

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ MODELING AND SIMULATION

doi: 10.17586/2226-1494-2024-24-2-267-275

УДК 519.233.22

О свойствах M -оценок, оптимизирующих весовую L_2 -норму функции влияния

Даниил Валерьевич Лисицин¹, Константин Викторович Гаврилов²✉

^{1,2} Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация

¹ lisitsin@ami.nstu.ru, <https://orcid.org/0009-0007-0723-6285>

² aenigma77@mail.ru✉, <https://orcid.org/0009-0001-9252-9942>

Аннотация

Введение. В работе развивается теория устойчивых M -оценок, относящихся к классу сниженных оценок, обладающих свойством устойчивости к асимметричному засорению. Многие известные сниженные оценки могут быть получены в рамках двух подходов д.т.н. А.М. Шурыгина: локально устойчивого подхода, основанного на анализе показателя неустойчивости оценки (L_2 -нормы функции влияния), или подхода, основанного на модели серии выборок со случайным точечным засорением (модели байесовского точечного засорения). Эти подходы удобны для построения различных устойчивых M -оценок и, по сравнению с классическими робастными процедурами, предоставляют более широкие возможности. Предложенное А.М. Шурыгиным в рамках первого из перечисленных подходов семейство условно оптимальных оценок может определяться как оптимизирующее асимптотическую дисперсию при ограничении на величину неустойчивости. Соответствующая задача допускает представление в форме оптимизации весовой L_2 -нормы функции влияния. Во втором подходе рассматривается специальным образом сформированная непараметрическая окрестность модельного распределения, и он тоже может быть сведен к анализу весовой L_2 -нормы функции влияния. Таким образом, данный критерий качества оценивания является достаточно общим и полезным для конструирования робастных оценок. **Метод.** Теория оценок, оптимальных с точки зрения весовой L_2 -нормы функции влияния, в настоящее время недостаточно развита. Так, для соответствующих семейств оценок остается нерешенным вопрос единственности членов семейства. Вопрос сводится к исследованию выпуклости (вогнутости) оптимизируемого функционала в зависимости от параметра, задающего семейство. **Основные результаты.** В работе в общем виде получено выражение для производной по параметру функционала качества оптимальной оценки. Получены неравенства для второй производной, необходимые для установления его выпуклости (вогнутости) по параметру. Полученные результаты применены для описания свойств условно оптимального семейства. Построены функции влияния ряда условно оптимальных оценок для параметров сдвига и масштаба нормальной модели. Исследованы характеристики этих оценок. **Обсуждение.** Показана устойчивость большинства рассмотренных оценок, что важно для их практического применения. Теоретические результаты могут быть полезны при исследовании свойств компромиссных оценок на базе двух критериев, а также при изучении минимаксных уровней засорения в рамках подхода А.М. Шурыгина на основе модели байесовского точечного засорения. Результаты работы могут найти применение в ситуациях целенаправленного искажения данных противником, в том числе в задачах, связанных с вредоносным машинным обучением.

Ключевые слова

M -оценки, робастная статистика, функция влияния, устойчивые оценки, сниженные оценки, условно оптимальные оценки

Ссылка для цитирования: Лисицин Д.В., Гаврилов К.В. О свойствах M -оценок, оптимизирующих весовую L_2 -норму функции влияния // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 2. С. 267–275. doi: 10.17586/2226-1494-2024-24-2-267-275

On the properties of M -estimators optimizing weighted L_2 -norm of the influence function

Daniil V. Lisitsin¹, Konstantin V. Gavrilov²✉

^{1,2} Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation

¹ lisitsin@ami.nstu.ru, <https://orcid.org/0009-0007-0723-6285>

² aenigma77@mail.ru✉, <https://orcid.org/0009-0001-9252-9942>

Abstract

The work develops the theory of stable M -estimators belonging to the class of redescending estimators, having the property of resistance to asymmetric contamination. Many well-known redescending estimators can be obtained within the framework of the locally stable approach of A.M. Shurygin, based on the analysis of the estimator instability functional (L_2 -norm of the influence function), or his approach based on the model of a series of samples with random point contamination (point Bayesian contamination model). These approaches are convenient for constructing various stable M -estimators and, in comparison with classical robust procedures, provide wider opportunities. The family of conditionally optimal estimators proposed by A.M. Shurygin within the framework of the first of the listed approaches can be defined as optimizing the asymptotic dispersion under a constraint on the value of instability. The corresponding problem can be represented in the form of optimization of the weighted L_2 -norm of the influence function. The second approach considers a specially formed nonparametric neighborhood of the model distribution, and it can also be reduced to the analysis of the weighted L_2 -norm of the influence function. Thus, this estimation quality criterion is quite general and useful for constructing robust estimators. The theory of estimators that are optimal in terms of weighted L_2 -norm of the influence function is currently underdeveloped. Specifically, for the corresponding families of estimators, the question of the uniqueness of family members remains unresolved. The question comes down to studying the convexity (concavity) of the optimized functional depending on the parameter defining the family. In the presented work, an expression is obtained in general form for the derivative with respect to the parameter of the quality functional of the optimal estimator. Inequalities are obtained for the second derivative necessary to establish its convexity (concavity) with respect to the parameter. Corollaries from these results are applied to describe the properties of a conditionally optimal family. The influence functions of a number of conditionally optimal estimators for the shift and scale parameters of the normal model are constructed. The characteristics of these estimators are studied. The stability of most of the considered estimators is shown, which is important for their practical application. The theoretical results obtained can be useful in studying the properties of compromise estimators based on two criteria as well as in studying minimax contamination levels within the framework of A.M. Shurygin's point Bayesian contamination model. The results of the work can be used in situations of purposed data corruption by an adversary including the problems related to adversarial machine learning.

