

УДК 678.04

DOI: [10.26102/2310-6018/2020.28.1.042](https://doi.org/10.26102/2310-6018/2020.28.1.042)

Сравнение точности аппроксимации экспериментальных данных методом наименьших относительных квадратов с методом наименьших квадратов

А.Б. Голованчиков¹, М.К. Доан², А.В. Петрухин³, Н.А. Меренцов⁴

Федеральное государственное бюджетное образовательное учреждение высшего образования «Волгоградский государственный технический университет», Волгоград, Российская Федерация

Резюме: Представлены результаты сравнения точности аппроксимации экспериментальных или табличных данных, полученной с использованием типового метода наименьших квадратов (МНК) и предлагаемого метода наименьших относительных квадратов (МНОК), на примере заданной табличной зависимости вязкости водноглицеринового раствора от массовой концентрации глицерина. Показывается преимущество последнего как по сумме и средним значениям локальных относительных отклонений расчетных данных вязкости искомого раствора полученных МНОК, с аналогичными данными, полученными типовым МНК, так и по наибольшим значениям этих относительных отклонений. Так, рассчитанные с использованием МНК средние относительные отклонения теоретических значений вязкости водного раствора глицерина от заданных табличных, по абсолютной величине равны 12,9 %, МНОК 5,8 %, то есть, ниже в 2 раза. Соответственно наибольшие относительные отклонения в МНК составляют 17,9 %, а МНОК – 10,6 %, то есть снижаются на 68 %. Предлагается определять условные значения параллельных опытов по экспериментальным данным основного опыта. Для этого расчет условных численных значений i -ого параллельного опыта определяется методом кусочной линейной аппроксимации $i-1$ и $i+1$ численных значений основного опыта или табличных данных. Проводится корреляционный анализ с определением коэффициентов корреляции, воспроизводимости, адекватности и значимости коэффициентов полученного уравнения регрессии.

Ключевые слова: линеаризация, аппроксимация, абсолютные и относительные отклонения, МНК и МНОК, коэффициент корреляции, воспроизводимость, адекватность, значимость.

Для цитирования: Голованчиков А.Б., Доан М.К., Петрухин А.В., Меренцов Н.А. Сравнение точности аппроксимации экспериментальных данных методом наименьших относительных квадратов с методом наименьших квадратов. *Моделирование, оптимизация и информационные технологии*. 2020;8(1). Доступно по: https://moit.vivt.ru/wp-content/uploads/2020/02/GolovanchikovSoavtors_1_20_2.pdf DOI: 10.26102/2310-6018/2020.28.1.042

Comparison of the accuracy of experimental data approximation using the least relative squares method with the least squares method

A.B. Golovanchikov, M.K. Doan, A.V. Petrukhin, N.A. Merentsov
*Federal State Budgetary Educational Institution of Higher Education
"Volgograd State Technical University", Volgograd, Russian Federation*

Abstract: The results of comparing the accuracy of approximation of experimental or tabular data obtained using the standard method of least squares (LSM) and the proposed method of least relative

squares (LRSM), for example, a given table dependence of the viscosity of a water-glycerine solution on the mass concentration of glycerol. The advantage of the latter is shown as the sum and average values of the local relative deviation of calculated data of viscosity of the desired solution obtained by LRSM, with similar data obtained by standard LSM and maximum values of these relative deviations. So, calculated using LSM average relative deviations of theoretical viscosity of an aqueous solution of glycerin from the specified table, in absolute value equal to 12.9%; LRSM of 5.8%, i.e., below 2 times. Accordingly the largest relative deviations in the LSM are 17.9%, and LRSM – 10.6 %, that is, reduced by 68%. It is proposed to determine the conditional values of parallel experiments based on the experimental data of the main experiment. To do this, the calculation of conditional numerical values of the i -th parallel experience is determined by the method of piecewise linear approximation of $i-1$ and $i+1$ numerical values of the main experience or table data. A correlation analysis is performed to determine the correlation coefficients, reproducibility, adequacy, and significance of the coefficients of the resulting regression equation.

