

На правах рукописи

Рахуба Максим Владимирович

Тензорные методы решения многомерных
частичных задач на собственные значения

01.01.07 — Вычислительная математика

АВТОРЕФЕРАТ

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

Москва 2017

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего профессионального образования «Московский физико-технический институт (государственный университет)».

Научный руководитель: *Оселедец Иван Валерьевич,*
доктор физико-математических наук, доцент
Сколковского института науки и технологий

Официальные оппоненты: *Корнеев Вадим Глебович,*
доктор физико-математических наук, профессор
Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет»

Хоромский Борис Николаевич,
доктор физико-математических наук, ведущий
научный сотрудник Института математики в
науках Общества Макса Планка, Германия

Ведущая организация: *Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Казанский (Приволжский) федеральный университет»*

Защита состоится «___» _____ 2017 г. в __:___ на заседании диссертационного совета Д 002.045.01 в Институте вычислительной математики Российской академии наук, расположенном по адресу: 119333, г. Москва, ул. Губкина, д. 8.

С диссертацией можно ознакомиться в библиотеке и на сайте ИВМ РАН
<http://www.inm.ras.ru>.

Автореферат разослан «___» _____ 2017 г.

Ученый секретарь
диссертационного совета,
д.ф.-м.н., профессор

Бочаров Геннадий Алексеевич

Общая характеристика работы

Актуальность темы исследования. Диссертация посвящена решению многомерных задач на собственные значения. Основным объектом исследования являются *многомерные массивы*, которые естественным образом возникают, например, при дискретизации многомерных дифференциальных уравнений на прямоугольной сетке. Многомерный массив размера $n \times \dots \times n$ с размерностью d задается с помощью n^d чисел. Это означает, что память и количество операций, необходимые для работы с таким массивом, растут экспоненциально с размерностью задачи d , что приводит к значительным затратам вычислительных ресурсов. Несложно оценить, что при $n = 2$ и $d = 300$ количество элементов такого массива равно 2^{300} и превышает оценку числа атомов во Вселенной 10^{80} .

Многомерные задачи на собственные значения возникают в ряде приложений, например, в задачах квантовой химии, базирующихся на решении уравнения Шредингера. Расчет больших молекул с помощью этого уравнения даже на современных суперкомпьютерах может занимать месяцы расчетного времени. Поэтому разработка новых быстрых методов решения многомерных задач на собственные значения является актуальной задачей.

Одним из способов преодоления экспоненциального роста числа параметров с размерностью задачи d является подход *тензорных разложений*, который получил активное развитие в последнее десятилетие. Тензорные разложения позволяют приближать с заданной точностью многомерные массивы с помощью небольшого числа параметров. Важной особенностью тензорного подхода является возможность строить тензорные разложения на основе стандартных алгоритмов вычислительной линейной алгебры. Это позволяет легко адаптировать алгоритмы к задачам, возникающим в различных приложениях. Однако, даже если мы обладаем априорной информацией о том, что искомое решение может быть представлено с помощью тензорных разложений с небольшим числом параметров, поиск этого решения может оказаться нетривиальной задачей. В частности, существующие тензорные алгоритмы для поиска собственных значений и собственных векторов имеют сильную зависимость от числа параметров разложения, а также от числа собственных значений, которые требуется найти. Поэтому необходимо разрабатывать новые эффективные методы решения этой задачи.

Цель диссертационной работы. Целью диссертационной работы является разработка новых тензорных методов решения многомерных частичных задач на собственные значения. Под частичной задачей понимается поиск части спектра, например, нескольких минимальных собственных значений.

Научная новизна. Предложены новые методы решения многомерных частичных задач на собственные значения с использованием тензорных разложений. Для случая линейного оператора предложен новый метод нахождения целевого собственного значения, базирующийся на методе Якоби-Дэвидсона и оптимизации на нелинейных многообразиях. Получены результаты о сходимости метода. Также предложен метод ALS II (ALS Inverse Iteration), базирующийся на методе попеременных наименьших квадратов (alternating least squares, ALS) и методе обратной итерации, получены оценки о локальной сходимости. Для попеременной минимизации функционалов предложена теория локальной сходимости, показывающая связь метода с мультипликативным методом Шварца. Предложен новый метод поиска нескольких собственных значений, базирующийся на итерационных методах и нелинейном предобуславливателе. На основе предложенных методов получено высокоточное решение уравнения Шредингера для расчета первых 84 колебательных уровней молекулы ацетонитрила.

На примере уравнений Хартри-Фока (ХФ) и Кона-Шэма (КШ) предложен новый метод решения задач на собственные значения с нелинейным оператором. Предложенный метод позволяет с высокой точностью решать уравнения ХФ и КШ. Для быстрого вычисления возникающих при решении уравнений ХФ и КШ сверток предложен новый быстрый алгоритм.

Практическая значимость работы. Предложенные методы могут быть использованы как для решения многомерных частичных задач на собственные значения, возникающих при дискретизации дифференциальных уравнений, так и для изначально дискретных задач, например, для расчета спектра в модели Гейзенберга, если известно, что решение может быть представлено с помощью тензорных разложений с малым числом параметров. Более того, предложенные методы могут быть использованы для решения задач малой размерности с помощью подхода квантизации^{1,2}. Реализованный программный код может быть адаптирован под конкретную прикладную за-

¹Khoromskij B. N. Tensor-structured preconditioners and approximate inverse of elliptic operators in \mathbb{R}^d // Constr. Approx. 2009. Т. 30. С. 599–620.

²Оселедец И. В. О приближении матриц логарифмическим числом параметров // ДАН. 2009. Т. 428, № 1. С. 23–24.

