

На правах рукописи

Голодов Валентин Александрович

**Интервальный подход к регуляризации неточно заданных
систем линейных уравнений**

05.13.17 – теоретические основы информатики

**Автореферат диссертации на соискание учёной степени кандидата
физико-математических наук**

Москва – 2014

Работа выполнена на кафедре экономико-математических методов и статистики Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Южно-Уральский государственный университет» (национальный исследовательский университет).

Научный руководитель – доктор физико-математических наук,
профессор Панюков Анатолий Васильевич.

Официальные оппоненты: доктор физико-математических наук,
Шарый Сергей Петрович,
Институт вычислительных технологий
Сибирского отделения РАН
профессор каф. мат. моделирования
Новосибирского госуниверситета;

кандидат физико-математических наук,
с.н.с. Иванко Евгений Евгеньевич,
Институт математики и механики
Уральского отделения РАН.

Ведущая организация – Федеральное государственное бюджетное
образовательное учреждение
высшего профессионального образования
«Пермский государственный
национальный исследовательский университет».

Защита диссертации состоится 13 ноября 2014 г., в 16 часов, на заседании диссертационного совета Д 002.017.02 при Федеральном государственном бюджетном учреждении науки Вычислительный центр им. А. А. Дородницына Российской академии наук, расположенном по адресу: 119333, г. Москва, ул. Вавилова, д. 40.

С диссертацией можно ознакомиться в библиотеке ВЦ РАН и на сайте <http://www.ccas.ru/>.

Автореферат разослан «___» _____ 2014 г.

Учёный секретарь
диссертационного совета
доктор физико-математических наук, профессор

Рязанов В.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Система линейных алгебраических уравнений используется при математическом моделировании объектов различной природы, в частности, физических, химических и социально-экономических процессов, процессов управления. В ходе качественного анализа и/или использования математических моделей необходимо решать различные системы линейных алгебраических уравнений.

В ситуации, когда система $Ax = b$ возникает как математическая модель реального явления, точные значения коэффициентов обычно неизвестны. Известны лишь приближённые, например, полученные в результате измерений, значения коэффициентов a_{ij} и погрешности измерения δ_{ij} .

В данной работе предлагается считать моделью **неточно заданной системы** линейных алгебраических уравнений интервальную систему линейных алгебраических уравнений, т.е. СЛАУ с интервальными коэффициентами. Таким образом, интервальная неопределённость $a_{ij} \in [a_{ij} - \delta_{ij}, a_{ij} + \delta_{ij}]$ является следствием природы самой математической модели. В том случае, когда система линейных алгебраических уравнений возникает как результат промежуточных вычислений, например, в ходе качественного и/или количественного исследования математических моделей, интервальная неопределённость является следствием аналитических оценок приближённых величин, выступающих в роли коэффициентов. Учет интервальной неопределенности в исходных данных позволяет существенно повысить качество моделирования и строить модели более адекватные исследуемым явлениям. В качестве примеров в диссертации приведены: интервальная модель межотраслевого баланса (модель Леонтьева), задача анализа смеси веществ по спектральным характеристикам компонентов методом Фирордта, также примерами могут служить: задача распознавания образов (анализ зашумленных изображений), поиск значений значений собственных функций оператора, задачи теории управления и многие другие.

На практике необходимо решать *плохо обусловленные* и даже *вырожденные* системы линейных алгебраических уравнений, неустойчивых к возмущениям и погрешностям входных данных. Для решения сильно вырожденных или несовместных СЛАУ применяют различные методы решения некорректных задач, например, псевдорешение по М.М. Лаврентьеву, регуляризация А.Н. Тихонова, матричной коррекции (И.И. Еремин, В.А. Горелик, В.И. Ерохин и др.) В данной работе предлагается подход к регуляризации исходной СЛАУ за счет её погружения в интервальную СЛАУ, данный подход назван **интервальной регуляризацией**.

В настоящее время широко распространены заблуждения, порождающие ошибки в расчетах: (1) распространение свойства ассоциативности операций сложения и умножения в поле действительных чисел на конечное множество машинных «действительных» чисел; (2) распространение свойства непрерывной зависимости от параметров решения системы, полученной после «эквивалентных» преобразований, на исходную систему. Это присутствует в популярных коммерческих пакетах «MatLab», «MathCad» и т.п., а также свободно распространяемом пакете «SciLab». Использование в параллельных вычислениях разного числа процессоров во многих случаях даёт существенно различающиеся результаты, демонстрируя необходимость применения надежных, а в некоторых случаях доказательных вычислений (см. работы К.И. Бабенко, Э.А. Бибердорф, В.А. Гаранжа, С.К. Годунов, А.И. Голиков, Ю.Г. Евтушенко, А.Н. Малышев, П.С. Панков, Н.И. Попова, Ju. Hall, J.H. Reif). Вопросам надежной реализации предлагаемого подхода посвящена отдельная глава.

