Московский экономический журнал. № № 9. 2025
Moscow economic journal. № № 9. 2025
Научная статья
Original article
УДК 330.43
doi: 10.55186/2413046X_2025_10_9_209

КЛЮЧЕВЫЕ МАТЕМАТИЧЕСКИЕ СВОЙСТВА ОБОБЩЕННОГО МЕТОДА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

KEY MATHEMATICAL PROPERTIES OF THE GENERALIZED MAXIMUM LIKELIHOOD METHOD

Краснослободцева Татьяна Петровна, доцент, к.ф-м.н., доцент кафедры высшей и прикладной математики, Институт тонких химических технологий им. М.В. Ломоносова, МИРЭА - Российский технологический университет, Москва

Михайлова Наталия Александровна, старший преподаватель кафедры высшей и прикладной математики, Институт тонких химических технологий им. М.В. Ломоносова, МИРЭА - Российский технологический университет, Москва

Тимченко Татьяна Владимировна, старший преподаватель кафедры высшей и прикладной математики, Институт тонких химических технологий им. М.В. Ломоносова, МИРЭА - Российский технологический университет, Москва

Krasnoslobodceva Tatyana Petrovna, Associate Professor, PhD in Physics and Mathematics, Associate Professor of the Department of Higher and Applied Mathematics, Lomonosov Institute of Fine Chemical Technologies, MIREA - Russian Technological University, Moscow

Mixajlova Nataliya Aleksandrovna, Senior Lecturer at the Department of Higher and Applied Mathematics, Lomonosov Institute of Fine Chemical Technologies, MIREA - Russian Technological University, Moscow

Timchenko Tatyana Vladimirovna, Senior Lecturer at the Department of Higher and Applied Mathematics, Lomonosov Institute of Fine Chemical Technologies, MIREA - Russian Technological University, Moscow

Аннотация. В современных экономических исследованиях и управленческой практике, а также в различных отраслях промышленности и сферах деятельности, широко используются разнообразные статистические методы. Прикладная статистика представляет собой научную дисциплину, занимающуюся разработкой методов обработки эмпирических данных. В рамках этой дисциплины традиционно выделяют три основных направления: дескриптивный анализ данных, теорию оценивания параметров и проверку статистических гипотез. Одним из фундаментальных подходов в теории оценивания является метод максимального правдоподобия, основанный на оптимизации соответствующей функции правдоподобия. В данной работе исследуется этот метод в наиболее общей формулировке и предлагается его расширение на случай, когда стандартная задача оптимизации функции правдоподобия не имеет решения. Впервые получено необходимое и достаточное условие состоятельности оценки максимального правдоподобия в общей постановке. Для достижения этого результата потребовалось применение аппарата математической статистики в пространствах произвольной природы, что относится к центральным разделам статистики нечисловых данных. Далее исследуется ситуация, когда задача максимизации функции правдоподобия не имеет решения. В этом случае для оценивания функции распределения из соответствующего множества предлагается использовать введенную в работе обобщенную оценку максимального правдоподобия. Данный подход имеет определенное концептуальное сходство с методом регуляризации А. Н. Тихонова, разработанным для решения некорректно поставленных операторных уравнений. В работе приведены примеры вычисления обобщенных оценок максимального правдоподобия. Показано, что к таким оценкам относятся, в частности, эмпирическая функция распределения и её симметризованный вариант, полученный в предположении симметрии оцениваемого распределения относительно нуля. Симметризованная функция распределения находит применение, в частности, при проверке гипотез об однородности связанных выборок. Сформулирован ряд нерешенных проблем, связанных с развитием обобщенного метода максимального правдоподобия. Решению этих проблем предполагается посвятить дальнейшие научные изыскания.