Keywords

M -estimators, robust statistics, influence function, stable estimates, redescending estimators, conditionally optimal estimators

For citation: Lisitsin D.V., Gavrilov K.V. On the properties of M -estimators optimizing weighted L_2 -norm of the influence function. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 2, pp. 267–275 (in Russian). doi: 10.17586/2226-1494-2024-24-2-267-275

Введение

Современные подходы к решению задачи оценивания параметров статистических моделей [1] направлены на обеспечение устойчивости получаемых решений к возможным отклонениям реальной ситуации от принятых в модели предположений [2–4]. Одним из наиболее трудных для формализации являлось асимметричное засорение симметрично распределенных данных, поскольку априорная информация о возможном искажении, как правило, отсутствует, а классические робастные подходы в этой ситуации не обладают достаточной устойчивостью. Для решения этой проблемы были предложены сниженные оценки [3, 4], в которых влияние периферийных наблюдений значительно снижено по сравнению с классическими робастными решениями. Однако эти решения оставались, по существу, эвристическими.

Важным шагом к получению теоретически обоснованных решений стала предложенная д.т.н. Александром Михайловичем Шурыгиным [2, 5] достаточно универсальная модель байесовского точечного засорения (БТЗ), которая позволяет описывать воздействие асимметричного засорения на набор данных.

Модель предполагает наличие серии выборок, каждая из которых имеет засорение в виде распределения, сосредоточенного в одной точке. Эта точка является фиксированной в пределах одной выборки, но имеет некоторое распределение по серии выборок. Каждому распределению засоряющей точки соответствует наилучшая оценка, полученная в результате минимизации ее асимптотического квадратичного отклонения [2]. Получающиеся оценки часто оказываются сниженными. В рамках модели БТЗ оптимальными являются многие известные сниженные оценки параметра сдвига, среди них бивес-оценка Тьюки, оценки Эндрюса, Бернулли, Смита, Мешалкина, оценка Хьюбера типа урезанного среднего, обобщенные оценки Шарбонье [2, 6, 7].

Еще один — локально устойчивый — подход А.М. Шурыгина [2] основан на показателе неустойчивости оценки (L_2 -норме функции влияния). Оптимизация неустойчивости при ограничении на асимптотическую дисперсию (или, что тоже, оптимизация асимптотической дисперсии при ограничении на величину неустойчивости) приводит к семейству условно оптимальных оценок [2, 8], которые, как правило, тоже являются сниженными.

Характерно, что и условно оптимальное семейство, и решения, получаемые в подходе на основе модели БТЗ, могут быть сведены к оптимизации весовой L_2 -нормы функции влияния [4]. Таким образом, данный подход является достаточно общим и востребованным: он может служить теоретической основой для конструирования широкого спектра оптимальных устойчивых оценок. Однако к настоящему времени оставался неизученным ряд свойств оптимальных оценок, в частности, открытым оставался вопрос единственности членов их семейств, который связан с вопросом выпуклости (вогнутости) оптимизируемого функционала в зависимости от параметра.

В настоящей работе в общем виде исследуются свойства оценок, оптимизирующих весовую L_2 -норму функции влияния при условии, что весовая функция зависит от скалярного параметра. Основные результаты — получение выражения для производной по параметру функционала качества и неравенства для второй его производной, необходимые для установления его выпуклости (вогнутости). В частности, для условно оптимального семейства [2, 8, 9] показана монотонность функций асимптотической дисперсии и неустойчивости оценки, что обеспечивает единственность элементов семейства [8]. В качестве применения теории рассмотрен ряд условно оптимальных оценок параметров сдвига и масштаба нормальной модели, исследованы устойчивость и другие характеристики этих оценок.

Элементы теории устойчивого оценивания

Пусть $x_1, \dots, x_m$ — независимые наблюдения случайной величины ξ , распределенной с плотностью $f(x, \theta)$, где $x \in X \subseteq R$ и параметр $\theta \in \Theta \subseteq R$. Здесь X и Θ — множества значений переменной x и параметра θ соответственно, R — множество вещественных чисел. M -оценка $\hat{\theta}$ неизвестного параметра может определяться как решение оценочного уравнения [1, 3]

$$\sum_{i=1}^m \psi(x_i, \hat{\theta}) = 0,$$

где $\psi(x, \theta)$ — *оценочная функция* параметра θ . Оценка должна удовлетворять условию асимптотической несмещенности вида [1, 10]

$$\mathbf{E} \psi(\xi, \theta) = \int_X \psi(x, \theta) f(x, \theta) dx = 0, \quad (1)$$

где $\mathbf{E}$ — оператор математического ожидания.

Дифференцируя (1) по θ и допуская возможность изменения порядка дифференцирования и интегрирования, можно записать следующие равенства [3, 10]:

$$\begin{aligned} N(\theta) &= \lim_{t \rightarrow \theta} \frac{d}{dt} \mathbf{E} \psi(\xi, t) = -\mathbf{E} \frac{\partial}{\partial \theta} \psi(\xi, \theta) = \\ &= \int_X \psi(x, \theta) \frac{\partial}{\partial \theta} f(x, \theta) dx, \end{aligned} \quad (2)$$

где функция $N(\theta)$ имеет смысл нормировочной характеристики оценочной функции.

Потребуем, чтобы в окрестности истинного значения параметра θ выполнялось условие асимптотической

несмещенности (1), были справедливы равенства (2), функция $N(\theta)$ была непрерывной и не равной нулю [11]. Более полный набор условий регулярности, обеспечивающий также $\sqrt{m}$ -состоятельность и асимптотическую нормальность оценок [1, 12, 13], приведен, например, в работе [5].