Потенциал имеющихся пакетов, поддерживающих символьные вычисления, не позволяет решать реальные проблемы математического и имитационного моделирования. Возможность обеспечения вычислений с произвольной точностью в программах пользователя даёт библиотека «GMP». Однако, ее использование требует от пользователя разработки собственного интерфейса для организации распределённых и параллельных вычислений. Возможность использования в параллельных вычислениях предоставляет своим объектам собственная разработка – библиотека «Exact Computation 2.0».

Интервальный анализ и его приложения являются активно развивающейся областью науки. Один из наиболее полных обзоров многообразия разработанных методов и областей применения дан в книге С.П. Шарого. Свой весомый вклад в исследование систем с интервальной неопределённостью в исходных данных внесли отечественные и зарубежные ученые: А.В. Лакеев, С.П. Шарый, И.А. Шарая, О. Beaumont, V. Kreinovich, B. Philippe, J. Rohn и многие другие. Системы с интервальной неопределённостью возникают в различных теоретических и прикладных областях науки, математические модели в экономике, например, модель межотраслевого баланса В.В. Леонтьева – изначально имеют интервальный характер, проверка физических моделей на реальном эксперименте требует обработки данных от приборов-датчиков результаты измерений которых лежат в интервалах, определяемых уровнями квантования прибора.

Современный уровень развития и распространения высокопроизводительных вычислительных систем кластерного типа актуализирует использование массового параллелизма уже на стадии разработки

алгоритмов. Современные ускорители достигли и превзошли в отдельных компонентах мощность традиционных процессоров, но программирование такого рода устройств требует новых параллельных методов и специальных подходов к разработке эффективных программных комплексов.

Целями диссертационной работы являются:

- 1) Предложение нового подхода к регуляризации СЛАУ за счет погружения в ИСЛАУ и перехода к поиску точки из допускового множества решений ИСЛАУ.
- 2) Развитие концепции псевдорешения ИСЛАУ ее применение в математическом моделировании.
- 3) Создание и тестирование масштабируемого алгоритма для вычисления псевдорешения интервальной системы линейных алгебраических уравнений в современной гетерогенной вычислительной среде.
- 4) Разработка программного комплекса для вычисления псевдорешения интервальной системы в гетерогенной вычислительной системе.

В соответствии с поставленной целью в диссертационной работе **решаются следующие задачи:**

- 1) Разработка алгоритма численного нахождения псевдорешения интервальных систем линейных алгебраических уравнений.
- 2) Адаптация алгоритма решения поставленной задачи к методам программирования гетерогенных вычислительных сред.
- 3) Разработка и тестирование программного комплекса поддержки безошибочных вычислений для параллельных алгоритмов в гетерогенных вычислительных системах.
- 4) Реализация программного комплекса для эффективного численного нахождения псевдорешения интервальных систем линейных алгебраических уравнений в гетерогенной вычислительной системе за счет масштабируемого параллельного алгоритма.

Основными **объектами исследования** являются: интервальная система линейных алгебраических уравнений, псевдорешение интервальной системы линейных алгебраических уравнений, методика вычисления псевдорешения интервальной системы, методы реализации дробно-рациональных вычислений без округления в распределённых вычислительных системах с графическими ускорителями, параллельный метод вычисления псевдорешения системы линейных алгебраических уравнений.

Методами исследования являются: математический анализ, алгебраические методы, анализ алгоритмов, численные методы, программирование, вычислительный эксперимент.

Научная новизна заключается: (1) в создании подхода к регуляризации неточно заданных систем за счет погружения в интервальную СЛАУ,

в создании нового подхода к коррекции правой части системы в задаче о допусках для интервальной системы линейных алгебраических уравнений, учет интервальной неопределенности в исходных данных повышает адекватность математического моделирования, (2) в разработке и реализации параллельного программного обеспечения для эффективного решения поставленной задачи в распределённых вычислительных системах с графическими ускорителями, (3) в разработке и реализации программного комплекса для безошибочных дробно-рациональных вычислений в гетерогенных вычислительных системах, реализации программного комплекса для численного решения задач математического моделирования.

Теоретическая значимость. Предлагаемый подход сводит задачу решения неточно заданной системы линейных уравнений к задаче поиска точки из допускового множества решений интервальной системы, подобная техника применяется впервые, метод решения нахождения точки из допускового множества ИСЛАУ продолжает ряд исследований М.М. Лаврентьева, Н.А. Хлебалина, J. Rohn, O. Beaumont, B. Philippe позволяет, вместе с определением совместности задачи о допусках для интервальной системы, решать задачу коррекции правой части системы. Исследуемый метод решения продолжает работы по применению техники линейного программирования в задачах интервального анализа. Коррекция задачи о допусках осуществляется за полиномиальное время. Расширение допустимых интервалов правой части системы позволяет предложить решение том в случае, когда допусковое множество интервальной системы пусто.

Комплекс программ для безошибочных дробно-рациональных вычислений продолжает и существенно развивает работы А.Н. Румянцева, библиотеки GMP на техническую базу современных кластерных решений с графическими ускорителями. Разработаны параллельные масштабируемые алгоритмы для эффективного использования GPGPU ускорителей.