Keywords: linearization, approximation, absolute and relative deviations, LSM and LRSM, correlation coefficient, reproducibility, adequacy, significance.

For citation: Golovanchikov A.B., Doan M.C., Petrukhin A.V., Merentsov N. A. Comparison of the accuracy of experimental data approximation using the least relative squares method with the least squares method. *Modeling, Optimization and Information Technology*. 2020;8(1). Available from: https://moit.vivt.ru/wp-content/uploads/2020/02/GolovanchikovSoavtors_1_20_2.pdf DOI: 10.26102/2310-6018/2020.28.1.042 (In Russ).

Введение

Известный типовой метод наименьших квадратов (МНК) предполагает возможность линеаризации исходного аналитического уравнения линейным алгебраическим уравнением:

$$y_i = b + a \cdot x_i, \quad (1)$$

и определением его коэффициентов, исходя из условия:

$$U = \sum_{i=1}^m (b + a \cdot x_i - y_i)^2 \rightarrow \min. \quad (2)$$

Однако затем обычно для определения точности аппроксимации оцениваются локальные относительные отклонения:

$$\delta_i = (b + a \cdot x_i - y_i) / y_i \quad (3)$$

как наибольшие, так и средние по абсолютной величине и по ним делается вывод о возможности описания заданной экспериментальной или табличной зависимости выбранным аналитическим уравнением. Поэтому предлагается вместо МНК при минимизации функции U на основании выражения (2) применять МНОК с минимизацией функции:

$$U_1 = \sum_{i=1}^m [(b + a \cdot x_i - y_i) / y_i]^2 \rightarrow \min \quad (4)$$

и сравнить локальные относительные отклонения теоретических значений от экспериментальных или табличных обоими методами по уравнению (3).

Целью работы является расчет сравнение точности аппроксимации экспериментальных или табличных данных с их значениями, полученными соответственно типовым методом наименьших квадратов (МНК) и предлагаемым методом наименьших относительных квадратов (МНОК) на примере заданной табличной зависимости вязкости водноглицеринового раствора от массовой концентрации глицерина.

Материалы и методы

Для примера в Таблице 1 приведены табличные значения вязкости водноглицеринового раствора от массовой концентрации глицерина (строки 1÷3).

Уравнения для расчета коэффициентов a и b уравнения (1) для функции минимизации U МНК по формуле (2) общеизвестны [1,2].

Уравнения для расчета этих коэффициентов для функции минимизации U_1 МНОК по формуле (3) приведены ниже [3]:

$$\left\{ \begin{aligned} Z &= \sum_{i=1}^m \left(\frac{1}{y_i} \right)^2 \cdot \sum_{i=1}^m \left(\frac{x_i}{y_i} \right)^2 - \left[\sum_{i=1}^m \left(\frac{x_i}{y_i^2} \right) \right]^2; \\ a &= \frac{\sum_{i=1}^m \left(\frac{1}{y_i} \right)^2 \cdot \sum_{i=1}^m \left(\frac{x_i}{y_i} \right) - \sum_{i=1}^m \left(\frac{1}{y_i} \right) \cdot \sum_{i=1}^m \left(\frac{x_i}{y_i^2} \right)}{Z}; \\ b &= \frac{\sum_{i=1}^m \left(\frac{1}{y_i} \right) \cdot \sum_{i=1}^m \left(\frac{x_i}{y_i} \right)^2 - \sum_{i=1}^m \left(\frac{x_i}{y_i} \right) \cdot \sum_{i=1}^m \left(\frac{x_i}{y_i^2} \right)}{Z}. \end{aligned} \right. \quad (5)$$

Необходимость описания вязкости водноглицеринового раствора непрерывной функцией была связана с нашими экспериментальными исследованиями по ультрафильтрации этого раствора на полупроницаемой мембране УАМ-100П с кратностью концентрирования ретанта до $K = 5$. При начальной концентрации исходного водноглицеринового раствора $C_n = 0,1 \text{ кгА/кг(А+В)}$, конечная концентрация ретанта $C_k = 0,5 \text{ кгА/кг(А+В)}$, что и учтено в Таблице 1.