дачу. Для этого достаточно задать в требуемом формате многомерный оператор, собственные значения которого необходимо найти.

Теоретическая значимость работы заключается в обосновании сходимости предлагаемых методов решения многомерных задач на собственные значения. Также в рамках диссертации разработана теория локальной сходимости метода попеременных направлений для минимизации функционалов.

Основные положения, выносимые на защиту. Основным результатом работы являются новые эффективные методы решения многомерных задач на собственные значения, их обоснование, а также применение к нескольким прикладным задачам. На защиту выносятся следующие результаты и положения.

1. Предложено обобщение метода Якоби-Дэвидсона с помощью подходов римановой оптимизации при ограничении на тензорные ранги решений. Получены результаты о сходимости метода.
2. Предложен ALS II метод, базирующийся на методе попеременных наименьших квадратов и методе обратной итерации. Получены оценки локальной сходимости.
3. Предложена концепция предобуславливания на многообразиях с помощью подхода попеременных наименьших квадратов. Концепция применена к методу LOBPCG. С помощью метода проведен высокоточный расчет колебательного спектра молекулы ацетонитрила.
4. Предложен новый тензорный метод решения уравнений Хартри-Фока и уравнений Кона-Шэма, являющихся задачами на собственные значения с нелинейным интегро-дифференциальным оператором.
5. Предложен быстрый алгоритм многомерной свертки в тензорных форматах на основе метода крестовой аппроксимации тензоров, который используется для быстрого вычисления оператора, возникающего в уравнениях Хартри-Фока и Кона-Шэма.

Апробация работы. Результаты диссертационной работы докладывались автором и обсуждались на следующих научных семинарах и конференциях:

- Zurich Colloquium in Applied and Computational Mathematics, Eidgenössische Technische Hochschule Zürich (ETH), 2017, Zürich
- Семинар “Вычислительная математика и приложения” ИВМ РАН (2017, рук. д.ф.-м.н. В.И. Агошков, д.ф.-м.н. А.Б. Богатырев, чл.-корр. Ю.В. Василевский, д.ф.-м.н. Ю.М. Нечепуренко, ак. Е.Е. Тыртышников)

- International Conference on Scientific Computation and Differential Equations, University of Bath, 2017, Bath
- Научная конференция “Ломоносов”, 2017, Москва
- 88-th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM), 2017, Weimar
- 5-th workshop on “High dimensional quantum dynamics: challenges and opportunities” 2016, Rostock
- Workshop “Quantum Dynamics: From Algorithms to Applications”, 2016, Greifswald
- 59-я научная конференция МФТИ, 2016, Москва
- 20-th Conference of ILAS, 2016, Leuven
- 4-th International Conference on Matrix Methods in Mathematics and Applications, Skolkovo Institute of Science and Technology, 2015, Moscow
- Workshop: Low-rank Optimization and Applications, Hausdorff Center for Mathematics, 2015, Bonn
- Workshop on Matrix Equations and Tensor Techniques, EPFL, 2013, Lausanne
- 56-я научная конференция МФТИ, 2013, Москва

Публикации. Результаты диссертации опубликованы в 10 печатных работах, из них 5 статей [1–5], 3 из которых [1–3] опубликованы в журналах, индексируемых Web of Science, а также 5 печатных работ [6–10] в сборниках тезисов конференций.

Личный вклад автора. Диссертационное исследование является самостоятельным законченным трудом автора. Лично автором предложена идея малорангового метода Якоби-Дэвидсона [4], а также идея метода предобуславливания на многообразиях [1]. Исследования и разработка алгоритмов в работах [1–4] осуществлены совместно с И.В. Оселедцем, вклад авторов равнозначен. Теоретический результат в совместной работе [5] принадлежит соавторам в равной степени. Реализация алгоритмов, а также подготовка численных экспериментов в работах [1–5] выполнены автором самостоятельно. Постановка задачи в работах [2], [3] выполнена И.В. Оселедцем. Автором совместно с Д.А. Колесниковым получены оценки локальной сходимости для предлагаемого метода ALS II, вклад авторов равнозначен.

Структура работы. Диссертация состоит из введения, четырех глав и заключения. Полный объем диссертации составляет 167 страниц с 19 рисунками и 12 таблицами. Список литературы содержит 152 наименования.

Содержание работы

Во **введении** обосновывается актуальность темы диссертации, формулируются цели и задачи, указывается значимость и научная новизна работы.

В **первой главе** описываются используемые в диссертационной работе тензорные разложения, а также подходы к решению частичной задачи на собственные значения в тензорных форматах. Для метода попеременных наименьших квадратов (alternating least squares, ALS) предлагается теория сходимости, показывающая связь метода с методом Шварца.

Под *тензором* в диссертации понимается многомерный массив чисел $\mathcal{X} \in \mathbb{R}^{n_1 \times \dots \times n_d}$. Число d называется *размерностью* тензора, n_k – размер k -й моды. В диссертационной работе используются два тензорных разложения: разложение Таккера³ и разложение тензорного поезда⁴ (ТТ разложение), которые формулируются в разделе 1.1. Для краткости изложения дадим определение только для ТТ разложения. Тензор \mathcal{X} представляется в виде ТТ разложения, если он может быть записан как

$$\mathcal{X}(i_1, \dots, i_d) = X^{(1)}(i_1)X^{(2)}(i_2) \dots X^{(d)}(i_d), \quad (1)$$

где $X^{(k)}(i_k) \in \mathbb{R}^{r_{k-1} \times r_k}$ называются *ТТ-ядрами*, причем $r_0 = 1$ и $r_d = 1$. Будем обозначать вектор с компонентами r_k за $\mathbf{r} = \{r_1, \dots, r_{d-1}\}$ и называть его *ТТ-рангом* тензора. Количество параметров ТТ разложения составляет $\mathcal{O}(dnr^2)$, $r = \max_k r_k$, $n = \max_k n_k$. Аналогичным образом определяется ТТ разложение для линейного оператора $\mathcal{A} : \mathbb{R}^{n_1 \times \dots \times n_d} \rightarrow \mathbb{R}^{n_1 \times \dots \times n_d}$:

$$\mathcal{A}(i_1, \dots, i_d, j_1, \dots, j_d) = A^{(1)}(i_1, j_1) \dots A^{(d)}(i_d, j_d),$$

где ТТ-ядра $A^{(k)}(i_k, j_k) \in \mathbb{R}^{R_{k-1} \times R_k}$, $R_0 = R_d = 1$. Для удобства будем использовать обозначение $R = \max_k R_k$.