Практическая значимость. Предложенный метод решения неточно заданных систем линейных алгебраических уравнений может быть использован в исследовании с помощью средств вычислительной техники, информационных процессов, задачах обработки зашумленных данных, относящихся к области исследования специальности 05.13.17 – теоретические основы информатики. Программное обеспечение позволяет получить решения неточно заданных систем в рамках предлагаемого подхода, подход показал свою эффективность в прикладных задачах из различных областей. Реализованные алгоритмы адаптированы для использования в гетерогенных вычислительных кластерах с графическими процессорами и эффективно масштабируются. Программный комплекс может быть

полезен для исследователей, которым необходимо использовать методы безошибочных вычислений. Реализованная функциональность позволяет решать некорректные задачи, неустойчивые к погрешностям входных данных, а также решать плохо обусловленные СЛАУ, востребованные в различных математических моделях, с абсолютной точностью.

Работа выполнена в рамках соглашения №14.В37.21.0395 с Министерством образования и науки Российской Федерации от 06 августа 2012 года. Правообладателем разработанных в рамках данной модели комплексов программ является ЮУрГУ.

Аппробация работы. Все результаты диссертационной работы, разработанные методы, алгоритмы и результаты вычислительных экспериментов докладывались и получили одобрение на следующих международных, всероссийских и внутривузовских конференциях: Третья научная конференция аспирантов и докторантов ЮУрГУ. (г. Челябинск, 2011), Алгоритмический анализ неустойчивых задач. Международная конференция, посвященная памяти В. К. Иванова (г. Екатеринбург, 2011), Четвертая научная конференция аспирантов и докторантов ЮУрГУ. (г. Челябинск, 2012), 15th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Verified Numeric (г. Новосибирск, 2012), Параллельные вычислительные технологии (ПаВТ'2012) (г. Новосибирск, 2012), Параллельные вычисления и задачи управления (РАСО'2012). Шестая международная конференция (г. Москва, 2012), Информационные технологии и системы. (оз. Банное, респ. Башкортостан, 2013), Пятая научная конференция аспирантов и докторантов ЮУрГУ. (г. Челябинск, 2013), GPU Technology Conference-2013 (San Jose, California, 2013).

Публикации. По теме диссертационного исследования опубликовано 15 работ (6 статей, 2 постера, тезисов конференции – 4, 3 свидетельства о государственной регистрации программы для ЭВМ). В их числе 2 статьи из перечня ведущих российских рецензируемых изданий, а также 1 статья в рецензируемом международном научном журнале.

КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **введении** обосновывается актуальность темы диссертации, выделяются и формулируются цели и задачи исследования, описывается структурно-логическая схема диссертационной работы.

В первой главе детально рассматривается математическая модель неточно заданной системы — интервальная система линейных алгебраических уравнений, области ее применения. Учет интервальной неопределенности в исходных данных позволяет существенно повысить

качество моделирования и строить модели более адекватные исследуемым явлениям. В качестве примеров приведены: интервальная модель межотраслевого баланса (модель Леонтьева), задача анализа смеси веществ по спектральным характеристикам компонентов методом Фирордта.

Даётся обзор известных подходов к решению различных постановок интервальных задач, приводятся данные о свойствах решений, а также производится сопоставление решений в рамках приводимых методов. Обозначены проблемы существующих программных средств для решения задачи о допусках для ИСЛАУ. Приводятся примеры математических подходов к решению некорректных задач, псевдорешение по М.М. Лаврентьеву, регуляризация А.Н. Тихонова, излагаются предпосылки предлагаемого подхода к регуляризации путем погружения системы в интервальную СЛАУ. Поскольку погрешность вычислений может существенно искажать результат, то следует уделять внимание влиянию погрешностей во входных данных и промежуточных вычислениях на численное решение. Дается краткий обзор применения доказательных и надежных вычислений при решении различных задач.

В данной работе рассматривается интервальная система линейных алгебраических уравнений

$$\begin{cases} \mathbf{a}_{11}x_1 + \mathbf{a}_{12}x_2 + \dots + \mathbf{a}_{1n}x_n = \mathbf{b}_1 \\ \mathbf{a}_{21}x_1 + \mathbf{a}_{22}x_2 + \dots + \mathbf{a}_{2n}x_n = \mathbf{b}_2 \\ \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ \mathbf{a}_{m1}x_1 + \mathbf{a}_{m2}x_2 + \dots + \mathbf{a}_{mn}x_n = \mathbf{b}_m \end{cases} \quad (1)$$

с интервалами $\mathbf{a}_{ij} = [a_{ij}, \bar{a}_{ij}]$ и $\mathbf{b}_i = [b_i, \bar{b}_i]$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, кратко,

$$\mathbf{A}x = \mathbf{b}. \quad (2)$$

Существуют различные постановки связанные с интервальными системами уравнений, с многими из них можно познакомиться в книге С.П. Шарого «Конечномерный интервальный анализ», существуют и отдельные постановки, такие как «Универсальное» решение ИСЛАУ, введенное и исследованное Л.Т. Ащепковым и Д.В. Давыдовым, постановка изложена в одноименной книге данных авторов.