Коэффициенты a и b , рассчитанные по уравнениям, соответствующим МНК и уравнениям системы (5) для МНОК, приведены ниже:

$$\text{МНК: } \begin{cases} a = 4,684 \\ b = -0,387 \end{cases}; \quad \text{МНОК: } \begin{cases} a = 3,83 \\ b = -0,131 \end{cases}$$

При этом аналитическая формула имеет вид:

$$\mu = K \cdot n^c \quad (6)$$

Тогда для

$$\text{МНК: } \begin{cases} K = 0,772 \\ n = 66,9 \end{cases}; \quad \text{МНОК: } \begin{cases} K = 0,898 \\ n = 38,32 \end{cases}.$$

где μ - вязкость водоглицеринового раствора, $\text{мПа}\cdot\text{с}$;
 K и n - коэффициенты.

Надо отметить, что попытка описать табличные значения вязкости водноглицеринового раствора от концентрации глицерина (строчки 2 и 3 Таблицы 1) давали локальные относительные отклонения в 100 и более процентов для степенной функции $\mu = K \cdot c^n$.

Уравнение (6) при логарифмировании обеих его частей линеаризуется в виде:

$$\ln(\mu) = \ln(K) + c \cdot \ln(n). \quad (7)$$

Тогда с учетом линейной зависимости (1):

$$x = c; y_t = \ln(\mu); b = \ln(K); a = \ln(n)$$

Для проведения корреляционного анализа, связанного с определением воспроизводимости параллельных опытов и сравнением расчетного критерия Кохрена с

$$S_{ao}^2 = \frac{m}{N-l} \sum_{i=1}^N \left(\frac{\mu_{t(i)}}{\mu_{c(i)}} - 1 \right)^2, \quad (11)$$

так как минимизируется сумма квадратов относительных отклонений теоретических значений концентрации от средних экспериментальных ее значений.

Кроме того, теоретические значения вязкости $\mu_{t(i)}$, рассчитанные МНК и МНОК также являются различными, так как коэффициенты аппроксимирующих уравнений K и n численно не равны. Последнее отражается и в проверке значимости этих коэффициентов по критерию Стьюдента.

Результаты

Таблица 1 – Заданные табличная и расчетная зависимости вязкости водноглицеринового раствора от массовой концентрации глицерина

Table 1-Specified table and calculated dependences of the viscosity of water glycerine solution on the mass concentration of glycerine

Наименование параметра	Величина параметра				
	1	2	3	4	5
1. Номер точки					
2. Концентрация глицерина в растворе, кгА/кг(В+А)	0,1	0,2	0,3	0,4	0,5
3. Табличное значение вязкости раствора, мПа·с (сПз)	1,311	1,769	2,501	3,75	6,05
4. Значения вязкости раствора, рассчитанные МНК, мПа·с	1,085	1,733	2,768	4,422	7,065
5. Относительные отклонения расчетных значений вязкости раствора от табличных, % (МНК)	-17,96	-2,05	10,69	17,93	16,78
6. Значения вязкости раствора, рассчитанные МНОК, мПа·с	1,286	1,887	2,767	4,095	5,953
7. Относительные отклонения расчетных значений вязкости раствора от табличных, % (МНОК)	-1,89	6,65	10,64	8,23	-1,6

Таблица 2 – Заданные табличные значения и условные значения вязкости раствора в параллельных опытах, рассчитанные по формуле (8), средние значения вязкости раствора и относительные отклонения

Table 2-Set table values and conditional values of the viscosity of the solution in parallel experiments, calculated using the formula (8), average values of the viscosity of the solution and relative deviations

