УДК 004.421

НАХОЖДЕНИЕ ПРИБЛИЖАЮЩЕЙ ФУНКЦИЙ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

Мансуров Кубанычбек Топчубаевич – к.ф.-м.н.,
mansurov2002@inbox.ru
Жулев Владислав Александрович– магистрант,
zhulev @gmail.ru
Абдухалимов Илхомжон Ибрахимович
магистрант, abduxalimov@gmail.ru
ОшТУ им. М.М.Адышева,
г. Ош, Кыргызская Республика

Аннотация: *Отличительной чертой физических задач от других предметных задач является то, что связь неизвестного с известным реализуется не только на основе физических законов, но и с помощью соответствующих экспериментов. В статье рассмотрен метод наименьших квадратов, приведены примеры обработки экспериментальных данных наиболее часто встречающихся в процессе физического практикума. К настоящему времени разработано несколько методов обработки результатов измерений. Наиболее употребительным, простым и обоснованным является метод наименьших квадратов (МНК).*

Ключевые слова: *метод наименьших квадратов (МНК), нахождение приближающей функций, функциональная зависимость, параметр, эксперимент, погрешность.*

ЭҢ КИЧИНЕ КВАДРАТТАР УСУЛУ МЕНЕН ЖАКЫНДАТУУ ФУНКЦИЯСЫН ТАБУУ

Мансуров Кубанычбек Топчубаевич – ф.-м.и.к.,
mansurov2002@inbox.ru
Владислав Александрович Жулев – магистрант,
zhulev @gmail.ru
Илхомжон Ибрахимович Абдухалимов
магистрант, abduxalimov@gmail.ru
М.М. Адышев атандагы ОшТУ,
Ош шаары, Кыргыз Республикасы

Аннотация: *Физикалык маселелердин башка предметтик маселелерден мүнөздүү өзгөчөлүгү- белгисиз менен белгилүүнүн ортосундагы байланышты табуу, физикалык закондордун гана негизинде эмес, тиешелүү эксперименттин жардамында ишке ашуусунда турат. Макалада эң кичине квадраттар ыкмасы каралат, физикалык практикада эң көп кездешкен эксперименталдык маалыматтарды иштетүүнүн мисалдары келтирилген. Бүгүнкү күнгө чейин, өлчөө натыйжаларын иштетүү үчүн бир нече ыкмалары иштелип чыккан. Эң кеңири таралган, жөнөкөй усулдардын бири болуп эң кичине квадраттар усулу саналат (ЭКУ).*

Ачкыч сөздөр: *эң кичине квадраттар усулу (ЭКУ), жакындатуу функциясын табуу, функционалдык көз карандылык, параметр, эксперимент, катачылык.*

FINDING THE APPROXIMATION FUNCTION BY THE LEAST SQUARE METHOD

Mansurov Kubanychbek Topchubaevich
Candidate of Physical and Mathematical Sciences,
mansurov2002@inbox.ru
Vladislav Alexandrovich Zhulev – master student
zhulev@gmail.ru

Abdukhalimov Ilhomjon Ibrahimovich master student, abduxalimov@gmail.ru of the OshTU named after. M.M. Adysheva, Osh city, Kyrgyz Republic

Abstract: *A distinctive feature of physical problems from other subject problems is that the connection between the unknown and the known is realized not only on the basis of physical laws, but also with the help of appropriate experiments.The article describes the method of Ordinary Least Squares, are examples of the experimental data most frequently encountered in the process of physical workshop. Several methods of processing of results of measurements are so far developed. The most common, simple and reasonable the method of ordinary least squares (OLS) is.*

Keywords*: the method of ordinary least squares (OLS), finding an approaching function, functional dependence, parameters, experiment, error.*

Введение

Сегодня исследователям очевидно, что существует потребность в передаче таких функций, как управление проведением эксперимента, обработка и анализ результатов измерений. Однако из-за недостаточного знания возможностей автоматизации активнее использовать ее сложно даже в тех сферах, где автоматизация эффективна или без нее, экспериментировать просто невозможно.

Метод наименьших квадратов - это один из методов регрессионного анализа, используемых для нахождения оценок параметров регрессии на основе минимизации суммы квадратов всех остатков. Основная идея этого метода состоит в том, что сумма квадратов ошибок рассматривается как критерий точности решения задачи, которую стремятся минимизировать. При использовании этого метода могут применяться как численный, так и аналитический подходы.

Нахождение приближающей функций - это приблизительное или точное определение любой величины из известных индивидуальных значений или других величин, связанных с ней. В исходном смысле это восстановление функции (точной или приближенной) до ее известных значений или значений ее производных в определенных областях. Основное применение этого метода - вычисление значения табличной функции для значений аргументов, отличных от узловых (косвенных), поэтому интерполяцию часто называют «искусством чтения таблиц между строками».