Задачей о допусках для ИСЛАУ (1) называется поиск решений принадлежащих допусковому множеству решений

$$\Xi_{tol}(\mathbf{A}, \mathbf{b}) = \{x \in \mathbb{R}^n \mid (\forall A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b)\}, \quad (3)$$

или в эквивалентной форме

$$\Xi_{tol}(\mathbf{A}, \mathbf{b}) = \bigcap_{A \in \mathbf{A}} \{x \in \mathbb{R}^n \mid (\exists b \in \mathbf{b})(Ax = b)\}, \quad (4)$$

В представлении (3) квантор всеобщности $\forall A \in \mathbf{A}$ эквивалентен теоретико-множественной операции пересечения. Поэтому из (4) следует, что допусковое множество наименее чувствительно среди всех других множеств решений ИСЛАУ к изменению интервальной матрицы системы $\mathbf{A}x = \mathbf{b}$, так как (4) не шире чем множество решений $\{x \in \mathbb{R}^n \mid (\exists b \in \mathbf{b})(Ax = b)\}$ самой «лучшей» и самой робастной(устойчивой) матрицы $A \in \mathbf{A}$. Хотя некоторые матрицы $A \in \mathbf{A}$ могут быть плохо обусловленными или вырожденными влияние «хороших» матриц обеспечивает ограниченность и хорошие оценки для Ξ_{tol} . Следовательно мы можем использовать допусковое множество как инструмент регуляризации для неустойчивых и плохо обусловленных задач, когда необходимо сгладить эффект изменения решения задачи, вызванный неточностями или возмущениями данных. Подобный подход можно назвать «интервальной регуляризацией».

Известные подходы к решению задачи о допусках имеют свои особенности, преимущества и недостатки.

Формально-алгебраический подход, предложенный и активно развиваемый Новосибирской научной школой, в частности, С.П. Шарым и его коллегами, требует реализации интервальной арифметики, например, можно воспользоваться `JInterval`.

Для ИСЛАУ с точечной абсолютно неособенной матрицей системы формальной решение единственно для любой интервальной правой части. Что касается интервальных систем уравнений с существенно интервальными матрицами, то для них единственность формальных решений является в настоящее время сравнительно малоисследованной. В рамках данной работы формальные решения не рассматривались.

Метод «распознающего» функционала, предложенный и разработанный С.П. Шарым, И.А. Шарой, основан на работе со специальным функционалом $Tol(\mathbf{A}, \mathbf{b})$ и его обобщениями. Данный метод позволяет проводить полное исследование задачи о допусках, включая исчерпывающие возможности по коррекции правой части системы и коррекции матрицы системы, естественно, за счет дополнительных вычислений.

Для данного метода в сложных случаях, например, когда Ξ_{tol} мало или содержит единственную точку, желательно применять точные вычисления или использовать другие процедуры доказательных вычислений, поскольку приближенные методы вычисления могут давать неверный результат.

Методы, сводящие различные аспекты задачи о допусках к задаче линейного программирования, разработанные на данный момент, позволяют решать задачи: определения пустоты Ξ_{tol} (работы J. Rohn), приближенного оценивания Ξ_{tol} (работы О. Beaumont, В. Philippe, Н.А. Хлебалина), но

не дают информации о коррекции данных системы в случае пустоты Ξ_{tol} . Точность проводимых вычислений также может оказаться критичной.

В рамках рассматриваемой задачи актуальным является поиск подхода, позволяющего получить точку допускового множества и обеспечить коррекцию правой части системы, в случае пустоты множества $\Xi_{tol}(\mathbf{A}, \mathbf{B})$. Уделить внимание прежде всего коррекции правой части системы имеет смысл, поскольку именно она содержит вектор наших «ожиданий» в ряде содержательных математических моделей.

Для этого, в продолжении с одной стороны работ Ж. Рohn и О. Beaumont, В. Philippe по исследованию ИСЛАУ и, с другой стороны, идеи псевдорешения, предложенной М.М. Лаврентьевым для задач математической физики, в данной работе предлагается подход к решению ИСЛАУ, позволяющий находить точку из допускового множества и корректировать правую часть системы в случае пустоты Ξ_{tol} .

Во второй главе описан предлагаемый численный метод решения ИСЛАУ, введено определение «псевдорешения ИСЛАУ», доказано его существование для любой ИСЛАУ. Приведен конструктивный способ поиска «псевдорешения ИСЛАУ», показаны его свойства. Показана необходимость точных дробно-рациональных вычислений для реализации предлагаемого численного метода нахождения псевдорешения, а также параллельных технологий MPI и CUDA для его эффективного вычисления.

В практических задачах система (1) часто даёт плохо обусловленную или вообще несовместную задачу о допусках. В этом случае разумным представляется введение понятия «псевдорешения» для ИСЛАУ.