Рассмотрим в качестве показателя качества оценивания квадрат весовой L_2 -нормы функции влияния с весом $s(x, \theta)$

$$U(\psi, s) = \frac{1}{N^2(\theta)_X} \int \psi^2(x, \theta) s(x, \theta) dx = \int IF^2(x, \theta) s(x, \theta) dx, \quad (3)$$

где $IF(x, \theta) = \psi(x, \theta)/N(\theta)$, IF — функция влияния Хампеля [4]. Функционал (3) возникает, например, как асимптотическое квадратичное отклонение оценки в модели БТЗ [2, 5, 14], где весовая функция $s(x, \theta)$ представляет собой плотность распределения засоряющей точки. В частности, асимптотическая дисперсия [1] M -оценки $\hat{\theta}$ определяется как $V(\psi) = U(\psi, f)$, а неустойчивость оценки [2, 14, 15] — $W(\psi) = U(\psi, 1)$. Минимизация функционала $V(\psi)$ приводит к *оценке максимального правдоподобия* (ОМП) [1], а минимизация $W(\psi)$ — к *оценке максимальной устойчивости* (ОМУ) [2]. Соответствующие значения оптимизируемых функционалов обозначим $V_{\text{ОМП}}$ и $W_{\text{ОМУ}}$. Также используем относительные характеристики оценки — *эффективность* [1, 2] $\text{eff} \psi = V_{\text{ОМП}}/V(\psi)$ и *устойчивость* [2].

В [2] показано, что функционал (3) достигает минимума по ψ на функции

$$\psi(x, \theta) = c(\theta) \left[\frac{\partial}{\partial \theta} \ln f(x, \theta) + \beta(\theta) \right] \frac{f(x, \theta)}{s(x, \theta)}, \quad (4)$$

где $c(\theta)$ — произвольная непрерывная функция, не равная нулю для всех $\theta \in \Theta$; функция $\beta(\theta)$ определяется из условия (1). Оценочную функцию (4) назовем оптимальной для заданной весовой функции $s(x, \theta)$.

Функция (4) является также решением задачи максимизации функционала (3) для функции $s(x, \theta) \leq 0$. Таким образом, весовая функция $s(x, \theta)$ в (4) может быть неотрицательной или неположительной, что расширяет исходное определение функционала (3) как квадрата весовой L_2 -нормы. Соответственно функционалу (3) при этом доставляется минимум либо максимум.

Далее для краткости, как правило, будем опускать аргументы функций. Например, (4) запишем в виде $\psi = c(\partial f / \partial \theta + \beta f) / s$.

Введем следующее соглашение. Когда функционал (3) или его частный случай (V или W) используется без указания аргументов, подразумевается, что в качестве оценочной функции в него подставляется выражение (4), зависящее от s . Если при этом весовая функция s является членом параметрического семейства, то функционал представляет собой функцию параметра, задающего это семейство. Штрихом будем обозначать производную функций по данному параметру.

Так, семейству условно оптимальных оценок [2, 8, 9, 11, 16] в зависимости от способа параметризации соответствуют функции $s = 1 + \gamma f$ или $s = \lambda + f$, где γ, λ — параметры, задающие семейство. Данное семей-

ство оценок может определяться как минимизирующее неустойчивость при ограничении на асимптотическую дисперсию (или минимизирующее асимптотическую дисперсию при ограничении на неустойчивость). Причем в задаче минимизации возможны два вида ограничений: ограничения-неравенства и ограничения-равенства. Последние приводят к более широким диапазонам параметров, задающих семейство, и потому оно названо *расширенным* условно оптимальным семейством [8].

Для обеспечения корректности приводимых далее рассуждений дополнительно потребуем, чтобы все полученные производные и интегралы существовали и были непрерывными функциями параметра, задающего семейство; для выписанных интегралов было допустимо, где это необходимо, внесение операции дифференцирования под знак интеграла.

Свойства оптимальных оценочных функций

Изучим свойства оптимальных оценочных функций, определяемых выражением (4). В работе [11] показано, что для величины (3) справедливо:

$$U = \frac{c}{N} = c^2 \int_X \psi^2 s dx,$$

где функция N определена в выражении (2), а c — в (4).

Перепишем равенства в виде [5]:

$$\frac{1}{U} = \frac{N}{c} = \frac{1}{c^2} \int_X \psi^2 s dx = \int_X \varphi^2 s dx, \quad (5)$$

где $\varphi = \psi/c = (\partial f / \partial \theta + \beta f)/s$.

Пусть функция s представляет собой параметрическое семейство, зависящее от некоторого скалярного параметра, от которого не зависит f . Тогда получим:

$$\varphi' = (\beta' f - \varphi s')/s. \quad (6)$$

Теорема 1. Для оценочной функции (4) справедливо

$$U' = \frac{1}{N^2} \int_X \psi^2 s' dx = \int_X \text{IF}^2 s' dx.$$

Доказательство. Действительно, используя равенства (5), (6) и (1), найдем

$$\left(\frac{N}{c}\right)' = 2\beta' \int_X \varphi f dx - \int_X \varphi^2 s' dx = \int_X \varphi \varphi' s dx = - \int_X \varphi^2 s' dx; \quad (7)$$

$$\begin{aligned} U' &= - \left(\frac{c}{N}\right)^2 \left(\frac{N}{c}\right)' = \left(\frac{c}{N}\right)' \int_X \varphi^2 s' dx = \\ &= \frac{1}{N^2} \int_X \psi^2 s' dx = \int_X \text{IF}^2 s' dx. \end{aligned} \quad (8)$$

Теорема доказана.