Предположим, что оператор \mathcal{A} является симметричным положительно определенным с простым минимальным собственным значением. Пусть

$$\mathcal{M}_{\mathbf{r}} \stackrel{\text{def}}{=} \{\mathcal{X} \in \mathbb{R}^{n_1 \times \dots \times n_d} : \text{ТТ-rank}(\mathcal{X}) = \mathbf{r}\}$$

обозначает множество тензоров с ТТ-рангом \mathbf{r} . Известно⁵, что $\mathcal{M}_{\mathbf{r}}$ образует гладкое подмногообразие в $\mathbb{R}^{n_1 \times \dots \times n_d}$. В разделе 1.3 ставится задача о поиске

³Tucker L. R. Some mathematical notes on three-mode factor analysis // Psychometrika. 1966. Т. 31. С. 279–311.

⁴Oseledets I. V. Tensor-train decomposition // SIAM J. Sci. Comput. 2011. Т. 33, № 5. С. 2295–2317.

⁵Holtz S., Rohwedder T., Schneider R. On manifolds of tensors of fixed TT-rank // Numer. Math. 2012. Т. 120, № 4. С. 701–731.

минимального собственного значения в тензорных форматах:

$$\begin{aligned} \mathfrak{K}(\mathcal{X}) &\stackrel{\text{def}}{=} \langle \mathcal{A}\mathcal{X}, \mathcal{X} \rangle \rightarrow \min, \\ \mathcal{X} &\in \mathcal{M}_r, \quad \|\mathcal{X}\| = 1, \end{aligned} \quad (2)$$

где скалярное произведение двух тензоров определяется как $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1, \dots, i_d} \mathcal{X}(i_1, \dots, i_d) \mathcal{Y}(i_1, \dots, i_d)$, а норма — $\|\mathcal{X}\| = \langle \mathcal{X}, \mathcal{X} \rangle^{1/2}$. Если точный собственный вектор приближается с высокой точностью ТТ разложением, то и решение оптимизационной задачи (2) является приближением собственного вектора. Отметим, что в главе 3 оператор \mathcal{A} является нелинейным, то есть $\mathcal{A} \equiv \mathcal{A}(\mathcal{X})$.

Кроме того, в диссертационной работе рассматривается поиск B минимальных собственных значений. Без ограничения по рангу эта задача тоже может быть поставлена в виде следующей задачи оптимизации

$$\begin{aligned} \text{trace}(\mathbf{X}^\top \mathbf{A} \mathbf{X}) &\rightarrow \min, \\ \mathbf{X}^\top \mathbf{X} &= I_B, \end{aligned} \quad (3)$$

где вместо многомерного оператора \mathcal{A} рассматривается соответствующая матрица \mathbf{A} , $\mathbf{X} = [\text{vec}(\mathcal{X}_1), \dots, \text{vec}(\mathcal{X}_B)] \in \mathbb{R}^{n_1 \dots n_d \times B}$, и vec обозначает векторизацию многомерного массива:

$$\text{vec}(\mathcal{X})_{i_1 \dots i_d} = \mathcal{X}(i_1, \dots, i_d), \quad i_k = 1, \dots, n_k, \quad k = 1, \dots, d.$$

Обычно ограничение на ранг для задачи (3) накладывается для блочного ТТ формата⁶, однако для больших B ранг такого представления быстро растет. Поэтому в диссертации рассматривается независимое ограничение на ранг каждого из векторов \mathcal{X}_i , $k = 1, \dots, B$, и используется обобщение классических итерационных методов.

Для решения задач минимизации в тензорных форматах существует несколько подходов, идеи которых используются в диссертации. Во-первых, в разделе 1.4 приводится описание обобщения классических итерационных методов с округлением по рангу, которое производится в конце каждой итерации для избежания роста количество параметров разложения.

Раздел 1.5 посвящен ALS подходу (попеременная оптимизация). Заметим, что подмножество ТТ тензоров со всеми фиксированными ядрами, кроме одного, образует линейное подпространство в пространстве всех тензоров. Это наблюдение может значительно упростить процесс оптимизации,

⁶Computation of extreme eigenvalues in higher dimensions using block tensor train format / S. V. Dolgov [и др.] // Computer Phys. Comm. 2014. Т. 185, № 4. С. 1207–1216.

так как обновление каждого ядра производится на линейном подпространстве, а не на всем нелинейном многообразии тензоров фиксированного ранга. Метод заключается в последовательном выборе ядер и оптимизации по выбранному ядру при условии зафиксированных остальных ядер. Процедура продолжается до сходимости. Формально обновление m -го ядра при минимизации функционала f может быть записано в виде

$$X^{(m)} \leftarrow \arg \min_Y f \left(X^{(1)}, \dots, X^{(m-1)}, Y, X^{(m+1)}, \dots, X^{(d)} \right).$$

В работе разработана теория локальной сходимости метода. В отличие от ранее предложенной теории⁷, в рамках предлагаемого подхода удается явно проиллюстрировать, что сходимость метода определяется через взаимосвязь классического мультипликативного метода Шварца и кривизны многообразия малоранговых тензоров. Для формулировки результата введем следующие обозначения.