№	Наименование параметра	Величина				
		1	2	3	4	5
1	Номер точки					
2	Концентрация глицерина в растворе, кгА/кг(А+В)	0,1	0,2	0,3	0,4	0,5
3	Табличное значение вязкости раствора, мПа·с (сПз)	1,311	1,769	2,501	3,75	6,05
4	Условное значение вязкости раствора, рассчитанное по формуле (7), мПа·с	1,36	1,906	2,76	4,276	7,355

5	Среднее значение вязкости раствора, рассчитанное по формуле (8), <i>мПа·с</i>	1,336	1,837	2,63	4,013	6,703
6	Значения вязкости раствора, рассчитанные МНК, <i>мПа·с</i>	1,188	1,85	2,802	4,49	6,996
7	Относительные отклонения расчетных значений вязкости раствора от средних, % (МНК)	-11,07	0,69	9,59	11,9	4,37
8	Значения вязкости раствора, рассчитанные МНОК, <i>мПа·с</i>	1,317	1,933	2,836	4,162	6,108
9	Относительные отклонения расчетных значений вязкости раствора от средних, % (МНОК)	-1,37	5,19	7,83	3,72	-8,88

Из результатов расчетов, приведенных в Таблице 2 видно, что наибольшие значения вязкости раствора, рассчитанные МНОК по сравнению с МНК уменьшаются в 1,3 раза, а средние с 7,52% до 5,4%, то есть в 1,4 раза, при этом численные значения коэффициентов K и n аппроксимирующего уравнения (6) равны:

$$\text{для МНК: } \begin{cases} K = 0,762 \\ n = 84,194 \end{cases}; \text{ для МНОК: } \begin{cases} K = 0,897 \\ n = 46,3 \end{cases}.$$

В Таблице 3 приведены результаты корреляционного анализа с учетом результатов двух параллельных опытов (табличных значений вязкости μ_i , условных ее значения $\mu_{n(i)}$ и средних значений $\mu_{c(i)}$).

Так как в формулах расчета коэффициента корреляции и воспроизводимости нет сравнения теоретических и заданных значений концентрации, то эти расчеты не зависят от метода МНК или МНОК и численно совпадают.

Таблица 3 – Сравнение коэффициентов корреляции, критериев, воспроизводимости в параллельных опытах, адекватности аппроксимирующего уравнения (6) экспериментальным данным и значимости коэффициентов этого уравнения, рассчитанных МНК и МНОК (число параллельных опытов, $p = 2$; число основных опытов $N = 5$; число коэффициентов уравнения (6), $m = 2$)

Table 3-Comparison of correlation coefficients, criteria, reproducibility in parallel experiments, the adequacy of the approximating equation (6) to experimental data, and the significance of the coefficients of this equation calculated by the OLS and IOC (the number of parallel experiments, $p = 2$; the number of main experiments $N = 5$; the number of coefficients of the equation (6), $m = 2$)

№	Наименование параметра	Величина параметра		
		МНК	МНОК	Табличное значение [1]
1	Коэффициент корреляции	0,992	0,992	-
2	Воспроизводимость в параллельных опытах	0,449	0,449	0,78 (Критерий Кохрена)
3	Адекватность аппроксимирующего уравнения (5)	1,51	1,38	5,41 (Критерий Фишера)
4	Значимость коэффициентов: K n	304,8 33658	728,3 37568	2,57 (Критерий Стьюдента)

Обсуждение

Как видно из результатов расчетов, приведенных в Таблице 1 (строчки 5 и 7), средние относительные отклонения теоретических значений вязкости от заданных табличных составляют 12,9% в МНК против 5,8% в МНОК, то есть МНОК снижает их по сравнению МНК в 2,23 раза, при этом 4 локальных значения относительных отклонений в МНК больше 10%, а в МНОК только одно из 5. Наибольшее относительное отклонение в МНК составляет 17,93%, а в МНОК – 10,64%, то есть последнее в 1,68 раза меньше.