Функционал U' в зависимости от вида s и способа ее параметризации может иметь различный смысл. Рассмотрим некоторые частные случаи теоремы 1.

Следствие 1 из теоремы 1. Пусть $s = \gamma f + h$, где γ — параметр, задающий семейство; h — независимая от

γ функция. Тогда для оптимальной оценочной функции (4) справедливо

$$U' = \frac{1}{N^2} \int_X \psi^2 f dx = U(\psi, f) = V.$$

Следствие 2 из теоремы 1. Пусть $s = f + \lambda h$, где λ — параметр, задающий семейство; h — независимая от λ функция. Тогда для оптимальной оценочной функции (4) справедливо

$$U' = \frac{1}{N^2} \int_X \psi^2 h dx = \int_X \text{IF}^2 h dx = U(\psi, h).$$

Отметим, что следствия из теоремы 1 при $h = 1$ описывают свойства условно оптимальных оценок, доказанные в [8] (теорема 1).

Следующий результат позволяет получить достаточные условия вогнутости (выпуклости) величины (3) как функции параметра, задающего семейство.

Теорема 2. Для оптимальной оценочной функции (4) справедливо одно из следующих неравенств:

$$U'' \leq \int_X \text{IF}^2 s'' dx, \text{ если } U > 0; \quad U'' \geq \int_X \text{IF}^2 s'' dx, \text{ если } U < 0.$$

Причем равенство в них достигается, только если s не зависит от x .

Доказательство. В следующих преобразованиях воспользуемся первым равенством в (7), не отбрасывая в нем нулевое слагаемое, а также равенствами (6) и (1):

$$\begin{aligned} \left(\frac{N}{c}\right)'' &= 2\beta'' \int_X \varphi f dx + 2(\beta')^2 \int_X \frac{f^2}{s} dx - 2\beta' \int_X \frac{\varphi f}{s} s' dx - \\ &- 2\beta' \int_X \frac{\varphi f}{s} s' dx + 2 \int_X \frac{\varphi^2}{s} (s')^2 dx - \int_X \varphi^2 s'' dx = \\ &= \int_X \frac{2}{s} (\beta' f - \varphi s')^2 dx - \int_X \varphi^2 s'' dx = \\ &= 2 \int_X (\varphi')^2 s dx - \int_X \varphi^2 s'' dx. \end{aligned}$$

Отсюда, используя (5), (7) и (8), находим:

$$\begin{aligned} \frac{1}{2} \left(\frac{N}{c}\right)^3 U'' &= \frac{1}{2} \left(\frac{N}{c}\right)^3 \left[- \left(\frac{N}{c}\right)^{-2} \left(\frac{N}{c}\right)' \right]' = \\ &= \left[\left(\frac{N}{c}\right)' \right]^2 - \frac{N}{2c} \left(\frac{N}{c}\right)'' = \\ &= \frac{N}{2c} \int_X \varphi^2 s'' dx + \left(\int_X \varphi \varphi' s dx \right)^2 - \int_X \varphi^2 s dx \int_X (\varphi')^2 s dx. \end{aligned}$$

Последние два слагаемых в полученном выражении есть неположительная величина в силу неравенства Коши–Буняковского. Равенство в нем достигается, только когда функции φ и φ' коллинеарны либо одна из них тождественно равна нулю, т. е. s не зависит соответственно от x либо от параметра, задающего семейство (в обоих случаях величина $\beta' = 0$). Случай, когда параметр отсутствует, не допускается по условию.

Таким образом, приходим к неравенству

$$\frac{1}{2} \left(\frac{N}{c}\right)^3 U'' \leq \frac{N}{2c} \int_X \varphi^2 s'' dx = \frac{1}{2} \left(\frac{N}{c}\right)^3 \int_X \text{IF}^2 s'' dx.$$

Домножим полученное выражение на величину $2U^3 = 2c^3/N^3$, знак которой соответствует знаку U . В результате получим неравенства в утверждении теоремы. Теорема доказана.

Следствие 1 из теоремы 2. Пусть $s = \gamma f + h \geq 0$, где $s \neq \text{const}$; γ — параметр, задающий семейство; h — независимая от γ функция. Тогда для оптимальной оценочной функции (4) $U'' < 0$.

Доказательство. Данное неравенство следует из теоремы 2, поскольку $s'' = 0$ и $U > 0$ в силу представления (3), кроме того, s зависит от x . Следствие доказано.

Следствие 2 из теоремы 2. Пусть $s = f + \lambda h$, где $s \neq \text{const}$; λ — параметр, задающий семейство; h — независимая от λ функция. Тогда для оптимальной оценочной функции (4) справедливо:

$$U'' < 0, \text{ если } s \geq 0; U'' > 0, \text{ если } s \leq 0.$$

Доказательство. Данные неравенства следуют из теоремы 2, поскольку $s'' = 0$, $U > 0$ при $s \geq 0$, $U < 0$ при $s \leq 0$, кроме того, s зависит от x . Следствие доказано.

Единственность членов условно оптимального семейства

Применим следствия 1 и 2 из теоремы 2 к условно оптимальному семейству, расширенному за счет отрицательных значений параметра [8]. В этом случае $h = 1$, а параметры, задающие семейство, удовлетворяют условиям

$$\begin{aligned} \gamma &> -1/\max_{x \in X} f; \\ \lambda &\geq 0 \text{ или } \lambda < -\max_{x \in X} f. \end{aligned}$$

Последние два неравенства соответствуют разным знакам U , поэтому исходную задачу минимизации величины (3) в данном случае следует понимать как $\int_X IF^2[s]dx \rightarrow \min_{\psi}$.