Рассмотрим C^1 функционал $f: \mathbf{V} \rightarrow \mathbf{V}$, \mathbf{V} – гильбертово пространство. Любому $x \in \mathbf{V}$ поставим в соответствие замкнутое подпространство $T(x)$ пространства \mathbf{V} . Затем предположим, что задано разложение $T(x)$ в сумму d замкнутых, вообще говоря, пересекающихся $T_i(x)$: $T(x) = T_1(x) + \dots + T_d(x)$. В случае тензоров фиксированного ранга $T(x)$ является касательным пространством к \mathcal{M}_r в точке x , а $T_i(x)$ являются линейными подпространствами, которые соответствуют тензорам со всеми фиксированными ядрами, кроме i -го ядра. Затем определим d отображений \mathbf{P}_i , $i = 1, \dots, d$, так, что для каждого $x \in \mathbf{V}$ линейный оператор $\mathbf{P}_i(x)$ является ортопроектором на подпространство $T_i(x)$. Соответственно, $\mathbf{P}(x)$ обозначает ортопроектор на $T(x)$. Далее, пусть \mathbf{S}_i , $i = 1, \dots, d$, обозначают (нелинейные) операторы на \mathbf{V} так, что $y = \mathbf{S}_i(x)$ удовлетворяет $y \in x + T_i(x)$, $\mathbf{P}_i(x) \nabla f(y) = 0$. Это означает, что \mathbf{S}_i отображает x в критическую точку f на гиперплоскости $x + T_i(x)$. Если, например, f является строго выпуклой и коэрцитивной, тогда оператор \mathbf{S}_i определяется единственным образом и соответствует минимизации f на $x + T_i(x)$. Попеременная оптимизация на “двигающихся” гиперплоскостях соответствует итерации следующего вида

$$x_{\ell+1} = \mathbf{S}(x_\ell) := (\mathbf{S}_d \circ \dots \circ \mathbf{S}_1)(x_\ell).$$

⁷Rohwedder T., Uschmajew A. On Local Convergence of Alternating Schemes for Optimization of Convex Problems in the Tensor Train Format // SIAM J. Num. Anal. 2013. Т. 51, № 2. С. 1134–1162.

Известно, что если спектральный радиус ρ производной оператора \mathbf{S} меньше единицы в неподвижной точке \bar{x} , то отсюда вытекает локальная сходимость метода. Для получения формулы для \mathbf{S}' определим сначала линейный оператор $\bar{\mathbf{N}}_i$ на \mathbf{V} так, что $\bar{\mathbf{N}}_i h := \mathbf{P}'_i(\bar{x}; h) \nabla f(\bar{x})$ для всех h . С помощью $\bar{\mathbf{A}}(x) = \nabla^2 f(x)$ обозначается гессиан f в x . Для краткости будем использовать следующие обозначения $\bar{\mathbf{P}}_i := \mathbf{P}_i(\bar{x})$, $\bar{\mathbf{P}} := \mathbf{P}(\bar{x})$, $\bar{\mathbf{A}} := \bar{\mathbf{A}}(\bar{x})$, $\bar{\mathbf{B}}_i := (\bar{\mathbf{P}}_i \bar{\mathbf{A}} \bar{\mathbf{P}}_i)^{-1}$. Доказана следующая теорема:

Теорема 1.2. *Предположим, что все \mathbf{P}_i и \mathbf{S}_i являются непрерывно дифференцируемыми в окрестности неподвижной точки \bar{x} , и что f является дважды непрерывно дифференцируемой в окрестности \bar{x} . Предположим, что все $\bar{\mathbf{B}}_i = (\bar{\mathbf{P}}_i \bar{\mathbf{A}} \bar{\mathbf{P}}_i)^{-1}$ существуют на $T_i(\bar{x})$. Тогда*

$$\mathbf{S}'(\bar{x}) = \prod_{i=d}^1 [(\mathbf{I} - \bar{\mathbf{P}}_i \bar{\mathbf{A}}) - \bar{\mathbf{B}}_i \bar{\mathbf{N}}_i]. \quad (4)$$

Отметим, что при $\bar{\mathbf{N}}_i = 0$ получается формула, аналогичная формуле для мультипликативного метода Шварца⁸ при ограничении на подпространство $\bar{T}(x)$ и замене $\bar{\mathbf{A}}$ на $\bar{\mathbf{P}} \bar{\mathbf{A}} \bar{\mathbf{P}}$. Поэтому справедлива теорема:

Теорема 1.3. *Предположим, что $\bar{\mathbf{N}}_i = 0$ и гессиан $\bar{\mathbf{A}}$ является положительно определенным на $T(\bar{x})$. Тогда $\rho(\bar{\mathbf{P}} \mathbf{S}'(\bar{x}) \bar{\mathbf{P}}) < 1$.*

Случай $\bar{\mathbf{N}}_i = 0$ возникает в двух важных сценариях. Во-первых, если подпространства $T_i(x)$ не меняются для всех x в окрестности \bar{x} , тогда $\mathbf{P}'_i(\bar{x}) = 0$. Этот случай возникает в классическом блочном покоординатном спуске, который также известен под названием *нелинейный блочный метод Гаусса-Зейделя*, и для которого известны результаты о сходимости. Во-вторых, операторы $\bar{\mathbf{N}}_i$ также равны нулю в неподвижных точках, удовлетворяющих $\nabla f(\bar{x}) = 0$. Это происходит, в частности, в интересующем нас случае, когда рассматривается малоранговая оптимизация, и глобальная критическая точка принадлежит рассматриваемому многообразию малоранговых тензоров. Этот результат используется при доказательстве сходимости ALS II метода в разделе 4.1 диссертации.