Определение 1 (1.2.1). *Псевдорешениями интервальной системы линейных уравнений $\mathbf{Ax} = \mathbf{b}$ назовем точки допускового множества системы $\mathbf{Ax} = \mathbf{b}(z)$ с расширенной правой частью $\mathbf{b}(z) = [\underline{\mathbf{b}} - z\mathbf{p}, \bar{\mathbf{b}} + z\mathbf{q}]$, где $\mathbf{p}, \mathbf{q}$ – константные положительные векторы, определяющие характер расширения исходя из содержательного смысла задачи, $z \geq 0$ – параметр, отвечающий за величину расширения.*

Существование объекта из определения [1] доказывает следующая теорема,

Теорема 1 (2.3.1). *Для любой системы интервальных уравнений $\mathbf{Ax} = \mathbf{b}$ при всех $z \geq \max\{0, \max_{i=1,\dots,m} \underline{b}_i/p_i, -\min_{i=1,\dots,m} \bar{b}_i/q_i\}$ множество $\Xi_{tol}(\mathbf{A}, \mathbf{b}(z)) \neq \emptyset$.*

На практике ценность представляет *наилучшее* возможное псевдорешение x^* , отвечающее значению $z^* = \inf_{\Xi_{tol}(\mathbf{A}, \mathbf{b}(z)) \neq \emptyset} z$.

Способ нахождения псевдорешения системы уравнений, а также минимального значения z^* параметра z – коэффициента, отвечающего за расширение интервалов в правой части $\mathbf{Ax} = \mathbf{b}$ даёт следующая теорема.

Теорема 2 (2.5.1). *Существует решение $x^{+*}, x^{-*} \in \mathbb{R}^n, z^* \in \mathbb{R}$ задачи линейного программирования*

$$z \rightarrow \min_{x^+, x^-, z}, \quad (5)$$

$$\sum_{j=1}^n (\bar{a}_{ij}x_j^+ - \underline{a}_{ij}x_j^-) \geq \underline{b}_i - zp_i, \quad i = 1, 2, \dots, m, \quad (6)$$

$$\sum_{j=1}^n (\bar{a}_{ij}x_j^+ - \underline{a}_{ij}x_j^-) \leq \bar{b}_i + zq_i \quad i = 1, 2, \dots, m, \quad (7)$$

$$x_j^+, x_j^-, z \geq 0, \quad j = 1, 2, \dots, n, \quad (8)$$

при этом $x^* = x^{+*} - x^{-*}$ является псевдорешением ИСЛАУ $Ax = b$.

Для частного случая интервальной системы – невырожденной точечной системы $Ax = b$, ее решение в обычном смысле и наилучшее возможное псевдорешение совпадают.

Утверждение 1 (2.5.1). *Если точечная система $Ax = b$ невырождена, то ее решение и псевдорешение интервальной системы $Ax = b$, где $A = [A, A], b = [b, b]$ совпадают.*

Полученная ЗЛП (5)-(7) потенциально имеет высокую степень вырожденности задачи, что будет приводить, при использовании приближенных вычислений, к закликиванию симплекс-метода и погрешностям при применении других методов решения задачи линейного программирования. Выходом является использование рациональных вычислений без округления, подобные разработки были предложены в совместной работе А.В. Панюкова, М.И. Германенко, В.В. Горбика библиотеке классов – «Exact Computation» 2009 года. В работах А.В. Панюкова и В.В. Горбиком доказывается что количество требуемых на каждой итерации бит памяти полиномиально зависит от числа бит, достаточных для представления элемента матрицы исходных данных. Симплекс допускает эффективное распараллеливание, при этом эффективность (т.е. отношение ускорения к числу процессоров) в асимптотике близка к 100 %,

В третьей главе рассматриваются проблемы использования аппаратных приближенных типов данных, которые реализованы в вычислительных процессорах и используются в программных кодах без учета погрешностей. Приводится обзор библиотек позволяющих производить точные дробно-рациональные вычисления. Описывается разработанный программный комплекс безошибочных дробно-рациональных вычислений на современных процессорах и графических ускорителях. Предлагаются методы адаптации разработанных компонентов библиотеки для эффективного проведения параллельных расчетов методами MPI. Дается формат входных и выходных данных, схема работы с разработанным программным комплексом для вычисления псевдорешения интервальной системы.

Применение безошибочной дробно-рациональной арифметики требует больших объёмов вычислений. Разработка параллельных алгоритмов для

проведения базовых арифметических операций в параллельном режиме с использованием массового параллелизма графических ускорителей (GPU) позволяет сократить время решение задач за счет эффективного использования ресурса параллелизма современных вычислительных систем.

Эту задачу помогает решить программная модель CUDA (Compute Unified Device Architecture) – программная модель, включающая описание вычислительного параллелизма и иерархической структуры памяти непосредственно в язык программирования. В целом работа нитей на GPU соответствует принципу SIMD (Single Instruction Multiple Data), однако есть существенное различие связанное с особенностями синхронизации в рамках этой архитектуры SIMT (Single Instruction Multiple Threads) и исполнению нитей блоками фиксированного размера.

