

doi: 10.17586/2226-1494-2025-25-3-527-535

УДК 519.233.22

**О свойствах компромиссных 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 M -оценок, оптимизирующих весовую L_2 -норму функции влияния. Указанный критерий качества оценивания является достаточно общим и, кроме того, позволяет получать решения, относящиеся к классу сниженных оценок, т. е. обладающие свойством устойчивости к асимметричному засорению. Такие оценки исследовались в рамках локально устойчивого подхода д.т.н. А.М. Шурыгина, основанного на анализе показателя неустойчивости оценки (L_2 -нормы функции влияния), и его подхода, основанного на модели байесовского точечного засорения. В работе исследуется компромиссное семейство оценок, для которого оптимизируемый функционал представляет собой выпуклую линейную комбинацию двух базовых критериев. Компромиссное семейство подобно условно оптимальному семейству оценок, предложенному А.М. Шурыгиным, но используемые критерии могут быть квадратами весовых L_2 -норм функции влияния с произвольными заранее известными весовыми функциями. **Метод.** Рассмотренная предметная область до настоящего времени оставалась малоизученной. При проведении исследований применялась ранее разработанная авторами настоящей работы теория, которая описывает свойства оценок, оптимизирующих весовую L_2 -норму функции влияния. **Основные результаты.** Получен ряд свойств компромиссных оценок, показана единственность элементов семейства. Отдельно рассмотрен член семейства, доставляющий равные значения критериям: показано, что данная оценка соответствует седловой точке оптимизируемого функционала, а также является минимаксным решением относительно базовых критериев на множестве всех регулярных оценочных функций. **Обсуждение.** Построенная теория проиллюстрирована на примере задачи оценивания математического ожидания нормального распределения в условиях целенаправленного вредоносного воздействия на набор данных (аналогичного атаке отравления данных во вредоносном машинном обучении).

Ключевые слова

M -оценки, робастная статистика, функция влияния, устойчивые оценки, сниженные оценки, седловая точка, отравление данных

Ссылка для цитирования: Лисицин Д.В., Гаврилов К.В. О свойствах компромиссных M -оценок, оптимизирующих весовую L_2 -норму функции влияния // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 527–535. doi: 10.17586/2226-1494-2025-25-3-527-535

**On the properties of compromise 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 paper develops a theory of M -estimators optimizing the weighted L_2 -norm of the influence function. The specified criterion of the estimation quality is quite general and, in addition, allows obtaining solutions related to the class of redescending estimators, i.e., possessing the property of stability to asymmetric contamination. Such estimators, in particular, were studied 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). In this paper, a compromise family of estimators is studied for which the optimized functional is a convex linear combination of two basic criteria. The compromise family is similar to the conditionally optimal family of estimators proposed by A.M. Shurygin, but the criteria used can be squares of the weighted L_2 -norms of the influence function with arbitrary pre-known weight functions. The considered subject area has remained little-studied to date. In the course of the research, we used a theory we had developed earlier, which describes the properties of estimators that optimize the weighted L_2 -norm of the influence function. As a result of the study, a number of properties of compromise estimators were obtained, and the uniqueness of the family elements was shown. A family member that delivers equal values to the criteria was considered separately: it was shown that this estimator corresponds to the saddle point of the optimized functional, and is also a minimax solution with respect to the basic criteria on the set of all regular score functions. The constructed theory is illustrated using the example of the problem of mathematical expectation estimating of a normal distribution under conditions of targeted malicious influence on a data set (similar to a data poisoning attack in malicious machine learning).

Keywords

M -estimators, robust statistics, influence function, stable estimates, redescending estimators, saddle point, data poisoning

For citation: Lisitsin D.V., Gavrilov K.V. On the properties of compromise M -estimators optimizing weighted L_2 -norm of the influence function. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 527–535 (in Russian). doi: 10.17586/2226-1494-2025-25-3-527-535

Введение

Современные подходы к решению задачи оценивания параметров статистических моделей [1] позволяют обеспечить устойчивость получаемых решений к возможным отклонениям реальной ситуации от принятых в модели предположений [2–4]. Среди таких отклонений часто рассматривают *засорение*, под которым будем понимать присутствие в массиве данных посторонних (засоряющих) наблюдений; с теоретической точки зрения в результате получается смесь исходного распределения наблюдений и засоряющего распределения. Одним из наиболее трудных для формализации являлось асимметричное засорение симметрично распределенных данных, поскольку априорная информация о возможном искажении, как правило, отсутствует. Классические робастные подходы в этой ситуации не обладают достаточной устойчивостью. Например, для математического ожидания нормально-го распределения оптимальными оказываются оценка Хьюбера — оценка максимального правдоподобия для плотности, нормальной в центре и экспоненциально убывающей на хвостах, и медиана — оценка максимального правдоподобия для распределения Лапласа, также имеющего экспоненциально убывающую плотность. Однако на практике хвосты могут быть более тяжелыми [3].

Для решения этой проблемы были предложены *сниженные оценки* [3, 4], в которых влияние периферийных наблюдений значительно снижено по сравнению с классическими робастными решениями или равно нулю. Однако все они получены в большей или меньшей степени эвристически.

Между тем наш соотечественник д.т.н. Александр Михайлович Шурыгин предложил довольно удачную *модель байесовского точечного засорения* (БТЗ) [2], которая позволяет описывать окрестность модельного распределения, возникающую в результате асимметричного засорения, и, в частности, рассматривать ситуации целенаправленного вредоносного воздействия (атаки) на набор данных [5].

Модель БТЗ предполагает наличие серии выборок, каждая из которых содержит засорение, сосредоточен-

ное 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 на практике оценки часто оказываются сниженными. В общем случае свойства оценок зависят от соотношения между модельной плотностью и плотностью распределения засоряющей точки (плотностью БТЗ). При неудачном выборе плотности БТЗ оценки могут не быть сниженными и даже вообще устойчивыми — например, когда оба распределения нормальные, но второе имеет меньшую дисперсию [6].

