошибки модели связаны авторегрессионной зависимостью первого порядка , то не учитывая первые наблюдения, преобразование заключается в следующем: из текущих значений переменных отнимаются предыдущие и умножаются на величину *r.* Величина *r*- это коэффициент автокорреляции при первом лаге.

(2)

Преобразование (2) называется авторегрессионным преобразованием. Для первых наблюдений применяется поправка Прайса - Уинстена - данные первых наблюдений умножаются на величину . Случайная ошибка полученной модели по предположению есть белый шум. Дальнейшее применение обычного метода наименьших квадратов позволит получить качественные оценки данной модели.

***Процедура Кохрейна-Оркатта***

Шаг 1. Оценка исходной модели [методом наименьших квадратов](https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BD%D0%B0%D0%B8%D0%BC%D0%B5%D0%BD%D1%8C%D1%88%D0%B8%D1%85_%D0%BA%D0%B2%D0%B0%D0%B4%D1%80%D0%B0%D1%82%D0%BE%D0%B2) и получение остатков модели.

Шаг 2. Оценка коэффициента автокорреляции остатков модели (формально её можно получить также как МНК-оценку параметра авторегрессии во вспомогательной регрессии остатков ).

Шаг 3. Авторегрессионное преобразование данных (с помощью оцененного на втором шаге коэффициента автокорреляции) и оценка параметров преобразованной модели обычным МНК.

Оценки параметров преобразованной модели и являются оценками параметров исходной модели, за исключением константы, которая восстанавливается делением константы преобразованной модели на 1*-r*. Процедура может повторяться со второго шага до достижения требуемой точности.

***Процедура Хилдрета - Лу***

В данной процедуре производится прямой поиск значения коэффициента автокорреляции, которое минимизирует сумму квадратов остатков преобразованной модели. А именно задаются значения *r* из возможного интервала (-1;1) с некоторым шагом. Для каждого из них производится авторегрессионное преобразование, оценивается модель обычным МНК и находится сумма квадратов остатков. Выбирается тот коэффициент автокорреляции, для которого эта сумма квадратов минимальна. Далее в окрестности найденной точки строится сетка с более мелким шагом и процедура повторяется заново. Процесс овторяться заново до достижения требуемой точности.

***Процедура Дарбина***

Используя авторегрессионное преобразование

. (3)

Раскрыв скобки и перенеся лаговую зависимую переменную вправо получим

. (4)

Введем обозначения , , . Тогда получим следующую модель:

, (5)

которую необходимо оценить с помощью обычного МНК.

При этом полученная оценка коэффициента [автокорреляции](https://ru.wikipedia.org/wiki/%D0%90%D0%B2%D1%82%D0%BE%D0%BA%D0%BE%D1%80%D1%80%D0%B5%D0%BB%D1%8F%D1%86%D0%B8%D1%8F) может быть использована для авторегрессионного преобразования и применения МНК для этой преобразованной модели для получения более точных оценок параметров.

Была реализована программа на языке С#. Программа реализует три процедуры: Кохрейна-Оркатта, Хилдрета-Лу и Дарбина, а также применяет обычный метод наименьших квадратов. Входные данные, которые будут использоваться для вычислений, получают из файла формата csv. Одна строка таблицы – это она строка файла. Файл должен состоять из двух столбцов, первый – значения независимой переменной *Х*, второй - значения зависимой переменной *Y*.

На основании имеющихся данных строится график зависимости*X* и *Y*. Выбираем вид функции регрессии: линейная или квадратичная. Далее по рассмотренным выше алгоритмам после выбора необходимой процедуры производим вычисления. На рисунках 1 - 3 показаны результаты работы программы для каждой отдельной процедуры.