На время выполнения ядра (кода на GPU) каждый блок получает в распоряжение часть быстрой разделяемой памяти, которую могут совместно использовать все нити блока. Поскольку не все нити блока выполняются физически одновременно, то для их взаимного согласования необходим механизм синхронизации. Для этой цели в CUDA предусмотрен вызов `__syncthreads()`, который блокирует дальнейшее исполнение ядра до тех пор, пока все нити блока не войдут в эту функцию.

Библиотека поддержки точных вычислений в гетерогенной среде должна функционировать в любой конфигурации, от минимальной, с одним процессором на вычислительном узле, до максимальной, когда узел системы состоит из нескольких центральных процессоров и 4–6 специализированных графических ускорителей. Для этого была разработана гибкая масштабируемая архитектура библиотеки, которая обеспечивает возможность проверки конфигурации системы в момент запуска приложения. Применение объектно-ориентированного подхода при создании программных комплексов позволило использовать все преимущества современного ООП. Современные классы `overlong` и `rational` реализованы в объектно-ориентированной парадигме на языке C++, классы оформлены в виде библиотеки «Exact Computation 2.0» в том числе с возможностью динамической загрузки.

Основные особенности *с точки зрения пользователя* программного комплекса поддержки рациональных вычислений:

- Класс `overlong` расширяет возможности целочисленных вычислений в гетерогенной вычислительной среде.
- Объектами класса `rational` являются обыкновенные дроби p/q , где p , q – объекты класса `overlong`.

- Объём памяти, занимаемый такими объектами, определяется значениями представляемых чисел, их диапазон ограничен только объёмом адресуемой памяти.
- Объекты класса `overlong` по умолчанию используют позиционную систему счисления по основанию 2^{32} .
- Для объектов классов `overlong` и `rational` определены все операторы, операции и бинарные отношения, используемые для стандартных числовых типов данных.
- Для объектов классов `overlong` и `rational` описаны все необходимые операции для пересылки в распределенной вычислительной среде с помощью функций библиотеки MPI.

Современная версия комплекса – библиотека классов «Exact Computation 2.0», зарегистрирована в Федеральной службе по интеллектуальной собственности. Основные особенности комплекса «Exact Computation 2.0» *с технологической точки зрения*:

- Для облегчения сопровождения и модификации классов операции с памятью инкапсулированы в отдельный класс `MemHandle`,
- выполнение базовых арифметических операций над числами полностью производится в рамках реализации класса `ArifRealization`.
- Архитектура позволяет использовать несколько типов вычислителей в рамках одного приложения, исполняемого на узлах различной конфигурации. Для этого достаточно иметь соответствующую реализацию классов `MemHandle` и `ArifRealization` для каждого типа вычислителя.
- Для пользователя библиотеки все нюансы, связанные с различными вычислительными элементами, скрыты интерфейсом библиотеки, интерфейс независим от конкретной реализации.
- Для объектов классов `overlong` и `rational` определены все операторы, операции и бинарные отношения, используемые для стандартных числовых типов данных.

Техника разбиения класса `overlong`, на три части «интерфейс-память-арифметика», позволяет гибко использовать возможности вычислительной системы. Учтена возможность использования современных гетерогенных вычислительных сред с GPU ускорителями Nvidia (возможность задействовать GPU от компании AMD и встроенные графические ускорители, например, Intel HD Graphics даёт язык OpenCL, однако эксперименты показали большую трудоемкость процесса кодирования при малом ускорении и, следовательно, малом приросте эффективности приложений в целом).

Параллельная архитектура и низкая тактовая частота ядер графического процессора требуют разработки параллельных алгоритмов базовых арифметических операций, поскольку традиционные последовательные алгоритмы не позволяют в полной мере использовать массовый параллелизм и легковесные нити, являющиеся отличительной чертой графического процессора.

Сравнение чисел. Для случая сравнения чисел разной длины или разного знака алгоритм тривиален, нетривиален случай двух двух чисел одинаковой длины в len разрядов. Эффективность разработанного алгоритма зафиксирована в следующем утверждении.

Утверждение 2 (3.3.1). *Алгоритм сравнения двух чисел одинаковой длины len требует $\log_2 len - 3$ параллельных шага. При этом быстрее выполнить операцию сравнения теоретически невозможно.*

Сложение (вычитание) чисел. Степень параллелизма алгоритма сложения (вычитания) зависит от входных данных (операндов) операции. Разработанный алгоритм имеет следующие характеристики.

Утверждение 3 (3.3.2). *Алгоритм сложения (вычитания) в лучшем случае выполняется полностью параллельно за 1 параллельный шаг, в среднем случае за 3 параллельных шага, в худшем случае является полностью последовательным и выполняется за число шагов равное длине большего из операндов.*

Умножение и деление чисел. Операция умножения одна из наиболее затратных по времени. Для выполнения операции параллельного умножения используется быстрая разделяемая между нитями блока `shared` память. В реализации существенно используется особенности архитектуры GPU Nvidia, а именно, полностью синхронное выполнение инструкций в рамках одного *warps*. Это избавляет от дополнительных синхронизаций. Для операции деления наиболее перспективными для вычислителя с массовым параллелизмом представляются итерационные схемы получения результата.