В рамках модели БТЗ являются оптимальными многие известные сниженные оценки параметра сдвига, среди них: бивес-оценка Тьюки, оценки Эндрюса, Бернуллы, Смита, Мешалкина (Уэлша), Фэйра, Рамсея, оценка Хьюбера типа урезанного среднего, обобщенные оценки Шарбонье [2, 5, 7].

Подход на основе БТЗ может быть сведен к анализу весовой L_2 -нормы функции влияния [4], где плотность БТЗ выступает в роли весовой функции. Поскольку последняя не обязательно должна быть нормированной, такая интерпретация удобна тем, что распространяет понятие наилучшей оценочной функции на весовые функции, не являющиеся плотностью, в том числе неинтегрируемые [5].

Таким образом, указанный критерий качества оценивания может служить теоретической основой [8] для конструирования широкого спектра оптимальных устойчивых оценок, имеющих важное практическое значение, в том числе робастных решений, которые характеризуются заданием непараметрической («полноразмерной») окрестности модельного распределения [3, 4]. При этом наиболее продуктивный путь построения оптимальных оценок заключается в разумном выборе весовой функции и последующем нахождении для нее оценочной функции, минимизирующей весовую L_2 -норму функции влияния. В частности, весовая функция может определяться как результат оптимизации самостоятельного критерия, например, некоторой меры неопределенности [5, 7].

Обзор основных результатов

В настоящей работе исследуются свойства *компромиссных оценок*, оптимальных с точки зрения компромиссного критерия в виде выпуклой линейной комбинации двух *базовых критериев* [9]. И компромиссный, и оба базовых критерия представляют собой квадраты весовых L_2 -норм функции влияния [8]. Для базовых критериев весовые функции заранее известны, а их выпуклая линейная комбинация выступает в роли весовой функции компромиссного критерия. Такая формулировка — способ задания семейства, подобного *условно оптимальным оценкам* [2, 10, 11], поскольку оптимизируемый функционал для них тоже может быть записан в виде выпуклой линейной комбинации нормированных функционалов *асимптотической дисперсии и неустойчивости* (квадрата L_2 -нормы функции влияния) [12].

Априорный выбор параметра линейной комбинации на практике может оказаться затруднительным. Важным результатом работы является возможность нахождения оценки, которая бы разрешала такую неопределенность. Указанная оценка соответствует седловой точке оптимизируемого функционала и названа *равнооптимальной* (ОРО). Она также удовлетворяет решению минимаксной задачи относительно базовых критериев.

С помощью компромиссного семейства обычно удается исправить рассмотренную ситуацию, когда используемая плотность БТЗ приводит к неустойчивой оценке. В результате можно перейти к компромиссу между исходным критерием качества оценки и критерием с константной весовой функции. Похожий способ обеспечения устойчивости рассмотрен в работе [13].

Критерий, аналогичный рассматриваемому компромиссному критерию, возникает в случае применения асимптотики сужающихся окрестностей засорения [3, 4, 14] к модели БТЗ. Полученные результаты позволяют решить задачу нахождения минимаксной оценки для такой модели при неизвестном размере окрестности, выражающемся через параметр линейной комбинации (подобно задаче поиска минимаксного размера окрестности в [14]).

Данная работа является продолжением исследований, опубликованных в работе [8]. Помимо изложения теории, в работе приведен пример построения устойчивых оценок параметра сдвига нормального распределения в условиях отвлекающего воздействия на набор данных.

Элементы теории устойчивого оценивания

Пусть $x_1, \dots, x_m$ — наблюдения случайной величины ξ , которая моделирует измеряемую числовую характеристику в некотором эксперименте. Для ее описания используется модельное распределение с плотностью $f(x, \theta)$, заданной в прямоугольнике $X \times \Theta$, где $X \subseteq R$ — множество значений переменной x (это возможные значения измеряемой характеристики), $\Theta \subseteq R$ — множество значений параметра плотности θ и R — множество действительных чисел. M -оценка неизвестного параметра θ может определяться [1, 3] как решение оптимизационной задачи

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^m \rho(x_i, \theta),$$

где $\rho(x_i, \theta): X \times \Theta \rightarrow R$ — некоторая функция, называемая функцией потерь.

Условие равенства нулю производной по θ оптимизируемой функции служит альтернативной (хотя и не эквивалентной [1]) формулировкой задачи M -оценивания:

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

называемой *оценочным уравнением*, где функция

$$\psi(x, \theta) = c(\theta) \partial \rho(x, \theta) / \partial \theta \quad (1)$$

называется *оценочной функцией* для параметра θ ; $c(\theta)$ — произвольная непрерывная функция, не равная нулю для всех $\theta \in \Theta$. Множитель $c(\theta)$ задает семейство *эквивалентных оценочных функций*.

Также оценочная функция (1) при модельном распределении должна удовлетворять условию асимптотической несмещенности, которое принимает вид [1, 15]

$$\mathbf{E} \psi(\xi, \theta) = \int_X \psi(x, \theta) f(x, \theta) dx = 0, \quad (2)$$

где $\mathbf{E}$ — оператор математического ожидания по плотности $f(x, \theta)$.

Дифференцируя (2) по θ и допуская возможность изменения порядка дифференцирования и интегрирования, можно записать следующие равенства [3, 15]:

$$\begin{aligned} N(\theta) &= -\lim_{t \rightarrow \theta} \frac{d}{dt} \mathbf{E} \psi(\xi, t) = -\mathbf{E} \frac{\partial}{\partial \theta} \psi(\xi, \theta) = \\ &= \int_X \psi(x, \theta) \frac{\partial}{\partial \theta} f(x, \theta) dx. \end{aligned} \quad (3)$$

Потребуем, чтобы в окрестности истинного значения параметра θ выполнялось условие асимптотической несмещенности (2), были справедливы равенства (3), функция $N(\theta)$ была непрерывной и не равной нулю [10]. Более полный набор условий регулярности, обеспечивающий состоятельность и асимптотическую нормальность оценок, приведен, например в [5, 16]. Специфика случая конечного X подробно рассмотрена в [17].