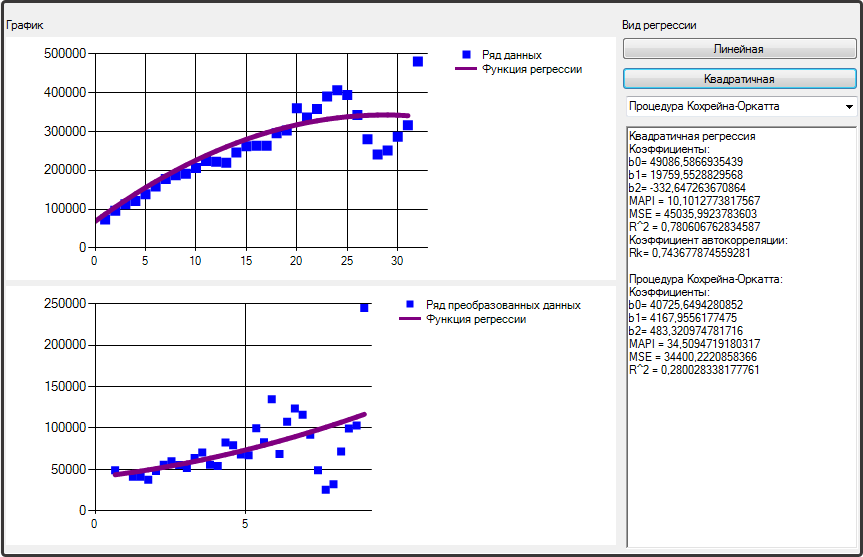


Рисунок 1 – Результат работы программы для процедуры Кохрейна-Оркатта

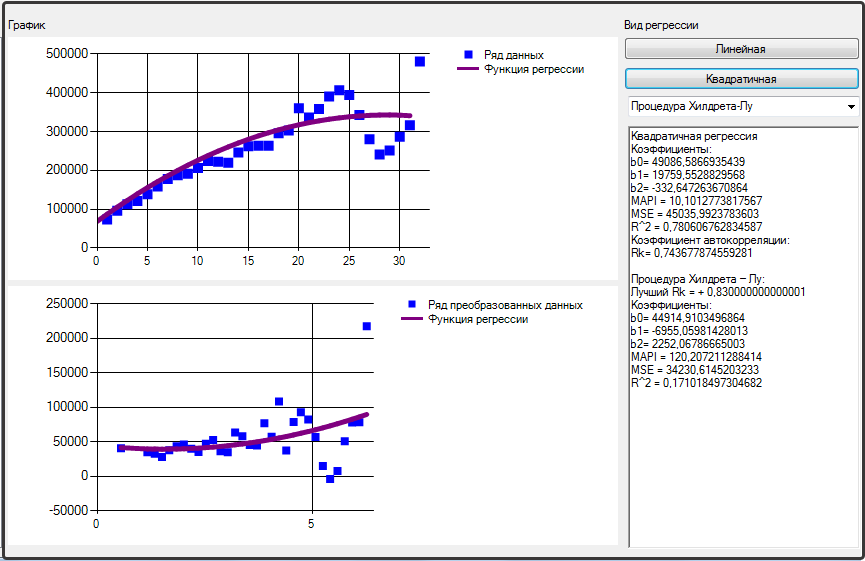


Рисунок 2 - Результат работы программы для процедуры Хилдрета-Лу

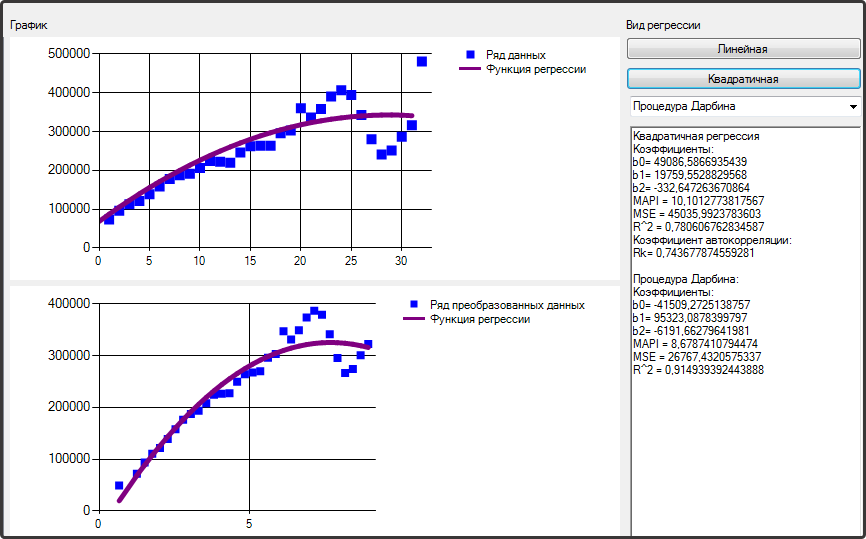


Рисунок 3 - Результат работы программы для процедуры Дарбина

Выходными данными являются параметры функции регрессии, среднеквадратическая ошибка *MSE*, средняя ошибка аппроксимации *MAPI* и коэффициент детерминации *R*2. Оценки *MSE*и *R*2 необходимы для исследования качества полученной модели, а оценка *MAPI* - точности прогнозной модели. Построенное уравнение регрессии можно считать удовлетворительным, если величина *MAPI*не превышает 8-10%.

Были проведены численные эксперименты, которые позволили сделать вывод об эффективности использования этих процедур на наборах данных, в которых присутствует автокорреляция.

**Библиографический список**

1. Дрейпер, Н. Прикладной регрессионный анализ[Текст]: пер. с англ. Ю. П. Адлером, В. Г. Горским. / Н. Дрейпер, Г. Смит. – книга 2, 2-е изд. – М. :Финансы и статистика, 2012. – 304 с.
2. Грачева, М. В. Моделирование экономических процессов [Текст]: учеб. пособие. / М. В. Грачева, Л. Н. Фадеева, Ю. Н. Черемных. – М.: Юнити, 2005. – 351 с.