Abstract. A variety of statistical methods are widely used in modern economic research and management practice, as well as in various industries and fields of activity. Applied statistics is a scientific discipline that develops methods for processing empirical data. Within the framework of this discipline, there are traditionally three main areas: descriptive data analysis, parameter estimation theory, and statistical hypothesis testing. One of the fundamental approaches in estimation theory is the maximum likelihood method, based on optimizing the corresponding likelihood function. In this paper, this method is investigated in its most general formulation and its extension is proposed for the case when the standard likelihood function optimization problem has no solution. For the first time, a necessary and sufficient condition for the consistency of the maximum likelihood estimate in the general formulation has been obtained. To achieve this result, it was necessary to apply the apparatus of mathematical statistics in spaces of arbitrary nature, which refers to the central sections of statistics of non-numerical data. Next, we study the situation when the problem of maximizing the likelihood function has no solution. In this case, it is proposed to use the generalized maximum likelihood estimate introduced in the paper to estimate the distribution function from the corresponding set. This approach has certain conceptual similarities with A. N. Tikhonov's regularization method, developed to

solve incorrectly posed operator equations. The paper provides examples of calculating generalized maximum likelihood estimates. It is shown that such estimates include, in particular, the empirical distribution function and its symmetrized version obtained under the assumption of symmetry of the estimated distribution relative to zero. The symmetrized distribution function finds application, in particular, in testing hypotheses about the homogeneity of related samples. A number of unsolved problems related to the development of the generalized maximum likelihood method are formulated. It is planned to devote further scientific research to solving these problems.
Ключевые слова: статистические методы экономики, математическая статистика, оценивание, метод максимального правдоподобия, статистика нечисловых данных, предельные теоремы

Keywords: statistical methods of economics, mathematical statistics, estimation, maximum likelihood method, statistics of non-numerical data, limit theorems

Введение

В сфере экономики, менеджмента и различных отраслей промышленности активно используются разнообразные статистические подходы. Статистика как прикладная дисциплина занимается разработкой методов систематизации и анализа эмпирической информации. В рамках данной науки, исходя из типа решаемых задач, традиционно выделяют три ключевых направления: дескриптивный анализ данных, оценивание параметров распределений и проверку статистических гипотез.

Среди множества методов параметрического оценивания особое место занимает подход максимального правдоподобия, фундаментом которого является процедура оптимизации специальной функции правдоподобия. В рамках данного исследования проводится анализ этого метода в его наиболее общей формулировке, а также рассматривается его расширение для ситуаций, когда стандартная задача оптимизации целевой функции не может быть решена традиционными способами.

1. Основные понятия

В рамках общепринятой вероятностной модели данных рассмотрим выборку $X_1, X_2, ..., X_n$, представляющую собой совокупность n независимых одинаково распределенных случайных величин. С математической точки зрения, случайная величина определяется как функция, заданная на пространстве элементарных событий и принимающая значения в некотором измеримом пространстве A. Хотя традиционно в качестве A рассматривается множество действительных чисел или конечномерное векторное пространство, современная теория вероятностей не накладывает таких ограничений. Пространство A может быть произвольным, в том числе нечисловой природы. Это позволяет естественным образом вводить такие понятия, как случайное бинарное отношение или случайное множество, рассматривая их как случайные величины со значениями в соответствующих пространствах (бинарных отношений или подмножеств данного множества).

С математической точки зрения важно отметить, что общее определение случайной величины предполагает задание σ-алгебры измеримых подмножеств пространства A. Однако если A является конечным множеством, естественно считать все его подмножества измеримыми, что позволяет избежать сложностей, связанных с вопросами измеримости.

Для введения понятия "плотность распределения вероятностей" в пространстве A необходимо задать некоторую эталонную меру μ. Функция f: A → [0, ∞) называется плотностью распределения случайной величины X относительно меры μ тогда и только тогда, когда для любого измеримого подмножества B ⊆ A выполняется равенство:

$$P(X \in B) = \int_B f(x)\mu(dx).$$

В случае, когда пространство A представляет собой множество действительных чисел, а мера μ соответствует стандартной мере Лебега, для которой мера отрезка единичной длины равна 1, правая часть формулы (1) представляет собой классический интеграл Римана или Лебега. Аналогичным образом, если пространство A является конечномерным евклидовым пространством, то выражение в правой части формулы (1) соответствует общеизвестному в математическом анализе многомерному интегралу.