Асимптотическая дисперсия [1] M -оценки $\hat{\theta}$ определяется выражением

$$V(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) f(x, \theta) dx = \mathbf{E} \text{IF}^2(\xi, \theta),$$

где $\text{IF}(x, \theta) = \psi(x, \theta) / N(\theta)$ — функция влияния Хампеля [4]. Функция влияния является показателем количественной робастности и отражает воздействие на оценку бесконечно малого точечного засорения. Величина $\gamma^* = \sup_{x \in X} |\text{IF}(x, \theta)|$ называется *чувствительностью к большой ошибке*, для качественно робастной оценки $\gamma^* < \infty$ [4].

Неустойчивостью оценки [2, 18, 19] называется величина

$$W(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) dx = \int_X IF^2(x, \theta) dx.$$

Показатель $W(\psi)$ представляет собой квадрат L_2 -нормы функции влияния, его минимизация приводит к оценке максимальной устойчивости (ОМУ) [2].

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

$$U(\psi, s) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) s(x, \theta) dx = \int_X IF^2(x, \theta) s(x, \theta) dx. \quad (4)$$

Такой функционал возникает, например, как асимптотическое квадратичное отклонение оценки в модели БТЗ [2, 5, 9, 16, 19], где весовая функция $s(x, \theta)$ представляет собой плотность распределения засоряющей точки.

В [2] показано, что функционал (4) достигает минимума по ψ на функции

$$\psi_s(x, \theta) = c(\theta) \left[\frac{\partial}{\partial \theta} \ln f(x, \theta) + \beta(\theta) \right] \frac{f(x, \theta)}{s(x, \theta)}, \quad (5)$$

где $c(\theta)$ — функция, имеющая тот же смысл, что и в (1); $\beta(\theta)$ — функция, определяемая из условия (2) путем подстановки в него оценочной функции (5) и решения получившегося уравнения относительно $\beta(\theta)$. Оценочную функцию (5) назовем оптимальной для заданной функции s .

Другие функции, зависящие от оценочной функции, далее отметим нижним индексом « s », если в них подставлена оптимальная оценочная функция (5). Результат (5) по построению справедлив для функций $s(x, \theta) \geq 0$, и далее будут рассматриваться только такие весовые функции. Для краткости записи выражений будем опускать аргументы функций там, где они понятны. Например, (5) можно записать в виде $\psi_s = c(\partial f / \partial \theta + \beta f) / s$. Перед обсуждением свойств компромиссных оценок приведем важные в настоящей работе результаты, доказанные в [8].

Пусть k — некоторый скалярный параметр, от которого зависит весовая функция s . Предположим, что все приведенные далее функции непрерывно зависят от параметра на своей области определения. Штрихом обозначим производную функций по данному параметру. Также для удобства введем обозначение $U(k) = U(\psi_s, s)$, поскольку это функция параметра.

Теорема 1. Для оценочной функции (5) справедливо

$$U'(k) = \frac{1}{N_s^2} \int_X \psi_s^2 s' dx = \int_X IF_s^2 s' dx.$$

Теорема 2. Для оценочной функции (5) справедливо неравенство

$$U''(k) \leq \int_X IF_s^2 s'' dx.$$

Причем равенство в нем достигается, только если функции s и s' не зависят от x . Это следует из усло-

вия коллинеарности функций ϕ и ϕ' [8, формула (6)]. Иными словами, неравенство в теореме 2 обращается в равенство, только когда весовые функции всех членов семейства не зависят от x (т. е. все члены семейства — ОМУ). Данный случай не представляет интерес, поэтому в качестве дополнительного условия регулярности примем, что неравенство в теореме 2 строгое.

Свойства компромиссного семейства оценок

Определим семейство компромиссных оценок как решение, оптимизирующее компромиссный критерий

$$U(\psi, k) = (1 - k)U_1(\psi) + kU_2(\psi), \quad (6)$$

где $U_1(\psi) = U(\psi, s_1)$ и $U_2(\psi) = U(\psi, s_2)$ — базовые критерии качества оценки, которым соответствуют весовые функции $s_1(x, \theta)$ и $s_2(x, \theta)$ (не зависящие от k); k — параметр, задающий семейство и определяющий компромисс между базовыми критериями.

Компромиссный критерий (6) является частным случаем (4) с весовой функцией

$$s = (1 - k)s_1 + ks_2, \quad (7)$$

и, таким образом, результат оптимизации критерия (6) имеет вид (5).

Определение 1. Компромиссной оценкой называется оценка с оценочной функцией вида (5) и весовой функцией (7), причем функции s , s_1 и s_2 должны быть одновременно неотрицательными.

Исходя из данного определения, для функции (7) выбирается диапазон допустимых значений параметра k . Очевидно, этот диапазон обязательно содержит отрезок $[0; 1]$, который традиционно определяет компромисс [9]. Более того, отрезок $[0; 1]$ соответствует парето-оптимальным решениям задачи двухкритериальной оптимизации [20]. Это непосредственно следует из условия оптимальной формулировки семейства, которая сводится к оптимизационной задаче $U_1(\psi) + \lambda U_2(\psi) \rightarrow \min_{\psi}$, где $\lambda \geq 0$ — параметр, задающий семейство. В текущей работе рассматривается расширенное множество значений параметра k , которое может содержать члены $k < 0$ и $k > 1$.

Наряду с $U(k)$, введем обозначения

$$U_1(k) = U_1(\psi_s) = U(\psi_s, s_1), \quad U_2(k) = U_2(\psi_s) = U(\psi_s, s_2).$$

Свойство 1. Применим теорему 1 к семейству оценок с весовой функцией (7):

$$U'(k) = \int_X IF_s^2 s' dx = \int_X IF_s^2 s'(s_2 - s_1) dx = U_2(k) - U_1(k).$$

Свойство 2. Продифференцируем равенство (6) по k и приравняем выражению, определяемому свойством 1. После приведения подобных получим

$$(1 - k)U_1'(k) + kU_2'(k) = 0.$$

Свойство 3. Применим теорему 2 к компромиссным оценкам с весовой функцией (7), учитывая, что $s'' = 0$. Получим $U''(k) < 0$.