В разделе 1.6 формулируются используемые результаты теории римановой оптимизации. Описывается многообразие тензоров фиксированного ранга, а также приводятся некоторые алгоритмические детали.

⁸Frommer A., Nabben R., Szyld D. B. Convergence of stationary iterative methods for Hermitian semidefinite linear systems and applications to Schwarz methods // SIAM J. Matrix Anal. Appl. 2008. Т. 30, № 2. С. 925–938.

Вторая глава посвящена новым методам решения задачи на собственные значения с линейными операторами в тензорных форматах. В разделе 2.1 предлагается обобщение метода Якоби-Дэвидсона на случай тензоров фиксированного ранга.

Классический метод Якоби-Дэвидсона (JD) для поиска целевого собственного значения матрицы A состоит из двух основных частей: *уравнения Якоби* и *ускорения с использованием подпространств*. Уравнение Якоби на k -й итерации метода имеет следующий вид:

$$(I - x_k x_k^\top)(A - \mathcal{K}(x_k)I)(I - x_k x_k^\top)\xi_k = -r_k, \quad x_k^\top \xi_k = 0. \quad (5)$$

Ускорение с использованием подпространств заключается в том, что на каждой итерации решение приближается в некотором базисе, причем базис расширяется с помощью решения (5). А именно, предположим, что на k -й итерации уже получен базис $\mathcal{V}_{b-1} = [v_1, \dots, v_{b-1}]$, $b \leq k$. На $(k+1)$ -й итерации базис расширяется с помощью вектора v_b , являющегося ортогонализацией ξ_k к \mathcal{V}_{b-1} : $\mathcal{V}_b = [\mathcal{V}_{b-1}, v_b]$.

Отметим несколько свойств JD метода. Во-первых, уравнение Якоби является системой линейных уравнений с матрицей, число обусловленности которой не ухудшается на подпространстве $x_k^\top \xi_k = 0$, когда $\mathcal{K}(x_k)$ приближается к точному собственному значению. При этом уравнение можно решать с помощью стандартных крыловских итерационных методов, благодаря которым автоматически поддерживается ортогональность поправки ξ_k и x_k даже, если система решается неточно. Более того, использование подпространств может значительно ускорить сходимость метода. В диссертационной работе предлагается обобщение JD метода при ограничении на тензорные ранги решения, причем у предлагаемого метода сохраняются отмеченные положительные свойства классического JD метода.

Известно, что JD метод может быть рассмотрен как метод Ньютона на сфере $S^{n-1} \stackrel{\text{def}}{=} \{x : \|x\| = 1\}$ с ускорением на подпространствах. Идея предлагаемого в диссертации метода заключается в том, чтобы рассмотреть метод Ньютона на пересечении единичной сферы и многообразия тензоров фиксированного ранга. Для удобства в этой главе рассматривается многообразие тензоров ранга \mathbf{r} , вытянутых в вектор: $\mathcal{M}_{\mathbf{r}} = \{\text{vec}(\mathcal{X}), \mathcal{X} \in \mathbb{R}^{n \times \dots \times n} : \text{TT-rank}(\mathcal{X}) = \mathbf{r}\}$, и вместо многомерного оператора \mathcal{A} рассматривается соответствующая матрица A . Определим $\mathcal{N}_{\mathbf{r}} \stackrel{\text{def}}{=} S^{n^d-1} \cap \mathcal{M}_{\mathbf{r}}$. В диссертации доказана следующая теорема:

Теорема 2.1. \mathcal{N}_r образует гладкое подмногообразие \mathbb{R}^{nd} размерности $\dim(\mathcal{M}_r) - 1$.

Показано, что аналог уравнения Якоби (5) на \mathcal{N}_r имеет вид:

$$(I - x_k x_k^\top) \left[P_{T_{\mathcal{X}_k} \mathcal{M}_r} (A - \mathcal{K}(x_k) I) P_{T_{\mathcal{X}_k} \mathcal{M}_r} \right] (I - x_k x_k^\top) \xi_k = -P_{T_{\mathcal{X}_k} \mathcal{M}_r} r_k, \quad (6)$$

$$\xi_k^\top x_k = 0, \quad \xi_k \in T_{\mathcal{X}_k} \mathcal{M}_r,$$

где $P_{T_{\mathcal{X}_k} \mathcal{M}_r}$ обозначает ортопроектор на касательное пространство \mathcal{M}_r в точке $x_k = \text{vec}(\mathcal{X}_k)$. Уравнение (6) выводится как один шаг неточного метода Ньютона на \mathcal{N}_r , в котором отброшено слагаемое, содержащее кривизну многообразия \mathcal{M}_r . Это слагаемое содержит обращения сингулярных чисел разверток. Сингулярные числа, в свою очередь, могут быть очень малы, если ранг для построения многообразия оказался больше, чем истинный ранг решения. Это ведет к неустойчивости при численных расчетах. В диссертации показывается связь предложенного метода и метода Гаусса-Ньютона для оптимизационных задач с ограничениями.

Показывается, что матрица системы (6) является симметричной, положительно определенной (если ищется минимальное с.з., $A > 0$) на подпространстве $\xi_k^\top x_k = 0$, $\xi_k \in T_{\mathcal{X}_k} \mathcal{M}_r$, а также, что ее число обусловленности не ухудшается при стремлении $\mathcal{K}(x_k)$ к точному собственному значению. Показано, что сложность решения уравнения (6) равна $\mathcal{O}(dnr^2R^2(r+n))$.