В параметрической статистике предполагается, что плотность распределения f(x) зависит от неизвестного параметра θ, который подлежит оценке на основе имеющейся выборки. Для решения этой задачи разработан ряд статистических методов. Один из наиболее фундаментальных подходов основан на анализе функции правдоподобия, которая определяется как совместная плотность распределения наблюдаемых данных, рассматриваемая как функция параметра θ при фиксированных значениях выборки.

Функция правдоподобия $L(\theta; x_1, ..., x_n)$ формально записывается как произведение плотностей отдельных наблюдений (в случае независимых одинаково распределенных данных):

$$g(X_1, X_2, ..., X_n; \theta) = f(X_1, \theta) f(X_2, \theta) ... f(X_n, \theta),$$

Таким образом, функция правдоподобия представляет собой плотность совместного распределения n элементов выборки в наблюдаемой точке n-мерного пространства $(X_1, X_2, ..., X_n)$.

Если выборка $X_1, X_2, ..., X_n$ получена из распределения с истинным значением параметра $\theta_0 \in \Theta$, то естественно предположить, что значение функции правдоподобия (2) достигает своего максимума по $\theta \in \Theta$ (при фиксированных $X_1, X_2, ..., X_n$) именно в точке $\theta_0 \in \Theta$. Это интуитивное соображение формирует теоретическое обоснование для оценивания неизвестного параметра $\theta_0 \in \Theta$ путем максимизации функции правдоподобия (2).

$$g(X_1, X_2, ..., X_n; \theta) \to \max_{\theta}, \ \theta \in \Theta.$$

2. Оценка максимального правдоподобия

При исследовании оценок максимального правдоподобия возникает ряд фундаментальных вопросов, требующих тщательного анализа:

- существование решения оптимизационной задачи;
- единственность полученного решения;
- состоятельность оценки максимального правдоподобия;
- сравнительная эффективность по отношению к другим оценкам параметра.

Если пространство параметров Θ является бикомпактным (т.е. из любого его открытого покрытия можно выбрать конечное подпокрытие), то существование оценки максимального правдоподобия гарантировано. Понятие бикомпактности обобщает свойства замкнутых ограниченных множеств в конечномерных евклидовых пространствах на топологические пространства произвольной природы.

Как показывают контрпримеры, решение задачи оптимизации может быть неединственным. При выполнении определенных условий регулярности [1] оценка максимального правдоподобия является состоятельной.

В асимптотическом смысле оценка максимального правдоподобия, как правило, является асимптотически эффективной среди всех оценок одномерного параметра и принадлежит к классу наилучших асимптотически нормальных оценок. Однако при конечных объемах выборок в некоторых ситуациях могут существовать оценки с меньшей дисперсией, например, несмещенные оценки.

В случаях, когда элементы выборки имеют нормальное распределение с неизвестными параметрами (математическим ожиданием и дисперсией), оценки максимального правдоподобия выражаются в явном виде через выборочное среднее и выборочную дисперсию. Однако для многих других распределений (гамма-распределение, бета-распределение) задача не имеет аналитического решения, что требует применения численных методов оптимизации.

В работах [2, 3] предлагается использовать одношаговые оценки (первую итерацию метода Ньютона-Рафсона), которые обладают теми же асимптотическими свойствами, что и оценки максимального правдоподобия, но задаются явными формулами, что устраняет вычислительные трудности.

Поскольку максимум функции положительных значений достигается в той же точке, что и максимум ее логарифма, для получения оценок максимального правдоподобия необходимо решать оптимизационную задачу для логарифмической функции правдоподобия.

Для анализа асимптотического поведения оценок максимального правдоподобия могут быть применены методы статистики нечисловых данных. Множество решений задачи оптимизации обозначается как Arg max и может содержать более одной точки.