Свойство 4. Согласно свойству 3 величина $U'(k)$ является убывающей функцией параметра k .

Свойство 5. Из свойств 1 и 3 следует, что при любом k имеет место $U'_2(k) < U'_1(k)$.

Свойство 6. Из свойств 2 и 5 следует, что $U_1(k)$ и $U_2(k)$ являются монотонными функциями параметра k на каждом из следующих промежутков: $k \leq 0$, $0 \leq k \leq 1$, $k \geq 1$. Рассмотрим эти промежутки подробнее.

Для $k \leq 0$ и $k \geq 1$, согласно свойству 2 функции $U'_1(k)$ и $U'_2(k)$ имеют совпадающие промежутки знакопостоянства и одинаковые знаки на них. Кроме того, данные функции в силу свойства 5 не имеют общих точек, а следовательно, одновременно являются положительными либо отрицательными. Таким образом, функции $U_1(k)$ и $U_2(k)$ возрастают или убывают одновременно. Причем, поскольку функционал (6) минимизируется, функции $U_1(k)$ и $U_2(k)$ являются убывающими при $k \leq 0$ и возрастающими при $k \geq 1$.

Для $0 \leq k \leq 1$, согласно свойству 2, функции $U'_1(k)$ и $U'_2(k)$ имеют совпадающие промежутки знакопостоянства и разные знаки на них. Кроме того, данные функции в силу свойства 5 не имеют общих точек, а следовательно, одна из них положительная, другая отрицательная. Заметим, что $U_1(k)$ достигает минимума, совпадающего с минимумом функционала (6), при $k = 0$, а $U_2(k)$ — аналогично при $k = 1$. Следовательно, функция $U_1(k)$ возрастает, а $U_2(k)$ — убывает.

Следующий результат устанавливает единственность членов компромиссного семейства оценок.

Теорема 3. Решение задачи оптимизации функционала (6) имеет взаимно однозначное соответствие с семейством компромиссных оценочных функций.

Доказательство. При каждом фиксированном k функционалу (6) соответствует выражение (4) с весовой функцией (7), а потому, согласно [2], задача его оптимизации имеет единственное решение вида (5). Нужно показать, что каждому такому оптимальному решению отвечает единственное значение k , определяющее компромисс.

Допустим, двум различным значениям k в выражении (6) отвечает одна оптимальная функция влияния IF_0 . Согласно свойству 1 для данных значений k

$$U'(k) = \int_X IF_0^2 (s_2 - s_1) dx,$$

где правая часть равенства одна и та же, однако левая часть согласно свойству 4 является монотонной функцией от k , следовательно принимает разные значения для разных k . Пришли к противоречию. Теорема доказана.

Равнооптимальная оценка

Ранее для условно оптимального семейства [2, 11] было введено понятие ОРО [10] как члена семейства, который доставляет равные значения относительно характеристикам устойчивости $stb \psi = W_{ОМУ}/W(\psi)$ и эффективности $eff \psi = V_{ОМП}/V(\psi)$, где $W_{ОМУ}$ и $V_{ОМП}$ — неустойчивость ОМУ и асимптотическая дисперсия ОМП. Распространим это понятие на компромиссные оценки.

Определение 2. ОРО называется оценка из компромиссного семейства, удовлетворяющая условию $U_1(k) = U_2(k)$.

Значение параметра k , при котором достигается это равенство, обозначим $k_{ОРО}$. Обратим внимание, что существование ОРО в семействе не гарантируется.

Свойство 7. Если ОРО существует, она единственна и соответствует решению максиминной задачи относительно аргументов функционала (6):

$$k_{ОРО} = \arg \max_k \min_{\psi} U(\psi, k).$$

Соответствующее ОРО значение критерия (6) обозначим $U_{ОРО} = \max_k U(k)$.

Действительно, согласно свойству 1 условие $U_1(k) = U_2(k)$ эквивалентно условию $U'(k) = 0$, которое вместе со свойством 3 образует достаточное условие максимума функции $U(k)$. С учетом свойства 4 условие $U'(k) = 0$ выполняется не более чем для единственного значения k , и оно является допустимым, поскольку функции $U(k)$, $U_1(k)$ и $U_2(k)$ неотрицательные (они равны в этой точке).

Свойство 8. Если ОРО существует, то ее оценочная функция $\psi_{ОРО}$ является решением минимаксной задачи относительно аргументов функционала (6):

$$\psi_{ОРО} = \arg \max_k \min_{\psi} U(\psi, k),$$

причем в точке решения функционал (6) принимает одно и то же значение, равное $U(\psi_{ОРО}, k)$ независимо от значения k .

Действительно, используя (6), можно записать

$$U(\psi, k) = U_1(\psi) + k[U_2(\psi) - U_1(\psi)] \leq U_1(\psi) + |k| |U_2(\psi) - U_1(\psi)|. \quad (8)$$

Второе слагаемое мажорирующего функционала в (8) достигает минимума, равного 0, при условии

$$U_1(\psi) = U_2(\psi) \quad (9)$$

для любого k (т. е. значение k является неопределенным). Полученному условию, в частности, отвечает ОРО, однако, кроме ОРО, это условие могло бы выполняться и для оценочных функций, не принадлежащих компромиссному семейству. Исходя из этого, нужно убедиться, что решение принадлежит данному семейству.

Когда условие (9) выполнено, неравенство в (8) обращается в равенство, в правой части которого остается только первое слагаемое, поэтому задача принимает вид

$$U_1(\psi) = (1 - k)U_1(\psi) + kU_2(\psi) = U(\psi, k) \rightarrow \min_{\psi},$$

ее решение для каждого k по определению принадлежит компромиссному семейству.