Согласно следствиям из теорем 1 и 2 справедливо неравенство

$$\frac{\partial^2 U}{\partial \gamma^2} = \frac{\partial V}{\partial \gamma} < 0$$

для всех допустимых значений $\gamma > -1/\max_{x \in X} f$; кроме того,

$$\frac{\partial^2 U}{\partial \lambda^2} = \frac{\partial W}{\partial \lambda} < 0, \text{ если } \lambda \geq 0; \frac{\partial W}{\partial \lambda} > 0, \text{ если } \lambda < -\max_{x \in X} f.$$

Отсюда следует единственность решения вариационных задач, определяющих семейство. Сформулируем это в виде следующих утверждений [8].

Утверждение 1. Задача с ограничениями-неравенствами. Пусть задана верхняя граница W_0 неустойчивости W или верхняя граница V_0 асимптотической дисперсии V оценки, так что соответственно $W_{\text{ОМУ}} \leq W \leq W_0$ или $V_{\text{ОМП}} \leq V \leq V_0$. Тогда каждое такое ограничение однозначно определяет условно оптимальную оценку, которой соответствуют некоторые неотрицательные значения параметров λ и γ .

Утверждение 2. Задача с ограничениями-равенствами. Пусть решается задача минимизации неустойчивости W при условии, что асимптотическая дисперсия оценки $V = V_0$, где $V_{\text{ОМП}} \leq V_0 < V_{\text{max}}$, V_{max} — предел асимптотической дисперсии оценки при $\gamma \rightarrow -1/\max_{x \in X} f$ или $\lambda \rightarrow -\max_{x \in X} f$. Тогда решением задачи является расши-

ренное условно оптимальное семейство оценок, причем каждое значение V_0 однозначно определяет оценку из данного семейства.

Оценивание параметров нормальной модели

На практике в качестве модельного распределения ошибок наблюдений часто выступает *нормальное распределение* [1] с плотностью

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right], x \in R, \quad (9)$$

где μ — параметр сдвига; σ — параметр масштаба. Потому для исследователя важно иметь в своем арсенале набор устойчивых оценок параметров данной модели.

В табл. 1 перечислены некоторые условно оптимальные оценки, соответствующие значениям параметра λ (задающего элемент семейства), а также оптимизационные задачи, приводящие к данным оценкам.

Присутствующая в таблице ОММ — наилучшая B -робастная оценка [4] в условно оптимальном семействе.

Рассмотрим семейство условно оптимальных оценок параметра сдвига μ в модели (9). Соответствующие оценочные функции, совпадающие с функциями влияния, определяются выражением (4) при $s = \lambda + f$ и $\beta = 0$, т. е. имеют вид

$$\psi(x, \mu) = \tilde{c} \frac{x - \mu}{1 + \lambda/f(x, \mu, \sigma)}, \quad (10)$$

где $\tilde{c} = c/\sigma^2$ — безразмерная константа, определяемая условием $N(\theta) = 1$ (она несущественна с точки зрения решения оценочного уравнения). Стандартное отклонение σ предполагается известным. Если оно неизвестно, может использоваться одна из его оценок (табл. 2) [2]. Для оценки параметра сдвига справедливы равенства $V_{\text{ОМП}} = \sigma^2$, $W_{\text{ОМУ}} = 4\sqrt{\pi}\sigma^3$ [2].

В табл. 2 для условно оптимальных оценок параметра сдвига μ приведено значение $\lambda\sigma$ (поскольку λ обратно пропорционально σ), значение $\tilde{c}$ и значения относительных характеристик эффективности и устойчивости. Кроме перечисленных в табл. 1 оценок, рассмотрена также одна безымянная оценка, относящаяся к расширению семейства в область $\lambda < 0$.

Для ОМУ общая формула (10) приводит к неопределенности, поэтому запишем отдельно выражение для данной функции влияния:

$$IF_{\text{ОМУ}}(x, \mu) = \sqrt{8}(x - \mu) \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right].$$

На рис. 1 показаны графики функций влияния рассмотренных в табл. 2 оценок при $\mu = 0$ и $\sigma = 1$, за

Таблица 1. Некоторые условно оптимальные оценки

Table 1. Some conditionally optimal estimators

Название	Сокращенное название	Уравнение для определения λ	Оптимизационная формулировка
Оценка максимального правдоподобия [1, 2]	ОМП	$\lambda = 0$	$V(\psi) \rightarrow \min_{\psi}$
Равнооптимальная оценка [11]	ОРО	$\text{eff}\psi = \text{stb}\psi$	$\min\{\text{eff}\psi, \text{stb}\psi\} \rightarrow \max_{\psi}$
Компромиссная оценка [2]	ОК	$\lambda = V_{\text{ОМП}}/W_{\text{ОМУ}}$	$\text{eff}^{-1}\psi + \text{stb}^{-1}\psi \rightarrow \min_{\psi}$
Равновесная оценка [17]	ОРВ	$\lambda = V/W > 0$	$V(\psi)W(\psi) \rightarrow \min_{\psi}$
Оценка максимальной устойчивости [2]	ОМУ	$\lambda = \infty$	$W(\psi) \rightarrow \min_{\psi}$
Минимаксная оценка	ОММ	$\max_x \text{IF} \rightarrow \min_{\lambda}$	—

Таблица 2. Характеристики некоторых оценок параметра сдвига

Table 2. Characteristics of some estimators of the shift parameter

$\lambda\sigma$	$\tilde{c}$	$\text{eff}\psi$, %	$\text{stb}\psi$, %	Название
0	1	100	0	ОМП
0,08304988	1,868355	85,04	85,04	ОРО
0,1410474	2,334381	81,15	90,74	ОК
0,1618861	2,496072	80,13	91,96	ОРВ
0,2916770	3,471798	75,93	95,97	ОММ
∞	∞	64,95	100	ОМУ
-0,5	-1,755425	42,51	86,30	—

исключением ОК и ОММ (чтобы линии графиков не располагались слишком плотно). ОК довольно близка к ОРВ, а ОММ занимает промежуточное положение между ОМУ и ОРВ.