В соответствии с законом больших чисел, выборочное среднее арифметическое сходится к математическому ожиданию при росте объема выборки. Соотношение (8), установленное в статистике нечисловых данных, описывает асимптотическое поведение оценок максимального правдоподобия. Из него следует, что для состоятельности оценки необходимо и достаточно, чтобы теоретическое среднее (7) состояло из единственной точки, совпадающей с истинным значением параметра $\theta_0 \in \Theta$.

Случай, когда пространство параметров Θ является подмножеством конечномерного евклидова пространства, детально исследован в классической математической статистике [1].

3. Основная идея метода максимального правдоподобия

При выполнении соответствующих условий регулярности, подробно описанных в авторитетных монографиях и учебных пособиях по математической статистике, решение оптимизационной задачи (4), известное как оценка максимального правдоподобия, обладает тремя фундаментальными свойствами: существованием, состоятельностью и асимптотической эффективностью.

$$Q(\Theta)=\left\{\int_{-\infty}^{x} f(x,\theta)dx,\ \ \theta\in\Theta\right\}.$$

Одним из ключевых условий регулярности в классической теории оценивания является существование взаимно однозначного соответствия между параметрическим пространством Θ и соответствующим множеством функций распределения Q(Θ). Это позволяет отождествить Θ с Q(Θ), рассматривая параметрическое пространство как подмножество множества всех функций распределения Q_0.

Однако в ряде практически важных случаев классическая оценка максимального правдоподобия не существует, поскольку условия регулярности не выполняются — в частности, когда некоторые допустимые распределения не обладают плотностями. Такая ситуация возникает, например, когда параметрическое пространство совпадает со всем множеством функций распределения: $\Theta = Q_0$. В этих случаях может быть применен метод обобщенного максимального правдоподобия.

Пусть X_1, X_2, ..., X_n — независимые одинаково распределенные случайные величины с неизвестной функцией распределения F, принадлежащей известному семейству $Q \subseteq Q_0$. Для оценивания F предлагается следующий подход. Пусть ρ — метрика Леви в Q_0. Рассмотрим последовательность подмножеств Q_1, Q_2, ..., Q_n, ... множества Q_0, замкнутых в топологии, порожденной метрикой Леви, таких что $Q_m \subset Q_p$ при $m < p$, а объединение всех Q_i всюду плотно в Q_0. Предположим, что оценка максимального правдоподобия существует в обычном смысле при $\Theta = Q \cap Q_n$. Обозначим эту оценку F_{nn}. Оценкой обобщенного максимального правдоподобия называется предел F_{nn} при $N \to \infty$, если данный предел существует.

Естественным образом возникают три фундаментальных вопроса:

а) При каких условиях существуют оценки F_{nn}?

б) Когда существует оценка обобщенного максимального правдоподобия?

в) Является ли оценка обобщенного максимального правдоподобия состоятельной?

Ответ на первый вопрос относительно прост. Если плотность $f(x) = F'(x)$ является непрерывной функцией от $F \in Q_n$ для любого x и любого N, причем Q_n — компактные множества, а Q замкнуто, тогда $Q \cap Q_n$ также компактны, и максимум непрерывной функции достигается, то есть оценка F_{nn} существует.

Что касается вопросов (б) и (в), то в общем случае ответы на них остаются неизвестными. Выдвигается гипотеза, что при определенных условиях регулярности оценка обобщенного максимального правдоподобия существует и является состоятельной.

Перспективными направлениями дальнейших исследований являются:

1. Доказательство или опровержение сформулированной гипотезы
2. Поиск конкретных условий существования и состоятельности оценок
3. Исследование частных случаев, таких как:

- Эмпирическая функция распределения
- Симметризованная функция распределения
- Оценки для параметрических семейств с сингулярными распределениями

Разработка общей теории обобщенного максимального правдоподобия представляет значительный интерес для современной математической статистики и ее приложений.

Заключение

В рамках проведенного исследования было получено необходимое и достаточное условие состоятельности оценки максимального правдоподобия в общей постановке. Для достижения этого результата потребовалось привлечение математического аппарата статистики в пространствах произвольной природы, представляющего собой центральный раздел статистики нечисловых данных.