Без ускорения с использованием подпространств можно показать, что $x_{k+1} = x_k + \xi_k$, вообще говоря, имеет ранг $2r$ и, значит, не принадлежит \mathcal{M}_r . Для того чтобы “вернуть” x_{k+1} на многообразии используется понятие *ретракции*. Доказано следующее утверждение:

Утверждение 2.2. Пусть R_r является ретракцией с касательного расслоения $T\mathcal{M}_r$ на \mathcal{M}_r , тогда

$$R(x, \xi) = \frac{R_r(x, \xi)}{\|R_r(x, \xi)\|}, \quad (7)$$

является ретракцией на \mathcal{N}_r .

Так как ретракция является нелинейной функцией, а уравнение (6) может решаться неточно, вводится линейный поиск $x_{k+1} = x_k + \tau_k \xi_k$, $\tau_k = \arg \min_{\tau} \mathcal{K}(R(x_k, \tau \xi_k))$. Для эффективности, τ_k находится с помощью правила Армико. В диссертации также доказывається сходимость регуляризованного метода на основе общей теории оптимизации на многообразиях⁹. А

⁹ Absil P.-A., Mahony R., Sepulchre R. Optimization algorithms on matrix manifolds. Princeton University Press, 2009.

именно, функционал \mathcal{K} регуляризируется с помощью норм псевдообратных матриц разверток. В случае $d = 2$ регуляризированный функционал имеет следующий вид $\mathcal{K}_\mu(x) = \mathcal{K}(x) + \mu^2 \|\mathcal{X}^+\|_F^2$, $x = \text{vec}(\mathcal{X})$. Справедлива теорема:

Теорема 2.3. *Любая предельная точка x_*^μ последовательности $\{x^{(k)}\}$ с регуляризованным функционалом \mathcal{K}_μ принадлежит \mathcal{M}_r и удовлетворяет $\text{grad } \mathcal{K}_\mu(x_*^\mu) = 0$.*

Далее вводится ускорение с использованием подпространств для тензорной версии JD метода. Проблема с использованием классической версии ускорения заключается в том, что линейная комбинация b векторов имеет ранг $b\mathbf{r}$, что приводит к заметному замедлению вычислений, а также к неустойчивости при ретракции при больших b . Поэтому в диссертации предлагается использовать подход *переноса векторов*, в котором накопленный базис проецируется на текущее касательное пространство: $\mathcal{V}_b = [P_{T_{\mathcal{X}_k} \mathcal{M}_r} \mathcal{V}_{b-1}, v_b]$, где v_b — ортогонализация ξ_k к $P_{T_{\mathcal{X}_k} \mathcal{M}_r} \mathcal{V}_{b-1}$. Важно отметить, что линейная комбинация любого числа векторов из касательного пространства $T_{\mathcal{X}_k} \mathcal{M}_r$ имеет ранг не больше $2\mathbf{r}$, что позволяет избежать роста рангов.

По аналогии с классическим JD методом показано, что если уравнение (6) решается точно, то итерация $x_{k+1} = R(x_k, \xi_k) \equiv R(x_k + \xi_k)$ может быть записана в виде аналога обратной итерации с адаптивными сдвигами:

$$\begin{aligned} \left[P_{T_{\mathcal{X}_k} \mathcal{M}_r} (A - \mathcal{K}(x_k)I) P_{T_{\mathcal{X}_k} \mathcal{M}_r} \right] \tilde{x} &= x_k, & P_{T_{\mathcal{X}_k} \mathcal{M}_r} \tilde{x} &= \tilde{x}, \\ x_{k+1} &= R(\tilde{x}). \end{aligned} \quad (8)$$

Проведено численное сравнение предложенного JD метода с другими подходами, в частности, с ALS минимизацией отношения Рэлея и методом, базирующемся на AMEn¹⁰ подходе. Предлагаемый метод показал лучшую сходимость, особенно в случае, когда линейные системы решаются неточно, что часто встречается на практике. Это можно объяснить с помощью нескольких факторов. Во-первых, на каждой итерации JD метода решаются линейные системы, число обусловленности которых не ухудшается при сходимости метода. Во-вторых, используется ускорение на подпространствах, что позволяет дополнительно ускорить сходимость. Более того, оптимизация проводится во всем касательном пространстве, а не на его подпространствах. Из положительных свойств метода следует также отметить возмож-

¹⁰Dolgov S. V., Savostyanov D. V. Corrected one-site density matrix renormalization group and alternating minimal energy algorithm // Numerical Mathematics and Advanced Applications — ENUMATH 2013. T. 103. 2015. С. 335–343.

ность эффективной параллелизации метода по ядрам разложения, а также отсутствие роста ранга.

В разделе 2.2 предлагается метод ALS II — комбинация обратной итерации и ALS подхода. Блочная версия итерации (ALS III) используется для коррекции собственных векторов, полученных с помощью блочного метода из раздела 2.3. Итерация ALS II имеет следующий вид:

$$\begin{aligned} P_{k+\frac{i-1}{d}}^{\neq i} (\mathcal{A} - \sigma \mathcal{I}) P_{k+\frac{i-1}{d}}^{\neq i} \mathcal{X}_{k+\frac{i}{d}} &= P_{k+\frac{i-1}{d}}^{\neq i} \mathcal{X}_k, \quad i = 1, \dots, d, \\ \mathcal{X}_{k+1} &:= \mathcal{X}_{k+1} / \|\mathcal{X}_{k+1}\|, \end{aligned} \quad (9)$$

где $P_{k+\frac{i-1}{d}}^{\neq i}$ являются ортопроекторами на части касательного пространства в точке $\mathcal{X}_{k+\frac{i-1}{d}}$, где все ядра, кроме i -го, фиксированы, σ обозначает сдвиг. Отметим сходство итерации (9) с итерацией (8). Отличие заключается в том, что в ALS II проекция выполняется не на все касательное пространство, а на его подпространства.