Все рассмотренные оценки, кроме ОМП, являются сниженными (их функции влияния имеют асимптотой ось абсцисс), а следовательно, обладают свойством устойчивости к асимметричному засорению. Оценкам с большим значением λ (при $\lambda > 0$) соответствует большее значение характеристики $\text{stb}\psi$; соответствующие

функции влияния быстрее сходятся к нулю при увеличении $|x|$, а также их $|\text{IF}(x, \mu)|$ достигает максимума при меньшем значении $|x|$. Оценка, которой соответствует $\lambda = -0,5/\sigma$, с описанной точки зрения имеет наибольшую среди исследуемых оценок устойчивость (хотя показатель $\text{stb}\psi$ при $\lambda < 0$ не способен это отразить количественно, поскольку его максимум достигается на ОМУ), однако она имеет невысокую эффективность и сравнительно большое значение $\max_x |\text{IF}(x, \mu)|$, что не очень хорошо [4], но может быть оправдано при значительном засорении наблюдений.

Построим условно оптимальные оценки параметра масштаба σ в модели (9). Согласно (4) соответствующие оценочные функции, совпадающие с функциями влияния, могут быть записаны в виде

$$\psi(x, \sigma) = \tilde{c}\sigma \frac{(x - \mu)^2/\sigma^2 - \tilde{\beta}}{1 + \lambda/f(x, \mu, \sigma)}, \quad (11)$$

где $\tilde{\beta} = 1 - \beta\sigma$ — безразмерная константа, определяемая из условия (1); $\tilde{c}$ имеет тот же смысл, что и в (10). Для оценки параметра масштаба справедливы равенства $V_{\text{ОМП}} = \sigma^2/2$, $W_{\text{ОМУ}} = 4\sqrt{\pi}\sigma^3$ [2].

В табл. 3 приведены характеристики условно оптимальных оценок параметра масштаба σ .

Для ОМУ общая формула (11) приводит к неопределенности, поэтому запишем отдельно выражение для данной функции влияния:

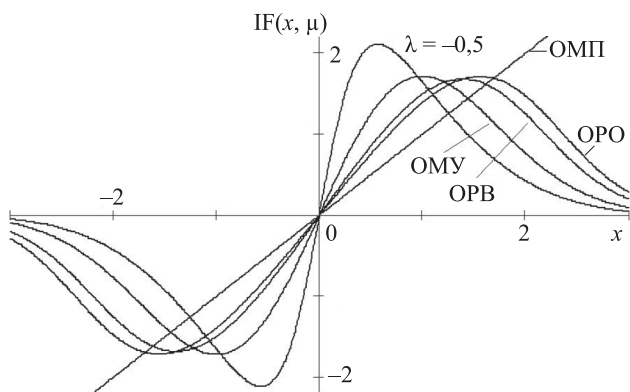


Рис. 1. Графики некоторых функций влияния параметра сдвига

Fig. 1. Graphs of some influence functions of the shift parameter

Таблица 3. Характеристики некоторых оценок параметра масштаба
Table 3. Characteristics of some estimators of the scale parameter

$\lambda\sigma$	$\tilde{\beta}$	$\tilde{c}$	eff ψ , %	stb ψ , %	Название
0	1	0,5	100	0	ОМП
0,04348800	0,7904612	1,101824	73,36	73,36	ОРО
0,07052370	0,7452787	1,349195	68,34	80,97	ОК
0,09066582	0,7211837	1,521623	65,72	84,49	ОРВ
∞	0,5	∞	43,30	100	ОМУ
-1,315140	0,4512637	-8,106702	38,68	99,20	ОММ
-0,5	0,2993957	-2,125707	24,90	85,75	—

$$IF_{\text{ОМУ}}(x, \sigma) = \sqrt{8}\sigma \left[\frac{(x - \mu)^2}{\sigma^2} - \frac{1}{2} \right] \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right].$$

На рис. 2 приведены графики функций влияния, перечисленных в табл. 3 оценок параметра масштаба при $\mu = 0$ и $\sigma = 1$, кроме ОК и ОММ.

ОК занимает промежуточное положение между ОРО и ОРВ, а ОММ довольно близка к ОМУ. Для изображенных на рис. 2 графиков, в целом, справедливы те же закономерности, что и для графиков на рис. 1.

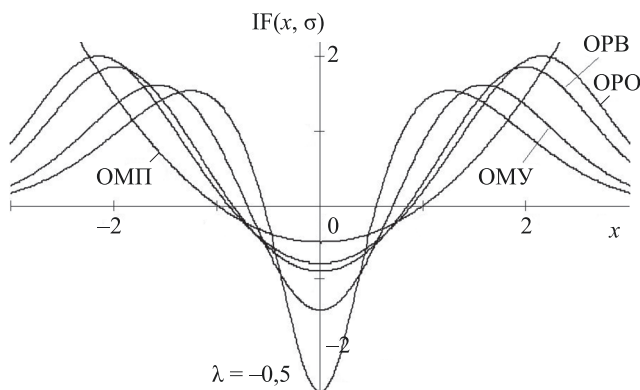


Рис. 2. Графики некоторых функций влияния параметра масштаба

Fig. 2. Graphs of some influence functions of the scale parameter

Если неизвестны оба параметра модели (9), тогда решается система оценочных уравнений для каждого из параметров.