Теорема 4. ОРО соответствует седловой точке функционала (6) относительно аргументов ψ и k , при этом имеет место

$$U(\psi_{ОРО}, k) = U(\psi_{ОРО}, k_{ОРО}) \leq U(\psi, k_{ОРО}).$$

Утверждение теоремы 4 следует из свойств 7 и 8.

Свойство 9. Если $0 \leq k_{\text{ОРО}} \leq 1$, то

$$U_{\text{ОРО}} = \min_k \max \{U_1(k), U_2(k)\} = \max_k \min \{U_1(k), U_2(k)\}.$$

Справедливость свойства 9 следует из свойства 6 для отрезка $0 \leq k \leq 1$.

Теорема 5. Если $0 \leq k_{\text{ОРО}} \leq 1$, то

$$\Psi_{\text{ОРО}} = \arg \min_{\Psi} \max \{U_1(\Psi), U_2(\Psi)\}.$$

Доказательство. По условию теоремы 5 значение k , соответствующее решению оптимизационной задачи в свойстве 8, принадлежит отрезку $[0; 1]$, поэтому решение этой задачи не изменится, если множество значений k , на котором производится оптимизация, ограничить указанным отрезком, т. е.

$$\Psi_{\text{ОРО}} = \arg \min_{\Psi} \max_{k \in [0; 1]} U(\Psi, k).$$

С учетом представления (6) очевидно, что минимизируемые функционалы в последнем выражении и в правой части утверждения теоремы 5 совпадают для каждой Ψ :

$$\max \{U_1(\Psi), U_2(\Psi)\} = \max_{k \in [0; 1]} U(\Psi, k).$$

Теорема 5 доказана.

Свойство 10. Если $\min_{\Psi} U_1(\Psi) = \min_{\Psi} U_2(\Psi)$, то $0 < k_{\text{ОРО}} < 1$.

Действительно, по условию имеем $U_1(0) = U_2(1)$. Рассмотрим разность $u(k) = U_2(k) - U_1(k)$. В силу свойства 6 на промежутке $0 \leq k \leq 1$ функция $U_1(k)$ возрастает, а $U_2(k)$ убывает, поэтому

$$u(0) = U_2(0) - U_1(0) = U_2(0) - U_2(1) > 0$$

и

$$u(1) = U_2(1) - U_1(1) = U_1(0) - U_1(1) < 0,$$

причем функция $u(k)$ убывает. Следовательно, значение $k_{\text{ОРО}}$, являющееся решением уравнения $u(k) = 0$, принадлежит интервалу $0 < k < 1$.

Данное свойство, в частности, делает удобным применение базовых критериев вида $\text{stb}_s^{-1}\Psi = U(\Psi, s)/U(\Psi_s, s)$ при некотором U , определяемом (4) [9]. Такие базовые критерии имеют минимальное значение, равное 1.

Обратим внимание, что для практического применения ОРО базовые критерии предварительно следует привести к единой шкале путем нормирования — в данном случае, умножением (делением) на положительные константы. С физической точки зрения оба слагаемых в (6) должны быть безразмерными или иметь одинаковую размерность, что обычно тоже обеспечивается нормированием [21]. Нормирование базовых критериев производится исследователем на основе имеющейся априорной информации и является неформальной процедурой. При недостатке априорной информации исследователь на свое усмотрение может воспользоваться некоторыми эвристиками: например, критерии могут нормироваться к своим минимальным значениям (это

будут функционалы вида $\text{stb}_s^{-1}\Psi$), или могут использоваться функционалы, которым соответствуют весовые функции s_1 и s_2 , являющиеся плотностями.

ОРО представляет собой объективный выбор одного члена компромиссного семейства, и, таким образом, разрешается неопределенность в выборе k . Получившийся компромисс удовлетворяет ряду свойств оптимальности. Так, теорема 4 гарантирует, что значение критерия (6) будет наилучшим при неблагоприятном значении $k = k_{\text{ОРО}}$, а любое другое значение k не ухудшит (правда, и не улучшит) значение этого критерия, т. е. ОРО — так называемое гарантирующее решение. Теорема 5 показывает, что базовые критерии одновременно не могут быть улучшены. Важность теоремы 5 состоит еще и в том, что она позволяет сформулировать ОРО с использованием только базовых критериев, без привлечения компромиссного.

Пример

Рассмотрим задачу оценивания математического ожидания нормального распределения в условиях БТЗ [2], плотность которого нормальная. Пусть модельная плотность имеет математическое ожидание θ и дисперсию σ_0^2 :

$$f(x, \theta) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{(x - \theta)^2}{2\sigma_0^2} \right], \quad (10)$$

а плотность БТЗ — математическое ожидание $\theta + \mu$ и дисперсию σ^2 :

$$s(x, \theta) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{(x - \theta - \mu)^2}{2\sigma^2} \right]. \quad (11)$$

Выбор нормальной плотности БТЗ можно обосновать непараметрическим способом (подобно тому, как это делается в теории робастности) — путем оптимизации некоторых информационных мер при дополнительных ограничениях по аналогии с изложенным в [5]. При этом главным отличием от [5] является необходимость ограничения на величину μ . Например, в соответствии с принципом максимальной энтропии можно максимизировать дифференциальную энтропию плотности $s(x, \theta)$ в классе распределений с фиксированным μ и дисперсией, ограниченной сверху величиной σ^2 .

Случай несовпадающих центров плотностей (10) и (11) интересен тем, что позволяет моделировать целенаправленное вредоносное воздействие (атаку) на набор данных. Обработка данных в подобных условиях все чаще упоминается в работах по робастности, по-видимому, из-за интенсивного развития такого направления, как вредоносное машинное обучение, в котором использование робастных методов считается перспективным [5, 22]. Рассматриваемый в настоящей работе способ воздействия в рамках этого направления называется *атакой отравления данных*. В связи с этим отметим, что некоторые специалисты [22] связывают трудности исследований в данном направлении со сложностью систем глубокого обучения: по их мнению, вначале целесообразно обратиться к изучению более простых моделей, и данный пример можно рассматри-

вать как шаг к реализации этой идеи на основе подхода А.М. Шурыгина.