Далее был исследован случай, когда стандартная задача максимизации функции правдоподобия не имеет решения. Для подобных ситуаций предложено использовать разработанную обобщенную оценку максимального правдоподобия, позволяющую оценивать функцию распределения из соответствующего множества. Данный подход обладает определенным концептуальным сходством с методом регуляризации Тихонова, разработанным для решения некорректно поставленных операторных задач [9].

В работе приведены конкретные примеры обобщенных оценок максимального правдоподобия. Показано, что к ним относятся, в частности:

- Эмпирическая функция распределения
- Её симметризованный вариант, полученный в предположении симметрии оцениваемого распределения относительно нуля

Симметризованная функция распределения находит практическое применение, в частности, при проверке гипотез об однородности связанных выборок.

Сформулирован ряд нерешенных проблем, связанных с дальнейшим развитием обобщенного метода максимального правдоподобия. К числу наиболее актуальных направлений для будущих исследований относятся:

1. Доказательство теорем существования для обобщенных оценок
2. Исследование условий их состоятельности и асимптотической нормальности
3. Разработка вычислительных алгоритмов для практической реализации
4. Изучение свойств симметризованных оценок для различных типов распределений

Решение этих задач представляет значительный интерес как для теоретической статистики, так и для прикладных исследований.

В то время как в одних регионах наблюдается избыток рабочей силы, другие сталкиваются с ее дефицитом. Это связано как с географическими и демографическими факторами, так и с различиями в уровне экономического развития регионов. Кроме того, сохраняется проблема несоответствия структуры.

Список источников

1. Боровков А.А. Математическая статистика. Изд. 5-е, стереотипное. - Санкт- Петербург : Лань, 2021. - 704 с.
2. Орлов А.И. Оценивание параметров: одношаговые оценки предпочтительнее оценок максимального правдоподобия // Научный журнал КубГАУ. 2015. №109. С. 208 – 237.
3. Орлов А.И. Прикладной статистический анализ. — М.: Ай Пи Ар Медиа, 2022. — 812 с.
4. Орлов А.И. Искусственный интеллект: нечисловая статистика : учебник. — М.: Ай Пи Ар Медиа, 2022. — 446 с.
5. Тюрин Ю.Н. Об оценивании функции распределения // Теория вероятностей и ее применения. 1970. Т. 15. № 3. С. 567-568.
6. Орлов А.И. О проверке однородности связанных выборок // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2016. № 123. С. 708–726.
7. Астафьев, Р. У. Методика формирования базы знаний для системы управления качеством программного обеспечения / Р. У. Астафьев // Научно-технологическое развитие 2025: сборник статей Международной научно-практической конференции, Петрозаводск, 26 июня 2025 года. – Петрозаводск: Международный центр научного партнерства «Новая Наука» (ИП Ивановская И.И.), 2025. – С. 126-130. – EDN WHXWFG.
8. Астафьев, Р. У. Роль имитационных моделей в системах поддержки принятия решений в области разработки программных продуктов / Р. У. Астафьев // Оптические технологии, материалы и системы (Оптотех - 2024): Международная научно-техническая конференция, Москва, 02–08 декабря 2024 года. – Москва: МИРЭА - Российский технологический

университет, 2024. – С. 789-790. – EDN JTFOGS.

9. Сидоров, А. А. Доказательство свойств средних степенных / А. А. Сидоров // Инновационные технологии в электронике и приборостроении: сборник докладов Российской научно-технической конференции с международным участием Физико-технологического института РТУ МИРЭА, Москва, 16–17 апреля 2020 года. Том 1. – Москва: МИРЭА - Российский технологический университет, 2020. – С. 287-293. – EDN ELMJXA.

10. Об одном аспекте в вопросе определения аналитичностифункции комплексного переменного / О. Ю. Козлова, Т. А. Манаенкова, А. И. Новикова [и др.] // Перспективные материалы и технологии (ПМТ-2024) : Сборник докладов Международной научно-технической конференции, Москва, 12–16 апреля 2024 года. – Москва: МИРЭА - Российский технологический университет, 2024. – С. 422-425. – EDN EMGWJP.