Как видно из расчетных параметров регрессионного анализа, приведенных в Таблице 3, при равенстве коэффициентов корреляции и воспроизводимости в параллельных опытах по критерию Кохрена, адекватность аппроксимирующего уравнения (6) в МНОК выше, чем при расчетах по типовому МНК (так как адекватность тем выше, чем меньше ее расчетное значение по сравнению с табличным значением критерия Фишера), а значимость коэффициентов K и n в МНОК выше, чем при расчетах по типовому МНК (так как значимость коэффициентов тем выше, чем сильнее их расчетное значение отличается от табличного значения критерия Стьюдента).

Результаты, приведенные в Таблице 3, подтверждают высокую корреляционную связь r в зависимости вязкости водноглицеринового раствора от концентрации глицерина ($r \rightarrow 1$) и воспроизводимости в параллельных опытах (расчетное значение критерия Кохрена $Ko = 5,41$), тоже касается адекватности аппроксимирующего уравнения (6), но здесь уже оба расчетных значения числа Фишера меньше табличного значения. Высока значимость рассчитанных коэффициентов K и n , которые значительно больше для обоих методов табличного числа Стьюдента.

Заключение

Таким образом, применение МНОК для аппроксимации табличных или экспериментальных данных с определением коэффициентов линеаризованного уравнения (1) по сравнению с МНК позволяет снизить средние значения относительных отклонений теоретических значений аппроксимирующей функции от экспериментальных или табличных в 2 и более раз, а наибольшие относительные отношения уменьшить на 60-70%. Хотя расчеты по обоим методам: типовому МНК и рассматриваемому в статье МНОК обеспечивают высокую корреляционную связь, воспроизводимости в параллельных опытах, адекватность и значимость коэффициентов аппроксимирующего уравнения (6), МНОК обеспечивает повышение точности расчетов средних и наибольших локальных относительных отклонений на 30%, а средних на 40%, повышение адекватности аппроксимирующего уравнения (6) на 8,6% (см. строку 2 в Таблице 3), при этом увеличивается значимость коэффициентов “ K ” в 2,38 раза, а “ n ” на 11,6%.

Благодарности

Статья выполнена по гранту Президента РФ в рамках научного проекта № МК-1287.2020.8.

ЛИТЕРАТУРА

1. Бондарь А.Г., Статюха Г.А. *Планирование эксперимента в химической промышленности (Основные положения, примеры и задачи)* .– Киев: Вища школа, 1976.
2. Закгейм А.Ю. *Введение в моделирование химико-технологических процессов* – М.: Химия, 1973.
3. Голованчиков А.Б., Тябин Н.В. Метод интерполирования экспериментальных данных при определении реологических свойств жидкостей .– *Инженерно-физический журнал*, 1981; XLI(1):70-73.
4. *Краткий справочник физико-химических величин* . Под редакцией Равделя А.А. и Пономаревой А.М. – Л.: Химия, 1983.
5. Голованчиков А.Б., Доан Минь Кыонг, Дулькин Т.А. Свид. о гос. регистрации программы для ЭВМ № 2018613318 от 7 марта 2018 г. Российская Федерация. *Программа для расчёта параметров линейного уравнения методом наименьших относительных квадратов* . ВолгГТУ. - 2018.
6. Ахназарова С.Л., Кафаров В.В. *Оптимизация эксперимента в химии и химической промышленности* .– М.: Высшая школа, 1978, 319 с.
7. Голованчиков А.Б., Доан Минь Кыонг, Дулькин Т.А. Свид. о гос. регистрации программы для ЭВМ № 2018613321 от 7 марта 2018 г. Российская Федерация. *Программа для расчёта критериев регрессионного анализа линейного уравнения* . ВолгГТУ. - 2018.
8. Голованчиков А.Б., Доан Минь Кыонг, Шибитова Н.В. Аппроксимация экспериментальных данных методом наименьших квадратов и методом наименьших относительных квадратов. *Энерго- и ресурсосбережение: промышленность и транспорт*. 2019;26(1):42-44.