В данном случае цель злоумышленника — вызвать смещение оценок математического ожидания, а воздействие моделируется точечным засорением как простейшим непараметрическим способом достигнуть этой цели (вообще, точечное засорение оказывается в некотором смысле наихудшим [15]). Статистик, полагая неизвестную ему засоряющую точку случайной, выбирает на основе некоторой априорной информации значения μ и σ^2 , определяющие ее распределение (подобный подход использован в [23]). Так, могут быть выбраны $\mu > 0$ или $\mu < 0$ в зависимости от того, что, по мнению статистика, выгодно злоумышленнику. Но если информации о цели злоумышленника нет, то статистик вынужден положить $\mu = 0$, получая известные оценки Мешалкина [2, 24]. В последнем случае предполагать $\sigma^2 \leq \sigma_0^2$ вряд ли имеет смысл, поскольку тогда в среднем искажение не приведет к ухудшению оценок.

Изучим робастность оптимальной оценки (5) с точки зрения значения γ^* . При $\sigma^2/\sigma_0^2 \leq 2/3$ модель не регулярна, при $2/3 < \sigma^2/\sigma_0^2 \leq 1$ оценкам соответствует неограниченное значение γ^* , но даже при $\sigma^2 > \sigma_0^2$ они могут иметь слишком большое, хотя и конечное значение γ^* .

В соответствии со сделанным анализом, как минимум в случае $\mu \neq 0$, защищенность оценок от воздействия злоумышленника не гарантирует их робастности. Однако робастность полезна в присутствии, наряду со злонамеренным, отличного от него «природного» искажения (а также как страховка от ошибок спецификации плотности БТЗ). Будем считать, что природа безразлична к задаче статистика, однако стремится к увеличению неопределенности, поэтому проявляет себя в равномерном распределении, максимизирующем дифференциальную энтропию [5]. Как отражающий качество оценивания при несобственной равномерной (равной 1) плотности БТЗ может рассматриваться функционал $W(\psi)$ (подробнее в [5]).

Итак, для достижения баланса между двумя аспектами устойчивости — защищенностью и робастностью — воспользуемся компромиссным функциона-

лом (6), в котором базовыми функционалами выберем критерий защищенности $U_1(\psi) = \text{stb}_s^{-1}\psi$, где s — плотность (11), и критерий робастности $U_2(\psi) = \text{stb}^{-1}\psi$. Функционалы $\text{stb}_s\psi$ [9] и $\text{stb}\psi$ [2] принимают значения в диапазоне $[0; 1]$, подобно $\text{eff}\psi$, и удобны в качестве относительных характеристик защищенности и робастности.

Предположим, что степени воздействия злоумышленника и природы статистику неизвестны, и он выбирает ОРО как гарантирующее решение.

На рисунке для $\theta = 0$ и $\sigma_0^2 = 1$ сравниваются функции влияния трех оценок: ОРО, оптимальной оценки для плотности БТЗ (11) и ОМУ (сплошные линии 1, 2 и 3). Также представлены графики плотностей (10) и (11), умноженных на 4 — пунктирные линии 4 и 5. Рисунок *a* соответствует случаю $\mu = 0,82$, $\sigma^2 = 0,6724$ с $k_{\text{ОРО}} \approx 0,249$, $\text{stb}_s\psi_{\text{ОРО}} = \text{stb}\psi_{\text{ОРО}} \approx 0,155$; рисунок *b* — случаю $\mu = 5$, $\sigma^2 = 4$ с $k_{\text{ОРО}} \approx 0,454$, $\text{stb}_s\psi_{\text{ОРО}} = \text{stb}\psi_{\text{ОРО}} \approx 0,665$.

В случае, соответствующем рисунку *a*, оптимальная оценка для плотности (11) неробастна ($\gamma^* = \infty$), поэтому необходимо использовать, например, компромиссную оценку. Однако ОРО имеет хотя и ограниченное, но достаточно большое значение γ^* , модуль функции влияния уже не мал в той области на левой части графика, где еще мала плотность $f(x, \theta)$. Возможно, в этом случае выгодно было бы использовать $k > k_{\text{ОРО}}$. Отметим характерное поведение функций влияния для этих двух оценок — прогиб под центральной областью плотности $s(x, \theta)$ для уменьшения ее влияния. Напротив, функция влияния ОМУ достигает максимального значения недалеко от центра $s(x, \theta)$, поэтому с большой вероятностью может оказаться слишком чувствительной к воздействию злоумышленника.

В случае, соответствующем рисунку *b*, оптимальная оценка для плотности (11) имеет, хотя и ограниченное, но достаточно большое значение γ^* , модуль функции влияния уже не мал в той области на левой части графика, где еще мала плотность $f(x, \theta)$. ОРО там выглядит более предпочтительной — значение γ^* не слишком большое и модуль функции влияния ненамного больше модуля функции влияния ОМУ. Однако ОМУ явно