Схема работы с программным комплексом для нахождения псевдорешения решения ИСЛАУ. На ввод программному комплексу подается соответствующим образом подготовленный текстовый файл, ответ выдается файл в двух формах: в точном – рациональные дроби, а также в приближенном, удобном для восприятия человеком. Все вычисления выполняются с рациональным типом данных.

В четвертой главе диссертации приводятся результаты численных экспериментов тестирования программного обеспечения для дробно-рациональных вычислений в гетерогенных вычислительных системах с графическими ускорителями. Строятся таблицы для выбора оптимальных

параметров для параллельных алгоритмов базовых арифметических операций.

В главе приводятся численные расчеты применения концепции псевдорешения для различных математических моделей, а также примеры учета интервальной неопределенности для повышения качества модели и учета погрешностей для проведения доказательного вычислительного эксперимента.

Вычислительный эксперимент проводился на компьютере с процессором Intel Core i7-950 3.06 ГГц, 6 Гб ОЗУ, GPU Nvidia 460(1 Гб GDDR5) GPU Nvidia 660Ti, под управлением ОС Win 7 x64, в качестве компилятора был выбран 64-разрядный Visual C++ 2011.

Производительность сравнения. Операция сравнения может быть реализована на уровне нижней теоретической оценки для параллельного алгоритма с len входными данными и результатом, зависящим от всех входов. Ускорение по сравнению с последовательной реализацией на CPU составило до 10-12 раз. Для GPU предпочтительны операции над более длинными числами, поскольку основную часть времени выполнения составляют расходы на запуск расчета на GPU(до 99% времени).

Производительность сложения (вычитания). При сложении случайных чисел, когда длина переноса в среднем не превосходит 2 разрядов, на первый план выходит масштабируемость операции. Для длинных переносов решающим фактором является частота ядер GPU, которая отвечает за скорость исполнения последовательных инструкций.

Ускорение по сравнению с последовательной реализацией на CPU составило до 15-20 раз. Основную часть времени выполнения составляют расходы на запуск расчета на GPU (до 99% времени), последовательные переносы большой длины снижают эффективность операции на GPU.

Эффективность реализации арифметических операций, достигается максимальной загрузкой вычислительных узлов и минимизацией накладных расходов. Загрузку вычислительных узлов устройства регулируется за счет размера блока параллельных нитей. В работе проведено сравнение и выбраны оптимальные характеристики: для ускорителей архитектуры FermiTM блок из 512 или 1024 нитей, для ускорителей архитектуры KeplerTM – блок в 1024 нити. Накладные расходы минимизированы за счет использования быстрой «памяти на чипе».

Производительность умножения. Операция умножения существенно зависит от объема быстрой (устанавливаемой непосредственно на чипе процессора) памяти ускорителя. Поэтому архитектура Kepler показывает лучшие результаты.

Результаты применения в математическом моделировании. Были рассмотрены две математические модели: модель Леонтьева с интервальными коэффициентами и хеометрический эксперимент анализа смеси веществ методом Фирордта. Результаты хорошо согласуются с решением полученным иными методами, а учет интервальной неопределенности повышает качество математической модели.

В приложениях приложены копии свидетельств о государственной регистрации программы для ЭВМ для компонентов программного комплекса, а также приводятся примеры выходных файлов.

Результаты, выносимые на защиту.

- Предложен новый подход к регуляризации СЛАУ за счет погружения в ИСЛАУ и перехода к поиску точки из допускового множества решений.
- Предложен новый подход к поиску точки допускового из множества интервальной системы линейных алгебраических уравнений, позволяющий наряду с исследованием совместности задачи о допусках для ИСЛАУ проводить коррекцию правой части системы. Учет интервальной неопределённости в исходных данных повышает качество математического моделирования.
- Разработан и протестирован масштабируемый параллельный алгоритм для нахождения псевдорешения интервальной системы линейных алгебраических уравнений в распределённых гетерогенных вычислительных системах с применением точных дробно-рациональных типов данных. Эффективность реализованного численного метода достигнута за счет применения технологии крупно-зернистого параллелизма.
- Разработан и протестирован комплекс программ для безошибочных дробно-рациональных вычислений. Выполнена реализация последовательных алгоритмов для базовых арифметических операций и их параллельных версий для распределённых вычислительных систем с GPU. Применение технологии GPGPU позволяет на порядок увеличить эффективность данного ПО по сравнению с последовательной версией. Разработан и реализован программный комплекс для численного решения задач математического моделирования.

