

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Зайцева Алексея Алексеевича

«Методы построения регрессионных моделей разнородных источников данных для индустриальной инженерии», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Диссертационная работа Зайцева Алексея Алексеевича посвящена вопросам построения регрессионных моделей по данным, в которых зависимая переменная может быть получена из разнородных источников, отличающихся по точности и стоимости получения зависимой переменной. С учетом подходов, обычно используемых в индустриальной инженерии для построения регрессионных моделей на основе данных, полученных из разных источников, диссидентом для исследования был выбран метод регрессии на основе гауссовских процессов. Рассматриваемые вопросы активно изучаются в нашей стране и за рубежом, однако у известных результатов есть ряд ограничений как теоретического, так и практического свойства. Так, большинство методов для выбора планов экспериментов для источников данных разной точности предполагают наличие достаточно точной регрессионной модели – но инженер обычно вынужден выбирать план экспериментов в условиях ограниченной информации об исследуемой зависимости. Кроме того, современные работы ограничивались рассмотрением минимаксной ошибки интерполяции только для одномерного пространства входных переменных (регрессоров). Таким образом, актуальность диссертационной работы не вызывает сомнений.

Научная новизна работы состоит в том, что в ней впервые исследованы ошибки интерполяции в минимаксном случае для моделирования по выборкам данных, в том числе и разной точности; предложен метод выбора оптимального соотношения между размерами выборок разнородных данных в условиях общего бюджетного ограничения; также разработаны более вычислительно эффективные по сравнению с ранее предложенными в литературе методы построения модели по разнородным данным большого размера. Разработанные алгоритмы вошли в программный комплекс, используемый для решения прикладных задач инженерного проектирования, который активно используется в том числе в компании Airbus.

Полученные в диссертации результаты обоснованы и достоверны. Теоретические результаты подтверждаются строгими доказательствами с использованием методов

спектрального анализа и математической статистики, а низкая трудоемкость предложенных численных методов обосновывается как с помощью точных оценок их вычислительной сложности, так и при помощи вычислительных экспериментов.

Диссертация логически структурирована и охватывает важные вопросы теоретического и практического характера. Она состоит из введения, пяти глав, заключения и списка литературы. Во введении содержится краткий обзор имеющихся в научной литературе результатов по теме диссертации и обосновывается актуальность работы.

В первой главе дается постановка задачи построения регрессионной модели по выборкам разнородных данных, приведены примеры реальных задач индустриальной инженерии, требующих дальнейшей разработки методов построения таких регрессионных моделей, приведено обоснование проведения исследования именно модели регрессии на основе гауссовских процессов, дана формальная математическая постановка. Вторая глава описывает теоретические свойства процедур построения регрессионных модели на основе гауссовских процессов и обосновывает с теоретической точки зрения и посредством вычислительных экспериментов возможность и полезность байесовского подхода к оценке параметров таких регрессионных моделей. Описанные результаты также проиллюстрированы вычислительными экспериментами.

В третьей главе приведены описания алгоритмов для моделирования по выборкам разнородных данных, основанные на низкоранговой аппроксимации ковариационной матрицы и на использовании черного ящика, способного в режиме реального времени предоставлять данные низкой точности. Показано, что предложенные алгоритмы дают достаточно точные результаты, получена вычислительная сложность. Глава также содержит описание комплекса программ, в котором реализованы предложенные в работе алгоритмы.

В четвертой главе приведены основные теоретические результаты исследования. Рассматриваются вопросы оценки ошибки интерполяции для однородных и разнородных данных. Удаётся получить ошибку интерполяции для многомерного пространства входных переменных для известной ковариационной функции случайного процесса и в минимаксном случае для класса достаточно гладких ковариационных функций (ранее такой результат был получен лишь в одномерном случае и для однородных данных в работе Голубева и Крымовой). Результаты обобщены на случай разнородных данных. Получено явное соотношение между размерами выборок разнородных данных, доставляющее минимум минимаксной ошибки интерполяции в условиях заданного общего ресурсного ограничения. Приведенные в данной главе численные эксперименты показывают, что несмотря на

предположения о бесконечном размере выборки и сеточной структуре плана эксперимента, полученные результаты могут использоваться для выбора плана эксперимента в реальных задачах. Основные результаты главы заключаются в получении минимаксной оценки интерполяции для многомерного пространства входных переменных и их использовании для получения качественных результатов для моделирования по выборкам разнородных данных.

Пятая глава содержит примеры использования разработанных в диссертации подходов для решения реальных задач и демонстрации свойств алгоритмов на искусственных примерах. Рассматривается задача построения регрессионных зависимостей прочности вращающегося диска от его геометрии, аэродинамического качества крыла самолета от его геометрии, а также режима полета и максимальной нагрузки на С-образный пресс от геометрии пресса. Также на ряде задач демонстрируется, что предложенный в исследовании метод выбора соотношения между размерами выборок разнородных данных позволяет в реальных задачах получать план эксперимента, при использовании которого для построения регрессионных моделей удается превзойти результаты, полученные с помощью эвристических подходов, применявшихся для выбора такого соотношения ранее.

Диссертационная работа выполнена на высоком уровне, написана ясным языком, хотя в тексте ее содержится некоторое количество опечаток. Автореферат правильно отражает содержание диссертационной работы.

По диссертации можно сделать следующие замечания:

1. В разделе 4.3 автором рассмотрен случай решетки с одинаковым шагом по каждой размерности. Не обсуждается вопрос, насколько принципиальным является данное ограничение при получении теоретических результатов.
2. В главе 4 не изучен вопрос об оптимальном плане эксперимента в случае наличия источников данных разной точности и известной ковариационной функции.
3. Часть рассмотренных прикладных задач не принадлежит к области индустриальной инженерии.

Указанные недостатки не влияют на общую положительную оценку диссертации.

Считаю, что диссертационная работа Зайцева А.А. на соискание ученой степени кандидата физико-математических наук является научно-квалификационной работой, в которой содержится решение важной прикладной задачи моделирования и которая соответствует всем требованиям ВАК, а ее автор – Зайцев Алексей Алексеевич – достоин

присуждения ему ученой степени кандидата физико-математических наук по специальности
05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Официальный оппонент,

ведущий научный сотрудник ИПУ РАН,

доктор физико-математических наук,

профессор

А.В. Назин

20 апреля 2017 г.

Адрес: Россия, 117997, Москва, ул. Профсоюзная, д.65

Телефоны: +7 495 334-89-10, +7 495 334-76-41

E-mail: nazine@ipu.ru