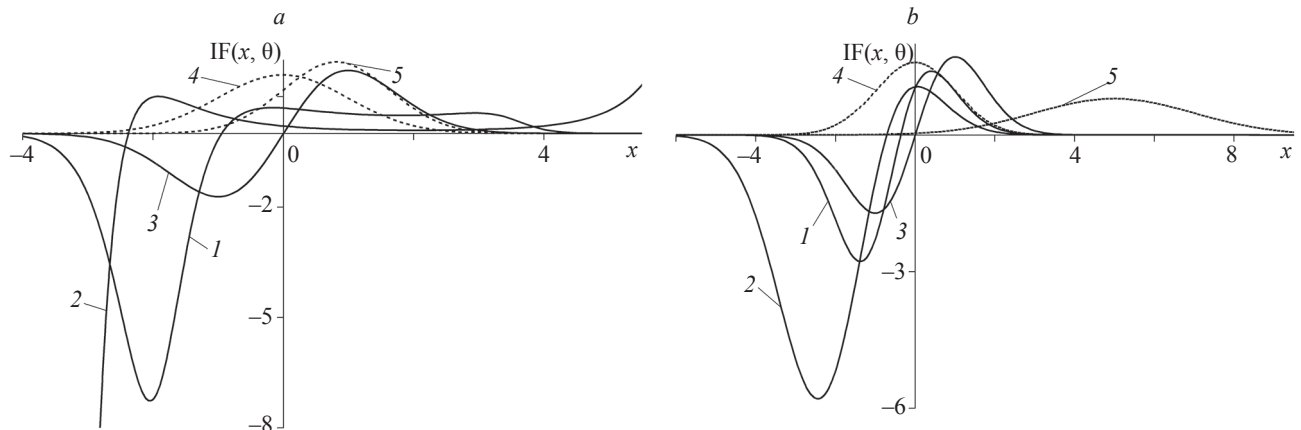


Рисунок. Графики некоторых функций влияния
Figure. Graphs of some influence functions

меньше защищена от искажения, сосредоточенного в правой части графика.

Малое значение функционалов $\text{stb}_s \psi_{\text{ОРО}}$ и $\text{stb} \psi_{\text{ОРО}}$ в случае, соответствующем рисунку, a , говорит о слишком большой потере качества с точки зрения критериев защищенности и робастности. В случае, соответствующем рисунку, b , потеря качества значительна, однако намного меньше. В то же время при $\sigma^2 \rightarrow \infty$ оценочные функции ψ_s и $\psi_{\text{ОРО}}$ стремятся к оценочной функции ОМУ, поэтому значение функционалов $\text{stb}_s \psi_{\text{ОРО}}$ и $\text{stb} \psi_{\text{ОРО}}$ будет стремиться к 1.

Наконец, отметим, что данные построения можно распространить на модель линейной регрессии и нейросетевые модели, включая смеси экспертов [25].

Заключение

В работе проведено теоретическое исследование компромиссного семейства оценок, которое обладает большей общностью, чем предложенное А.М. Шурыгиным семейство условно оптимальных оценок, и отличается способом параметризации. Получены следующие результаты:

- выведен ряд свойств семейства компромиссных оценок;
- показана единственность членов семейства компромиссных оценок;

Литература

1. Боровков А.А. Математическая статистика. Новосибирск: Наука. Сиб. предприятие РАН: Изд-во Ин-та математики, 1997. 771 с.
2. Шурыгин А.М. Прикладная стохастика: робастность, оценивание, прогноз. М.: Финансы и статистика, 2000. 223 с.
3. Huber P., Ronchetti E. *Robust Statistics*. 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. Использование принципа максимальной энтропии для конструирования робастных оценок при байесовском точечном засорении. Часть I // Прикладная математика и вопросы управления. 2024. № 1. С. 55–72. <https://doi.org/10.15593/2499-9873/2024.1.04>
6. Лисицин Д.В. Устойчивые методы оценивания параметров статистических моделей. Новосибирск: Изд-во НГТУ, 2013. 76 с.
7. Лисицин Д.В., Гаврилов К.В. Использование принципа максимальной энтропии для конструирования робастных оценок при байесовском точечном засорении. Часть II // Прикладная математика и вопросы управления. 2024. № 2. С. 18–33. <https://doi.org/10.15593/2499-9873/2024.2.02>
8. 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. <https://doi.org/10.17586/2226-1494-2024-24-2-267-275>
9. Лисицин Д.В. Устойчивое оценивание параметров модели по многомерным неоднородным неполным данным // Научный вестник НГТУ. 2013. Т. 50. № 1. С. 17–30.
10. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров модели при асимметричном засорении данных // Научный вестник НГТУ. 2008. Т. 30. № 1. С. 33–41.
11. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016. 352 p. <https://doi.org/10.1002/9781119264507>
12. Гаврилов К.В., Веретельникова Е.Л. Об одном способе выбора компромисса в семействе условно оптимальных оценок //

- введено понятие равнооптимальной оценки, показано, что ей соответствует седловая точка оптимизируемого функционала;
- показано, что равнооптимальная оценка является минимаксным решением относительно базовых критериев на множестве всех регулярных оценочных функций;
- построенная теория применена к задаче оценивания математического ожидания нормального распределения в условиях целенаправленного вредоносного воздействия на набор данных.

Среди членов компромиссного семейства особую роль играет равнооптимальная оценка. В силу своих свойств оптимальности она может представлять интерес в качестве решения, когда необходим некоторый априорный выбор одной оценки из семейства.

В ближайшие планы входит применение полученных результатов к задаче о минимаксной оценке при неизвестном размере окрестности байесовского точечного засорения в асимптотике сужающихся окрестностей, а также распространение построений из примера на задачу оценивания параметров линейной регрессии. В дальнейшем планируется распространение этих построений и на нейросетевые модели, включая смеси экспертов.