11. SIDOROV Andrei, 2024, THE IMPACT OF ANNOUNCEMENTS ON CRYPTOCURRENCY PRICES, Revista Economică, Lucian Blaga University of Sibiu, Faculty of Economic Sciences, vol.76(4), pages 69-94, December. DOI: https://doi.org/10.56043/reveco-2024-0035

References

1. Borovkov A.A. Matematicheskaya statistika. Izd. 5-e, stereotipnoe. - Sankt-Peterburg : Lan`, 2021. - 704 s.

2. Orlov A.I. Ocenivanie parametrov: odnoshagovy`e ocenki predpochtitel`nee ocenok maksimal`nogo pravdopodobiya // Nauchny`j zhurnal KubGAU. 2015. №109. S. 208 – 237.

3. Orlov A.I. Prikladnoj statisticheskij analiz. — M.: Aj Pi Ar Media, 2022. — 812 c.

4. Orlov A.I. Iskusstvenny`j intellekt: nechislovaya statistika : uchebnik. — M.: Aj Pi Ar Media, 2022. — 446 c.

5. Tyurin Yu.N. Ob ocenivanii funkcii raspredeleniya // Teoriya veroyatnostej i ee primeneniya. 1970. T. 15. № 3. S. 567-568.
6. Orlov A.I. O proverke odnorodnosti svyazanny`x vy`borok // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2016. № 123. S. 708–726.
7. Astaf`ev, R. U. Metodika formirovaniya bazy` znanij dlya sistemy` upravleniya kachestvom programmnogo obespecheniya / R. U. Astaf`ev // Nauchno-texnologicheskoe razvitie 2025: sbornik statej Mezhdunarodnoj nauchno-prakticheskoj konferencii, Petrozavodsk, 26 iyunya 2025 goda. – Petrozavodsk: Mezhdunarodny`j centr nauchnogo partnerstva «Novaya Nauka» (IP Ivanovskaya I.I.), 2025. – S. 126-130. – EDN WHXWFG.
8. Astaf`ev, R. U. Rol` imitacionny`x modelej v sistemax podderzhki prinyatiya reshenij v oblasti razrabotki programmny`x produktov / R. U. Astaf`ev // Opticheskie texnologii, materialy` i sistemy` (Optotex - 2024): Mezhdunarodnaya nauchno-texnicheskaya konferenciya, Moskva, 02–08 dekabrya 2024 goda. – Moskva: MIRE`A - Rossijskij texnologicheskij universitet, 2024. – S. 789-790. – EDN JTFOGS.
9. Sidorov, A. A. Dokazatel`stvo svojstv srednix stepenny`x / A. A. Sidorov // Innovacionny`e texnologii v e`lektronike i priborostroenii: sbornik dokladov Rossijskoj nauchno-texnicheskoj konferencii s mezhdunarodny`m uchastiem Fiziko-texnologicheskogo instituta RTU MIRE`A, Moskva, 16–17 aprelya 2020 goda. Tom 1. – Moskva: MIRE`A - Rossijskij texnologicheskij universitet, 2020. – S. 287-293. – EDN ELMJXA.
10. Ob odnom aspekte v voprose opredeleniya analitichnostifunkcii kompleksnogo peremennogo / O. Yu. Kozlova, T. A. Manaenkova, A. I. Novikova [i dr.] // Perspektivny`e materialy` i texnologii (PMT-2024) : Sbornik dokladov Mezhdunarodnoj nauchno-texnicheskoj konferencii, Moskva, 12–16 aprelya 2024 goda. – Moskva: MIRE`A - Rossijskij texnologicheskij universitet, 2024. – S. 422-425. – EDN EMGWJP.

11. SIDOROV Andrei, 2024, THE IMPACT OF ANNOUNCEMENTS ONCRYPTOCURRENCY PRICES, Revista Economică, Lucian Blaga University of Sibiu, Faculty of Economic Sciences, vol.76(4), pages 69-94, December. DOI: https://doi.org/10.56043/reveco-2024-0035