Для случая $d = 2$ получена оценка локальной сходимости метода через сходимость ALS итерации для минимизации отношения Рэля. Отметим, что локальная сходимость ALS минимизации следует из Теоремы 1.3. Пусть \mathcal{X}_* — собственный вектор, отвечающий минимальному собственному значению симметричного положительно определенного оператора \mathcal{A} , а $s_{\min}(\mathcal{X}_*)$ обозначает минимальное сингулярное число \mathcal{X}_* . Справедлива теорема:

Теорема 2.3. Пусть $d = 2$ и существует такая окрестность точки $\mathcal{X}_* \in \mathcal{M}_r$, $\|\mathcal{X}_*\| = 1$, что для любого начального приближения из этой окрестности ALS минимизация отношения Рэля сходится со скоростью сходимости $\rho_{\text{als}} < 1$, тогда найдется $\sigma_0 < \lambda_1(\mathcal{A})$: для любого $\sigma \in (\sigma_0, \lambda_1(\mathcal{A}))$ существует такая окрестность, что для любого \mathcal{X}_0 из этой окрестности, ALS II (9) также сходится:

$$\|\mathcal{X}_{k+1} - \mathcal{X}_*\| \leq \left(\rho_{\text{als}} + c \left| \frac{\lambda_1(\mathcal{A}) - \sigma}{\lambda_2(\mathcal{A}) - \sigma} \right| \right) \|\mathcal{X}_k - \mathcal{X}_*\|, \quad c = \frac{C}{s_{\min}(\mathcal{X}_*)} \cdot \frac{\lambda_n(\mathcal{A})}{\lambda_2(\mathcal{A}) - \lambda_1(\mathcal{A})},$$

где C является константой, не зависящей от $\sigma, \mathcal{A}, \mathcal{X}_*, \mathcal{X}_0$.

В разделе 2.2 предлагается блочный решатель, базирующийся на методе LOBPCG¹¹ и ALS оптимизации в качестве предобуславливателя. Для простоты рассмотрим случай поиска одного собственного значения с использо-

¹¹Knyazev A. V. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method // SIAM J. Sci. Comput. 2001. Т. 23, вып. 2. С. 517–541.

ванием LOBPCG:

$$\begin{aligned}\mathcal{R}_k &:= \mathcal{B}(\mathcal{A}\mathcal{X}_k - \lambda_k\mathcal{X}_k), & \mathcal{P}_{k+1} &:= \alpha_2\mathcal{R}_k + \alpha_3\mathcal{P}_k, \\ \mathcal{X}_{k+1} &:= (\alpha_1\mathcal{X}_k + \mathcal{P}_{k+1})/\|\alpha_1\mathcal{X}_k + \mathcal{P}_{k+1}\|,\end{aligned}\tag{10}$$

где \mathcal{B} — предобуславливатель, а $\alpha = (\alpha_1, \alpha_2, \alpha_3)^\top$ выбирается из условия минимизации отношения Рэлея. Отметим, что если $\mathcal{R}_k, \mathcal{X}_k, \mathcal{P}_k$ заданы в ТТ формате, то такие операции, как сложение, умножение матрицы на вектор, вычисление нормы могут быть получены в ТТ формате с использованием явных формул. Для избежания роста рангов дополнительно требуется выполнять операцию округления. Ключевой частью тензорной версии метода LOBPCG является умножение вектора на предобуславливатель \mathcal{B} . В качестве предобуславливателя используется $\mathcal{B} \approx (\mathcal{A} - \sigma\mathcal{I})^{-1}$, а именно, предлагается минимизировать функционал энергии

$$\begin{aligned}\langle (\mathcal{A} - \sigma\mathcal{I})\mathcal{Y}, \mathcal{Y} \rangle - 2\langle (\mathcal{A}\mathcal{X}_k - \lambda_k\mathcal{X}_k), \mathcal{Y} \rangle \rightarrow \min, \\ \text{TT-rank}(\mathcal{Y}) = \mathbf{r}\end{aligned}\tag{11}$$

с помощью нескольких проходов ALS минимизации, то есть вдоль многообразия тензоров фиксированного ранга. Назовем конструкцию такого нелинейного предобуславливателя предобуславливателем на многообразии (*manifold preconditioner, MP*), а метод — MP LOBPCG.

Пусть теперь требуется найти B минимальных собственных значений и собственных векторов. В блочной версии метода вместо трех тензоров $\mathcal{R}_k, \mathcal{X}_k, \mathcal{P}_k$, необходимо рассматривать $3B$ тензоров $\mathcal{R}_k^{(l)}, \mathcal{X}_k^{(l)}, \mathcal{P}_k^{(l)}$, $l = 1, \dots, B$. Векторы $\mathcal{R}_k^{(l)}$ вычисляются независимо по аналогии с (11). Для вычисления $\mathcal{X}_{k+1}^{(l)}, \mathcal{P}_{k+1}^{(l)}$ требуется считать линейные комбинации с $2B$ тензорами, что может привести к значительному росту ранга и увеличению стоимости вычислений. Для избежания роста рангов предлагается использовать *метод крестовой аппроксимации*¹², который позволяет вычислять ТТ разложение тензора по небольшому числу его элементов. Идея заключается в следующем: сначала из ТТ-представления вычисляются элементы тензоров, которые требуется сложить, затем найденные элементы складываются, и по ним строится ТТ-представление с помощью интерполяционной формулы. Отметим, что для быстрого вычисления матрично-векторных умножений также используется ALS оптимизация.