References

1. Borovkov A.A. *Mathematical Statistics*. Amsterdam: Gordon and Breach, 1998, 570 p.
2. Shurygin A. M. *Applied Stochastics: Robustness, Estimation, Prediction*. Moscow, Finansy i statistika Publ., 2000, 223 p. (in Russian)
3. Huber P., Ronchetti E. *Robust Statistics*. 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. The use of maximum entropy principle to construct robust estimators under point Bayesian contamination. Part I. *Applied Mathematics and Control Sciences*, 2024, no. 1, pp. 55–72. (in Russian). <https://doi.org/10.15593/2499-9873/2024.1.04>
6. Lisitsin D.V. *Robust Methods for Parameters Estimating of Statistical Models*. Novosibirsk, NSTU Publ., 2013, 76 p. (in Russian)
7. Lisitsin D.V., Gavrilov K.V. The use of maximum entropy principle to construct robust estimators under point Bayesian contamination. Part II. *Applied Mathematics and Control Sciences*, 2024, no. 2, pp. 18–33. (in Russian). <https://doi.org/10.15593/2499-9873/2024.2.02>
8. 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). <https://doi.org/10.17586/2226-1494-2024-24-2-267-275>
9. Lisitsin D.V. Robust estimation of model parameters in presence of multivariate nonhomogeneous incomplete data. *Science Bulletin of the Novosibirsk State Technical University*, 2013, vol. 50, no. 1, pp. 17–30. (in Russian)
10. Lisitsin D.V., Gavrilov K.V. On stable estimation of models parameters in presence of asymmetric data contamination. *Science Bulletin of the Novosibirsk State Technical University*, 2008, vol. 30, no. 1, pp. 33–41. (in Russian)
11. Shevlyakov G.L., Oja H. *Robust Correlation: Theory and Applications*. John Wiley & Sons, 2016, 352 p. <https://doi.org/10.1002/9781119264507>
12. Gavrilov K.V., Veretelnikova E.L. On one way of choice of compromise in the family of conditionally optimal estimates //

- Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 60–68. <https://doi.org/10.17223/19988605/67/7>
13. Coretto P., Hennig C. Robust improper maximum likelihood: tuning, computation, and a comparison with other methods for robust Gaussian clustering // *Journal of the American Statistical Association*. 2016. V. 111. N 516. P. 1648–1659. <https://doi.org/10.1080/01621459.2015.1100996>
 14. 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>
 15. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: статистическая обработка неоднородных совокупностей. М.: Статистика, 1980. 210 с.
 16. 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>
 17. Лисицин Д.В., Гаврилов К.В. Оценивание параметров распределения ограниченной случайной величины, робастное к нарушению границ // Научный вестник НГТУ. 2016. Т. 63. № 2. С. 70–89. <https://doi.org/10.17212/1814-1196-2016-2-70-89>
 18. 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>
 19. 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*. 1994. V. 3. P. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
 20. Ногин В.Д. Принятие решений в многокритериальной среде: количественный подход. М.: Физматлит, 2005. 176 с.
 21. Подиновский В.В. Идеи и методы теории важности критериев в многокритериальных задачах принятия решений. М.: Наука, 2019. 103 с.
 22. Li F., Lai L., Cui Sh. *Machine Learning Algorithms: Adversarial Robustness in Signal Processing*. Springer, 2022. 104 p. <https://doi.org/10.1007/978-3-031-16375-3>
 23. Grobthans M., Sawade C., Brückner M., Scheffer T. Bayesian games for adversarial regression problems // *Proc. of the 30th International Conference on Machine Learning. International Conference on Machine Learning*. 2013, V. 28. P. 55–63. <https://doi.org/10.5555/3042817.3042943>
 24. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983. 471 с.
 25. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2020. 1104 с.
 12. 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, pp. 60–68. (in Russian). <https://doi.org/10.17223/19988605/67/7>
 13. Coretto P., Hennig C. Robust improper maximum likelihood: tuning, computation, and a comparison with other methods for robust Gaussian clustering. *Journal of the American Statistical Association*, 2016, vol. 111, no. 516, pp. 1648–1659. <https://doi.org/10.1080/01621459.2015.1100996>
 14. 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>
 15. Smolyak S.A., Titarenko B.P. *Stable Estimation Methods: Statistical Processing of Heterogeneous Aggregates*. Moscow, Statistika Publ., 1980, 210 p. (in Russian)
 16. 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>
 17. Lisitsin D.V., Gavrilov K.V. Estimation of distribution parameters of a bounded random variable robust to bound disturbance. *Science Bulletin of the Novosibirsk State Technical University*, 2016, vol. 63, no. 2, pp. 70–89. (in Russian). <https://doi.org/10.17212/1814-1196-2016-2-70-89>
 18. 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>
 19. 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*, 1994, vol. 3, pp. 315–340. https://doi.org/10.1007/978-94-011-0854-6_15
 20. Nogin V.D. *Decision Making in Multicriteria Environment: a Quantitative Approach*. Moscow, Fizmatlit Publ., 2004, 176 p. (in Russian)
 21. Podinovskiy V.V. *Ideas and Methods of Criteria Importance Theory in Multicriteria Decision-Making Problems*. Moscow, Nauka Publ., 2019, 103 p. (in Russian)
 22. Li F., Lai L., Cui Sh. *Machine Learning Algorithms: Adversarial Robustness in Signal Processing*. Springer, 2022, 104 p. <https://doi.org/10.1007/978-3-031-16375-3>
 23. Grobthans M., Sawade C., Brückner M., Scheffer T. Bayesian games for adversarial regression problems. *Proc. of the 30th International Conference on Machine Learning. International Conference on Machine Learning*, 2013, vol. 28, pp. 55–63. <https://doi.org/10.5555/3042817.3042943>
 24. Aivazyaz S.A., Yenyukov I.S., Meshalkin, L.D. *Applied Statistics: Bases of Modelling and Initial Data Processing*. Moscow, Finansy i statistika, Publ., 1983, 471 p. (in Russian)
 25. Haykin S. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1998, 842 p.

Авторы

Лисицин Даниил Валерьевич — доктор технических наук, профессор, Новосибирский государственный технический университет, Новосибирск, 630073, Российская Федерация, [sc 57199594779](https://orcid.org/0009-0007-0723-6285), <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, [sc 58966384700](https://orcid.org/0009-0001-9252-9942), <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), <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, [sc 58966384700](https://orcid.org/0009-0001-9252-9942), <https://orcid.org/0009-0001-9252-9942>, aenigma77@mail.ru

Статья поступила в редакцию 25.02.2025
Одобрена после рецензирования 08.04.2025
Принята к печати 25.05.2025

Received 25.02.2025
Approved after reviewing 08.04.2025
Accepted 25.05.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»