При выполнения лабораторных работ по курсу физика, для любых условий измерение какой-либо величины происходит по крайней мере 3 раза и полученный результат надо дать в виде $\sigma = \sigma \pm \Delta\sigma$. Это применяется для уменьшения прямых и косвенных ошибок прибора. На основе полученных результатов определяем зависимость искомой величины от заданных параметров. Метод наименьших квадратов требует с использования многочисленных числовых данных и вычислений. Поэтому важным для этого случая является компьютерная программа, которая отвечает всем условиям.

Цели задачи: создание программы для нахождения аппроксимирующей функции полученных экспериментальных результатов методом наименьших квадратов.

Для достижения этой цели необходимо решить следующие задачи:

- разработать методики обучения методом наименьших квадратов;
- рассмотреть метод наименьших квадратов, попарно линейную регрессию;
- разработать комплекс программ для получения данных и аппроксимация функции методом наименьших квадратов;
- разработать структуры данных для хранения и обработки экспериментальных данных;

– доказать, что найденная функция принимает минимальное значение, если коэффициенты являются решениями системы.

Допустим функциональная зависимость величины u от другой функциональной величины z известно, но параметры a, b, c,... этой зависимости неизвестны. Например, получена таблица(1) значений u_i, при некоторых значениях $z_i (i=1,2,...,N)$. Требуется найти такие значения параметров a, b, c,... при которых функция $u = f(z,a,b,c,...)$ наилучшим образом описывает экспериментальные данные.

Результаты проведенных измерений Таблица1.

N	A	B
1	1	1.1
2	3	1.98
3	5	3.2

МНК утверждает, что «наилучшей» кривой будет такая, для которой сумма квадратов отклонений экспериментальных значений u_i от значений функции $f(z,a,b,c,...)$ минимальна. Таким образом, для определения параметров a, b, c,... необходимо найти минимум функции:

$$\Phi = \sum_{i=1}^{N} (u_i - f(z,a,b,c,...))^2 \qquad (1)$$

Отметим, что Ф рассматривается здесь как функция параметров a, b, c,..., так как величины u_i, z_i известны из экспериментальных данных.

В общем случае нахождение минимума функции (1) удается сделать далеко не всегда. Поэтому для практической реализации МНК часто применяют следующий искусственный прием: находят некоторое функциональное преобразование $y = y(u), x = x(z)$, которое приводит исследуемую зависимость $u = f(z)$ к линейному виду[3,с.60]:

$$y = a \cdot x + b \qquad (2)$$

для реализации МНК наиболее проста. Если построить график этой зависимости, то будет видно, что точки лежат **примерно** на одной прямой и показано на рис.1.

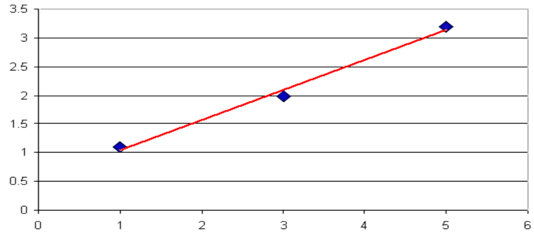

Рис.1. График зависимости

Некоторые преобразования будут рассмотрены ниже при изложении конкретных примеров. Примеры преобразований такого типа приведены в табл. 2.

Таблица 2.

Примеры преобразований

Вид зависимости	Преобразование	Параметры
$u = A \cdot z^{\gamma}$	$y = \ln u$ $x = \ln z$	$\ln u = \ln A + \gamma \cdot \ln z$ $a = \ln A \quad b = y$
$u = A \cdot e^{\gamma}\ u = A \cdot e^{\gamma}$	$y = \ln u$ $x = z$	$\ln u = \ln A + \gamma \cdot z$ $a = \ln A \quad b = y$

$u = \dfrac{A \cdot z}{1 + B \cdot z}$	$y = \dfrac{1}{u}$ $x = \dfrac{1}{z}$	$\dfrac{1}{u} = \dfrac{1}{A \cdot z} + \dfrac{B}{A}$ $a = \dfrac{1}{A},\ \ b = \dfrac{B}{A}$

Подставим выражение (2) в выражение (1) $\Phi = \sum_{i=1}^{N} (y_i - a \cdot x_i - b)^2$ (3) и получим уравнения для определения параметров a и b. Для этого вычислим производные функции Ф по a и b и приравняем их к нулю.

$$\begin{aligned} &\frac{d\Phi}{da} = 0; a \cdot \sum_{i=1}^{N} x_i^2 + b \cdot \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \cdot x_i \\ &\frac{d\Phi}{db} = 0; a \cdot \sum_{i=1}^{N} x_i + b \cdot N = \sum_{i=1}^{N} y_i \end{aligned} \qquad (4)$$

