

## ОТЗЫВ

официального оппонента доктора физико-математических наук, профессора Г.М. Кошкина на диссертационную работу Браништи Владислава Владимировича «**Методы и алгоритмы настройки проекционной оценки плотности вероятности случайного вектора в условиях малых выборок**», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 – **теоретические основы информатики**

### 1. Актуальность темы диссертации

Задача оценивания плотности вероятности случайного вектора возникает при разработке методов распознавания образов, фильтрации, распознавания и синтеза изображений. Имеющиеся в настоящее время методы оценивания плотности вероятности можно разделить на параметрические и непараметрические. Параметрические методы используются в случае, когда известна структура закона распределения с точностью до параметров, и задача сводится к построению статистических оценок этих параметров. Однако часто на практике возникают ситуации непараметрической неопределённости, когда структура закона распределения неизвестна. При использовании непараметрических методов представляет интерес исследование сходимости получаемых оценок к истинной функции плотности вероятности по заданной метрике и оценка скорости сходимости. В связи с этим возникает задача оптимальной настройки оценок функции плотности вероятности.

В большинстве работ, посвящённых этой проблеме, исследования преимущественно выполняются в предположении, когда объём выборки  $n \rightarrow \infty$ . Однако в анализе данных задачу оценивания плотности часто приходится решать при малых объёмах выборки  $n$ , например, при обработке биомедицинских данных и данных, касающихся производства и эксплуатации дорогостоящих технических систем. Исследования показали,

что результаты, полученные при  $n \rightarrow \infty$ , могут оказаться неоптимальными для малых  $n$ .

В данной диссертационной работе разрабатываются и исследуются проекционные оценки, которые, в отличие от других непараметрических оценок, не содержат в себе всей выборки и допускают компактное аналитическое выражение. Это оказывается более удобным при теоретическом анализе, в приложениях, а также повышает быстродействие алгоритмов классификации и восстановления зависимостей.

В связи с вышеизложенным, выполненные Браништи В.В. исследования при малых объёмах выборки  $n$ , что особенно важно для приложений, а также разработанные им соответствующие методики и алгоритмы для решения обширного класса конкретных практических задач, являются весьма актуальными.

## **2. Общая методология и методика исследования**

При исследовании проекционных оценок, а также при разработке эффективных алгоритмов при малых объёмах выборки использовались методы теории вероятностей, математической статистики, функционального анализа и теории меры, а также матричного анализа. При численных расчётах функционалов качества получаемых оценок применялся метод статистических испытаний.

## **3. Степень обоснованности и достоверности научных положений, выводов и рекомендаций**

Все полученные в диссертации научные результаты, выводы и рекомендации опираются на корректное применение методов математической статистики и функционального анализа, подтверждаются численными результатами.

## **4. Оценка проведенного исследования и полученных результатов**

Диссертация, объемом 125 страниц, включает введение, три главы основного содержания, заключение, список использованных источников из 160 наименований.

В первой главе представлены основные понятия и определения, необходимые теоретические выкладки. Дается обзор и сравнительный анализ основных методов оценивания функции плотности вероятности. Приводятся необходимые сведения из теории меры и функционального анализа.

Приводится постановка задачи оптимизации оценки функции плотности вероятности. Рассматриваются проекционные оценки плотности вероятности, приводятся формулы, которые позволяют настраивать коэффициенты и длину ряда этих оценок. Для оценки Розенблатта–Парзена даются различные методы настройки параметра размытости  $h$ .

**Вторая глава** посвящена исследованию проекционной оценки. В ней решается проблема применимости проекционной оценки для оценивания функций плотности вероятности с несуммируемым квадратом. Предлагается проекционная оценка плотности вероятности, в которой параметры настраиваются обобщенным методом моментов. Применение обобщенного метода моментов для оценивания коэффициентов проекционной оценки сводится к решению системы линейных уравнений. Было проведено сравнение (см. табл. 2.1) предложенных методов настройки проекционной оценки с традиционными методами, которое показало, что независимо от используемого базиса и восстанавливаемого распределения обобщенный метод моментов на малых выборках даёт близкие или лучшие результаты. Предложенный подход распространяется и на многомерный случай.