REFERENCES

1. Bondar A. G. *Planning an experiment in the chemical industry (Main provisions, examples and tasks)* . Bondar A. G., Statyukha G. A.-Kiev: Vyshcha SHKOLA, 1976.
2. Zakheim A. Yu. *Introduction to modeling of chemical and technological processes*-Moscow: Chemistry, 1973.
3. Golovanchikov A. B., Tyabin N. V. Method of interpolating experimental data in determining the rheological properties of liquids .– *Engineering and physical journal*, 1981; XLI (1):70-73.
4. *Brief reference of physical and chemical quantities* . edited by Ravdel A. A. and Ponomareva A. M.-L.: Chemistry, 1983.
5. .Golovanchikov A. B., Doan min Kyong, Dulkin T. A. SVID. about the state registration of the computer program No. 2018613318 of March 7, 2018, the Russian Federation. *Program for calculating the parameters of a linear equation using the method of least relative squares*. Volgstu. - 2018.
6. Akhnazarova S. L., Kafarov V. V.-*Optimization of experiment in chemistry and chemical industry* . M.: Higher school, 1978.
7. Golovanchikov A. B., Doan Minh Quang, T. A. Dulkin. SVID. about the state registration of the computer program No. 2018613321 of March 7, 2018, the Russian Federation. *The program for the calculation of the criteria of the regression analysis linear equation* . VSTU. - 2018.
8. Golovanchikov A. B., Doan min Kyong, Shibitova N. V. Approximation of experimental data using the least squares method and the least relative squares method . *Energy and resource saving: industry and transport*. 2019;26(1):42-44.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Голованчиков Александр Борисович, д-р. тех. наук, профессор, кафедра «Процессы и аппараты химических и пищевых производств», ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, Российская Федерация.

e-mail: pahp@vstu.ru

ORCID: [0000-0002-4576-4229](https://orcid.org/0000-0002-4576-4229)

Доан Минь Кьонг, аспирант, кафедра «Процессы и аппараты химических и пищевых производств», ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, Российская Федерация.

e-mail: victorydoan@yandex.ru

ORCID: [0000-0001-7986-938X](https://orcid.org/0000-0001-7986-938X)

Меренцов Николай Анатольевич, к.т.н., доцент, кафедра «Процессы и аппараты химических и пищевых производств», ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, Российская Федерация.

e-mail: pahp@vstu.ru

ORCID: [0000-0001-5946-0667](https://orcid.org/0000-0001-5946-0667)

Петрухин Алексей Владимирович, к.т.н., доцент, кафедра «Системы автоматизированного проектирования и поискового конструирования», ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, Российская Федерация.

e-mail: cad@vstu.ru

ORCID: [0000-0002-5356-8789](https://orcid.org/0000-0002-5356-8789)

Alexander B. Golovanchikov, Dr. Tech. Sci., Professor, "Processes and devices of chemical and food productions» Department, Federal State Budgetary Educational Institution of Higher Education "Volgograd State Technical University", Volgograd, Russian Federation

Cuong M. Doan, PhD Student, "Processes and devices of chemical and food productions» Department, Federal State Budgetary Educational Institution of Higher Education "Volgograd State Technical University", Volgograd, Russian Federation

Nykolay A. Merentsov, PhD, Assistant Professor, "Processes and devices of chemical and food productions» Department, Federal State Budgetary Educational Institution of Higher Education "Volgograd State Technical University", Volgograd, Russian Federation

Alexey V. Petrukhin, PhD, Assistant Professor, «Computer-aided design and search engineering systems» Department, Federal State Budgetary Educational Institution of Higher Education "Volgograd State Technical University", Volgograd, Russian Federation