Обсуждение

В работе получены следующие результаты о свойствах оценок, оптимизирующих весовую L_2 -норму функции влияния:

— доказаны две теоремы, согласно которым найдены выражение для производной по параметру функции

оценок качества оптимальной оценки и неравенства, необходимые для установления его выпуклости (вогнутости) по параметру;

- сформулированы следствия из теорем, относящиеся к условно оптимальному семейству оценок;
- в частности, для условно оптимального семейства установлены свойства монотонности функций неустойчивости и асимптотической дисперсии в зависимости от параметров, задающих семейство, откуда следует единственность элементов семейства;
- построены функции влияния ряда оптимальных оценок параметров сдвига и масштаба нормальной модели, исследованы характеристики этих оценок.

Доказанные теоремы оказались полезными при исследовании свойств компромиссных оценок на базе двух критериев, а также при изучении минимаксных уровней засорения в рамках модели БТЗ для задач, аналогичных исследованным в работе [18]. Соответствующие материалы готовятся к опубликованию. Можно считать перспективным использование полученных результатов в условиях, когда искажение данных производится целенаправленно действующим противником — в постановках задач, близких к используемым во вредоносном машинном обучении [19].

Заключение

Полученные в работе теоретические результаты служат инструментом для анализа однопараметрических семейств оценок, обладающих оптимальностью в смысле минимума весовой L_2 -нормы функции влияния. В первую очередь, данный инструмент позволяет ответить на вопрос о единственности членов семейства.

При подходящем выборе весовой функции оптимальные оценки оказываются сниженными, т. е. обладают устойчивостью к асимметричному засорению наблюдений. Это важно для практического применения данных оценок.

Теоретические результаты работы проиллюстрированы на примере ряда условно оптимальных оценок А.М. Шурыгина, для которых приведены все необходимые характеристики.