**В третьей главе** рассматриваются следующие задачи анализа данных: восстановление функции регрессии, классификация и оценивание количества информации. Предложены алгоритмы настройки коэффициентов и длины ряда проекционной оценки плотности вероятности, предназначенные для решения прикладных задач на малых выборках. Алгоритмы реализованы на языке Wolfram Language. Приводится сравнение разработанных алгоритмов с известными алгоритмами настройки на малых выборках.

## **5. Научная новизна полученных результатов**

В диссертационной работе Браништи В.В. получены следующие новые научные результаты, представляющие теоретический и практический интерес:

1. Впервые использовано весовое расширение пространства  $L_2$  при построении проекционной оценки для любых плотностей распределений.
2. Разработан новый метод настройки коэффициентов проекционной оценки плотности вероятности, являющийся обобщением метода моментов.
3. Предложен новый метод оценивания длины ряда проекционной оценки на базе метода моментов или его обобщения.
4. Разработаны алгоритмы настройки коэффициентов и длины ряда проекционной оценки плотности вероятности, которые

ориентированы на решение задач восстановления зависимостей, классификации и оценивания количества информации.

## **6. Значение выводов и рекомендаций, полученных в диссертации для науки и практики**

Представленные в диссертационной работе Браништи В.В. исследования обладают научной новизной и достоверностью, все полученные выводы научно обоснованы. Разработанные методики и алгоритмы имеют теоретическую ценность и могут быть использованы при решении задач восстановления зависимостей, классификации и оценивания количества информации.

## **7. Замечания по диссертационной работе**

1. Зачем в формуле (1.17)  $v$  нормируется на 1 модуль ядра? Обычно просто ядро нормируется на 1.

2. В работе имеются погрешности и неточности. Так, на с. 17  $k$ -мерная функция распределения обозначается как  $n$ -мерная, на с. 23 не определяется  $D\{\varphi_j(\xi)\}$ .

3. На с. 72 сначала говорится о максимизации оценки функционала (1.6), хотя вроде необходимо его минимизировать. Далее, функционал (2.21) то требуется максимизировать, то минимизировать, а в формуле (2.22) ищется минимум (см. также Таблицу 2.6). В то же время на с. 88 и с. 95 говорится только о максимизации (2.36) и (2.37), но на с. 103 речь идет снова о минимизации функционала. Думаю, что в указанных случаях должна идти речь о минимизации функционалов.

4. В ряде случаев алгоритмы восстановления функции регрессии на основе простой оценки Розенблатта–Парзена успешнее справляются с задачей по сравнению с алгоритмами на основе проекционной оценки (см., например, с. 98). Возникает вопрос о сравнении проекционных оценок с усложненными оценками Розенблатта–Парзена, обеспечивающими улучшенную скорость сходимости их среднеквадратических ошибок к нулю.

Приведенные недостатки вполне устранимы, замечания носят, в основном, характер пожеланий, поэтому все это не снижает существенным образом ценности проделанной работы, что позволяет дать общую положительную оценку результатов исследований.

## Заключение

Диссертационная работа Браништи В.В. обладает внутренним единством, содержит новые научные результаты и положения, выдвигаемые для публичной защиты, что свидетельствует о личном вкладе автора диссертации в науку. Материалы диссертации достаточно полно изложены в публикациях автора. Автореферат соответствует содержанию диссертации.

На основании вышеизложенного считаю, что диссертационная работа Браништи В.В. «Методы и алгоритмы настройки проекционной оценки плотности вероятности случайного вектора в условиях малых выборок» соответствует требованиям п. 9 "Положения о присуждении учёных степеней" постановления Правительства Российской Федерации от 24.09.2013 г. № 842, а ее автор, Браништи Владислав Владимирович, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.17 – теоретические основы информатики.

Официальный оппонент,  
профессор кафедры системного анализа и математического моделирования федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Томский государственный университет», доктор физико-математических наук (05.13.18 – Математическое моделирование, численные методы и комплексы программ), профессор



Геннадий Михайлович Кошкин

4.06.2019 г.

Сведения об организации:

634050, г. Томск, пр. Ленина, 36; Тел.: +7382 252-95-82

E-mail: [rector@tsu.ru](mailto:rector@tsu.ru); <http://www.tsu.ru>

Кафедра системного анализа и математического моделирования образована в результате реорганизации кафедры теоретической кибернетики.