Данная система является линейной и легко решается:

$$\begin{aligned} a &= \frac{N\sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} y_i \cdot \sum_{i=1}^{N} x_i}{N\sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i^2)^2}; \\ b &= \frac{\sum_{i=1}^{N} y_i \cdot \sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{N} x_i \cdot \sum y_i \cdot x_i}{N\sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i^2)^2}. \end{aligned} \qquad (5)$$

Однако полученные выражения не очень удобны для практических расчетов, поэтому перепишем их в несколько иной форме.

$$S_x^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\frac{\sum_{i=1}^{N} x_i}{N} \right)^2 = (x^2) - (x)^2 \qquad (6)$$

Для этого обозначим (угловые скобки $\langle ... \rangle$ означают среднее арифметическое по экспериментальным данным) и запишем:

$$a = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{S_x^2} \qquad (7)$$

Из второго уравнения системы (4) выразим $b = \langle y \rangle - a \cdot \langle x \rangle$. Выражения (6), (7) позволяют достаточно быстро с помощью непрограммируемого калькулятора рассчитать параметры линейной зависимости (2). Сформулируем условия, при которых полученные таким способом значения параметров являются оптимальными (несмещенными, состоятельными, эффективными оценками [3,с.62].

1. Результаты измерений являются независимыми.
2. Погрешности измерений подчиняются нормальному распределению.
3. Величины x_i известны точно.

Практически МНК в изложенной форме применяют, если погрешности измерений y_i значительно (более чем на порядок) превосходят погрешности измерений величин x_i. При выполнении этих условий параметры a, b линейно выражаются через результаты измерений y_i, (погрешностями измерений x_i пренебрегаем), поэтому погрешность определения параметров может быть найдена стандартным методом как погрешность косвенного измерения. Несколько громоздкие выкладки приводят к следующим формулам для оценок погрешностей:

$$\Delta a = 2\sqrt{\frac{1}{N-2}\left(\frac{S_y^2}{S_x^2} - a^2\right)} \qquad (8)$$

где $S_y^2 = \langle y^2 \rangle - \langle y \rangle^2$, остальные обозначения сохраняем прежними:

$$\Delta b = \sqrt{S_x^2 + \langle x \rangle^2 \Delta a} \qquad (9)$$

Таким образом, формулы (6) – (9) полностью исчерпывают МНК для анализа линейной зависимости. Формулы (7) – (8) дают оценки только случайных погрешностей измерений. Их использование полностью оправдано, если этот тип погрешностей преобладает, что чаще всего бывает на практике. Свидетельством такого преобладания является заметный разброс точек (y_i, x_i) на графике, когда эти точки не ложатся точно на прямую. Отметим, что постоянная систематическая приборная погрешность не влияет на определение параметра a и является аддитивной добавкой к погрешности параметра b, т.е. если приборная погрешность измерения величин y_i равна Δy_{pr}, то $\Delta a_{pr} = 0; \Delta b_{pr} = \Delta y_{pr}$.

Отметим также, что в некоторых случаях необходимо проводить несколько измерений величины u при одном и том же значении z. В этом случае никаких модификаций МНК не требуется. Достаточно рассматривать эти значения как независимые, т.е. включать в расчеты пары z_i, u_i с одними и теми же значениями z_i. Иными словами, одному значению z может соответствовать несколько значений u. Естественно, не могут быть все z одинаковыми, иначе в формуле (5) в знаменателе окажется нуль.

Заключение

Вычсления по МНК обычно проводят на ЭВМ, используя стандартные программы. В результате вычислений следует написать уравнение прямой и провести ее на графике. Это и будет расчетная линия, имеющая наименьшее расхождение с результатами эксперимента. Во многих практически важных случаях (и в частности, при оценке сложных нелинейных связей) количество неизвестных параметров может быть достаточно много и поэтому реализация метода наименьших квадратов оказывается эффективной лишь при использовании современной вычислительной техники. Информация, представленная в настоящем статье, может стать основой для дальнейшей проработки и исследования метода в других областях науки.

Литература:

1. Вержбицкий В.М. Численные методы. Линейная алгебра и нелинейные уравнения./ В.М. Вержбицкий - М.: Оникс 21 век, 2005.
2. Лапчик М.П. Численные методы./ М.П. Лапчик, М.И. Рагулина, Е.К. Хеннер - М.: Академия, 2005
3. Самарский А.А. Введение в численные методы: учеб. пособие / А.А. Самарский - М.: Лань, 2008.С-62
4. Фадеев Д.К., Фадеева В.Н. Вычислительные методы линейной алгебры./ Д.К. Фадеев, В.Н. Фадеева- М.: Лань, 2002