Литература

References

1. Borovkov A.A. *Mathematical Statistics*. Amsterdam: Gordon and Breach, 1998. 570 p. <https://doi.org/10.1201/9780203749326>
2. Шурыгин А.М. Прикладная стохастика: робастность, оценивание, прогноз. М.: Финансы и статистика, 2000. 224 с.
3. Huber P., Ronchetti E. *Robust Statistics*. 2nd ed. John Wiley & Sons, 2009. 354 p. <https://doi.org/10.1002/9780470434697>
4. Hampel F., Ronchetti E., Rousseeuw P., Stahel W. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 2005. 536 p. <https://doi.org/10.1002/9781118186435>
5. Лисицин Д.В., Гаврилов К.В. Максиминная задача оценивания параметров в условиях байесовского точечного засорения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 56–64. <https://doi.org/10.17223/19988605/62/6>
6. Лисицин Д.В., Гаврилов К.В. Оценивание параметров распределения ограниченной случайной величины, робастное к нарушению границ // Научный вестник Новосибирского государственного технического университета. 2016. № 2(63). С. 70–89. <https://doi.org/10.17212/1814-1196-2016-2-70-89>
7. Lisitsin D.V., Usol'tsev A.G. Minimum gamma-divergence estimation for non-homogeneous data with application to ordered probit model // *Applied methods of statistical analysis. Statistical computation and simulation: proceedings of the International Workshop. Novosibirsk, 18–20 Sept. 2019. Novosibirsk: NSTU, 2019. P. 227–234.*
8. Лисицин Д.В., Гаврилов К.В. О свойствах условно оптимальных оценок // Научный вестник Новосибирского государственного технического университета. 2015. № 1(58). С. 76–93. <https://doi.org/10.17212/1814-1196-2015-1-76-93>
9. Лисицин Д.В. Устойчивое оценивание параметров модели по многомерным неоднородным неполным данным // Научный вестник Новосибирского государственного технического университета. 2013. № 1(50). С. 17–30.
10. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: статистическая обработка неоднородных совокупностей. М.: Статистика, 1980. 210 с.
11. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров модели при асимметричном засорении данных // Научный вестник Новосибирского государственного технического университета. 2008. № 1(30). С. 33–40.
12. DasGupta A. *Asymptotic Theory of Statistics and Probability*. New York: Springer, 2008. 722 p. <https://doi.org/10.1007/978-0-387-75971-5>
13. Van der Vaart A.W. *Asymptotic Statistics*. Cambridge: Cambridge University Press, 1998. 443 p. <https://doi.org/10.1017/CBO9780511802256>
14. Shurygin A.M. New approach to optimization of stable estimation // *Proc. of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach. V. 3. Engineering and Scientific Applications*. Springer, Dordrecht, 1994. P. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
15. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending M -estimators // *Journal of Statistical Planning and Inference*. 2008. V. 138. N 10. P. 2906–2917. <https://doi.org/10.1016/J.JSPI.2007.11.008>
16. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016. 319 p. <https://doi.org/10.1002/9781119264507>
17. Гаврилов К.В., Веретельникова Е.Л. Об одном способе выбора компромисса в семействе условно оптимальных оценок // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. в печати.
18. Rieder H., Kohl M., Ruckdeschel P. The cost of not knowing the radius // *Statistical Methods and Applications*. 2008. V. 17. N 1. P. 13–40. <https://doi.org/10.1007/s10260-007-0047-7>
19. Есипов Д.А., Бучаев А.Я., Керимбай А., Пузикова Я.В., Сайдумаров С.К., Сулименко Н.С., Попов И.Ю., Кармановский Н.С. Атаки на основе вредоносных возмущений на системы обработки изображений и методы защиты от них // Научно-технический вестник информационных технологий, механики и оптики. 2023. № 4(23). С. 720–733. <https://doi.org/10.17586/2226-1494-2023-23-4-720-733>
1. Borovkov A.A. *Mathematical Statistics*. Amsterdam: Gordon and Breach, 1998. 570 p. <https://doi.org/10.1201/9780203749326>
2. Shurygin A.M. *Applied stochastics: robustness, estimation, prediction*. Moscow, Finansy i statistika Publ., 2000. 224 p. (in Russian)
3. Huber P., Ronchetti E. *Robust Statistics*. 2nd ed. John Wiley & Sons, 2009. 354 p. <https://doi.org/10.1002/9780470434697>
4. Hampel F., Ronchetti E., Rousseeuw P., Stahel W. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, 2005. 536 p. <https://doi.org/10.1002/9781118186435>
5. Lisitsin D.V., Gavrilov K.V. Maximin problem of parameter estimation in conditions of point Bayesian contamination. *Tomsk State University Journal of Control and Computer Science*, 2023, no. 62, pp. 56–64. (in Russian). <https://doi.org/10.17223/19988605/62/6>
6. Lisitsin D.V., Gavrilov K.V. Estimation of distribution parameters of a bounded random variable robust to bound disturbance. *Scientific Bulletin of NSTU*, 2016, no. 2(63), pp. 70–89. (in Russian) <https://doi.org/10.17212/1814-1196-2016-2-70-89>
7. Lisitsin D.V., Usol'tsev A.G. Minimum gamma-divergence estimation for non-homogeneous data with application to ordered probit model. *Applied methods of statistical analysis. Statistical computation and simulation. Proceedings of the International Workshop. Novosibirsk, 18–20 Sept. 2019. Novosibirsk: NSTU, 2019, pp. 227–234.*
8. Lisitsin D.V., Gavrilov K.V. On properties of conditionally optimal estimates. *Scientific Bulletin of NSTU*, 2015, no. 1(58), pp. 76–93. (in Russian). <https://doi.org/10.17212/1814-1196-2015-1-76-93>
9. Lisitsin D.V. Robust estimation of model parameters in presence of multivariate nonhomogeneous incomplete data. *Scientific Bulletin of NSTU*, 2013, no. 1(50), pp. 17–30. (in Russian)
10. Smolyak S.A., Titarenko B.P. *Stable estimation methods: statistical processing of heterogeneous aggregates*. Moscow, Statistika Publ., 1980, 210 p. (in Russian)
11. Lisitsin D.V., Gavrilov K.V. On stable estimation of models parameters in presence of asymmetric data contamination. *Scientific Bulletin of NSTU*, 2008, no. 1(30), pp. 33–40. (in Russian)
12. DasGupta A. *Asymptotic Theory of Statistics and Probability*. New York, Springer, 2008, 722 p. <https://doi.org/10.1007/978-0-387-75971-5>
13. Van der Vaart A.W. *Asymptotic Statistics*. Cambridge, Cambridge University Press, 1998, 443 p. <https://doi.org/10.1017/CBO9780511802256>
14. Shurygin A.M. New approach to optimization of stable estimation. *Proc. of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach. V. 3. Engineering and Scientific Applications*. Springer, Dordrecht, 1994, pp. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
15. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending M -estimators. *Journal of Statistical Planning and Inference*, 2008, vol. 138, no. 10, pp. 2906–2917. <https://doi.org/10.1016/J.JSPI.2007.11.008>
16. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016, 319 p. <https://doi.org/10.1002/9781119264507>
17. Gavrilov K.V., Veretel'nikova E.L. On one way to choose a compromise in a family of conditionally optimal estimators. *Tomsk State University Journal of Control and Computer Science*, 2024, no. 67, in press. (in Russian)
18. Rieder H., Kohl M., Ruckdeschel P. The cost of not knowing the radius. *Statistical Methods and Applications*, 2008, vol. 17, no. 1, pp. 13–40. <https://doi.org/10.1007/s10260-007-0047-7>
19. Esipov D.A., Buchaev A.Y., Kerimbay A., Puzikova Ya.V., Saidumarov S.K., Sulimenko N.S., Popov I.Yu., Karmanovskiy N.S. Attacks based on malicious perturbations on image processing systems and defense methods against them. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 720–733. (in Russian). <https://doi.org/10.17586/2226-1494-2023-23-4-720-733>

Авторы

Лисицин Даниил Валерьевич — доктор технических наук, профессор, профессор, Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация, [sc 57199594779](https://orcid.org/0009-0007-0723-6285), lisitsin@ami.nstu.ru

Гаврилов Константин Викторович — кандидат технических наук, доцент, Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация, <https://orcid.org/0009-0001-9252-9942>, aenigma77@mail.ru

Authors

Daniil V. Lisitsin — D.Sc., Full Professor, Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation, [sc 57199594779](https://orcid.org/0009-0007-0723-6285), lisitsin@ami.nstu.ru

Konstantin V. Gavrilov — PhD, Associate Professor, Novosibirsk State Technical University, Novosibirsk, 630073, Russian Federation, <https://orcid.org/0009-0001-9252-9942>, aenigma77@mail.ru

Статья поступила в редакцию 04.12.2023
Одобрена после рецензирования 18.02.2024
Принята к печати 17.03.2024

Received 04.12.2023
Approved after reviewing 18.02.2024
Accepted 17.03.2024



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»