Публикации автора по теме диссертации

Статьи, опубликованные в научных журналах и изданиях, которые включены в перечень, российских рецензируемых научных журналов и изданий для опубликования основных научных результатов диссертаций:

1. Панюков А.В., Голодов В.А. Подход к решению систем линейных алгебраических уравнений с интервальной неопределенностью в исходных данных. // Вестник Южно-Уральского Государственного Университета. Серия: «Математическое моделирование и программирование». – 2013. – Т. 6, № 2. – С. 109–119.
2. Панюков А.В., Голодов В.А. Программная реализации алгоритма решения системы линейных алгебраических уравнений с интервальной неопределенностью в исходных данных //Управ. большими системами. М.: ИПУ РАН. – 2013. – № 43. – С. 78–94.
3. Panyukov A.V., Golodov V.A. Computing Best Possible Pseudo-Solutions to Interval Linear Systems of Equations//Reliable Computing. Volume 19. – 2014. – Issue 2. – pp. 215–228.

Другие научные публикации:

4. Голодов В.А. Библиотека классов «Exact Computation 2.0». Государственная регистрация программы для ЭВМ №2013612818 от 14.03.2013. //Программы для ЭВМ, базы данных, топологии интегральных микросхем. Оф. бюллетень. – М.: ФИПС, 2013. – № 3. – С. 251. Доступна по адресу: <http://195.208.85.248/Archive/EVM/2013/201302/DOC/RUNW/000/002/013/612/818/document.pdf>
5. Голодов В.А., Панюков А.В. Программа «ExactLinPSolutor 1.0». Государственная регистрация программы для ЭВМ №2014610445 от 9.01.2014. // Программы для ЭВМ, базы данных, топологии интегральных микросхем. Оф. бюллетень. – М.: ФИПС, 2014. – № 2. – Доступна по адресу: <http://195.208.85.248/Archive/EVM/2014/2014.02.20/DOC/RUNW/000/002/014/610/445/document.pdf>
6. Голодов В.А., Панюков А.В. Программа «ExactISLAYSolutor 1.0». Государственная регистрация программы для ЭВМ №2014610333 от 9.01.2014. // Программы для ЭВМ, базы данных, топологии интегральных микросхем. Оф. бюллетень. – М.: ФИПС, 2014. – № 2. – Доступна по адресу: <http://195.208.85.248/Archive/EVM/2014/2014.02.20/DOC/RUNW/000/002/014/610/333/document.pdf>
7. Голодов В.А. Разработка программной реализации некоторых операций длинной арифметики для гетерогенных вычислительных систем //Научный поиск: материалы четвертой научной конференции аспирантов и докторантов. Естественные науки. – Челябинск: Издательский центр ЮУрГУ, 2011. – С. 16–20.

8. Панюков А.В., Голодов В.А. Вычисление псевдорешений систем линейных алгебраических уравнений с интервальной неопределенностью коэффициентов // Алгоритмический анализ неустойчивых задач: тез. Докл. Международной конференции, посвященной памяти В.К. Иванова, Екатеринбург, 31 октября – 5 ноября. – Екатеринбург: ИММ УрО РАН, 2011. – С. 261–262.
9. Голодов В. А. Адаптация библиотеки «Exact Computational» для гетерогенных вычислительных сред // Научный поиск [текст]: материалы четвертой научной конференции аспирантов и докторантов. Естественные науки. – Челябинск: Издательский центр ЮУрГУ, 2012.
10. Панюков А.В., Голодов В.А. Техника программной реализации алгоритма решения системы линейных алгебраических уравнений с интервальной неопределенностью в исходных данных // «Параллельные вычисления и задачи управления» РАСО'2012. Шестая международная конференция, Москва, 24-26 октября. Труды в 3 т. – М.: ИПУ РАН, 2012. – Т. 2. – С. 155–166.
11. Panyukov A.V., Golodov V.A. Computing the best possible pseudo-solutions to interval linear systems of equations // 15th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Verified Numeric (SCAN'2012, Novosibirsk, Russia, September 23-29, 2012): Book of abstracts. – Institute of Computational Technologies Publisher, 2012. – Pp. 134–135.
12. Голодов В.А. Распределённые символьные дробно-рациональные вычисления на процессорах x86 и x64 // Параллельные вычислительные технологии (ПаВТ'2012): труды международной научной конференции (Новосибирск, 26 - 30 марта). – Челябинск: Издательский центр ЮУрГУ, 2012. – С. 719.
13. Панюков А. В., Голодов В. А. Масштабируемость алгоритмов арифметических операций в позиционных системах счисления. // Информационные технологии и системы: тр. Второй междунар. науч. конф., Банное, Россия, 27 февр.–3 марта. – 2013. – Т. 6. – С. 109–119.
14. Panyukov A. V., Golodov V. A. Solving interval linear system of equations using gpu computing. // GPU Technology Conference 2013 (San Jose, California, March 18–21, 2013). <http://www.gputechconf.com/posters>. – Nvidia Corporation, 2013.
15. Panyukov A. V., Golodov V. A. Scalability of algorithms for arithmetic's operations in radix notation // GPU Technology Conference 2013 (San Jose, California, March 18–21, 2013). <http://www.gputechconf.com/posters>. – Nvidia Corporation, 2013.