¹²Oseledets I. V., Tyrtshnikov E. E. TT-cross approximation for multidimensional arrays // Linear Algebra Appl. 2010. Т. 432, № 1. С. 70–88.

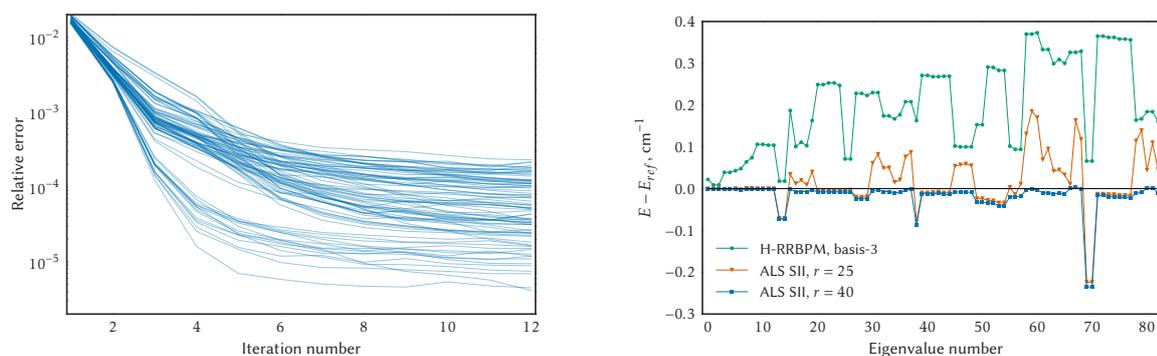


Рис. 1: Сходимость MP LOBPCG (слева) и точность полученных энергий (справа) для молекулы CH_3CN . Референсные энергии E_{ref} получены с помощью сеток Смоляка. Отрицательные значения $E - E_{\text{ref}}$ соответствуют более точным значениям по сравнению с E_{ref} . ALS SII использует решение MP LOBPCG в качестве начального приближения.

В разделе 2.3.3 приводятся результаты расчетов колебательного спектра молекул. Гамильтониан в нормальных координатах q_i имеет вид

$$\mathcal{H} = \sum_{i=1}^d \frac{\omega_i}{2} \left(-\frac{\partial^2}{\partial q_i^2} + q_i^2 \right) + V(q_1, \dots, q_d), \quad (12)$$

где V — поверхность потенциальной энергии (PES), которая обычно задается в виде полинома. Задача дискретизована с помощью псевдоспектрального метода на эрмитовой сетке. Рис. 1 иллюстрируют сходимость и точность метода для расчета колебательного спектра молекулы ацетонитрила CH_3CN . Приводится сравнение с H-RRBPM¹³ методом и методом, базирующемся на сетках Смоляка¹⁴. В диссертации показано, что комбинация MP LOBPCG и ALS SII дает более точные значения энергий при меньших затратах памяти, чем H-RRBPM метод и метод Смоляка. Для наиболее точного базиса (basis-3) H-RRBPM метода предлагаемый метод выигрывает также по скорости: 3 часа против 30 минут расчетного времени.

В **третьей главе** рассматривается задача на собственные значения с нелинейным оператором на примере уравнений Хартри-Фока (ХФ) и Кона-Шэма (КШ), которые используются при расчете электронного спектра атомов и молекул. В силу нелинейности и наличия нелокальных преобразова-

¹³Thomas P. S., Carrington Jr T. Using Nested Contractions and a Hierarchical Tensor Format to Compute Vibrational Spectra of Molecules with Seven Atoms // J. Phys. Chem. A. 2015. С. 13074–13091.

¹⁴Avila G., Carrington Jr T. Using nonproduct quadrature grids to solve the vibrational Schrödinger equation in 12D // J. Chem. Phys. 2011. Т. 134, № 5. С. 054126.

ний типа свертки решение этих уравнений с вычислительной точки зрения является затратным и для больших молекул может занимать недели расчетного времени. Поэтому необходимо разрабатывать новые быстрые методы их решения. Для решения этих уравнений используется дискретизация на равномерной сетке и разложение Таккера, благодаря которому сложность алгоритма растет линейно относительно количества узлов сетки по одному координатному направлению.

В разделе 3.1 приводится формулировка уравнений. Рассмотрим систему с замкнутыми оболочками с N_e электронами и $N = N_e/2$ орбиталями, пусть также N_{nuc} обозначает число ядер с зарядами Z_α , расположенными в $\mathbf{R}_\alpha \in \mathbb{R}^3, \alpha = 1, \dots, N_{nuc}$. Тогда уравнения ХФ/КШ могут быть записаны как

$$H(\Phi) \phi_i(\mathbf{r}) \equiv \left(-\frac{1}{2} \Delta + V(\Phi) \right) \phi_i(\mathbf{r}) = \lambda_i \phi_i(\mathbf{r}), \quad \int_{\mathbb{R}^3} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}, \quad (13)$$

где ϕ_i обозначает неизвестные функции орбитали, $\Phi = [\phi_1 \dots \phi_N]$, а λ_i обозначает энергии орбиталей, $i = 1, \dots, N$. Потенциал V имеет различную форму для уравнений ХФ и КШ:

$$V(\Phi) = - \sum_{a=1}^{N_{nuc}} \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} + \int_{\mathbb{R}^3} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \begin{cases} \int_{\mathbb{R}^3} \frac{\sum_{j=1}^N \phi_j(\mathbf{r}) \phi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \cdot d\mathbf{r}' & \leftarrow \text{ХФ}, \\ V_{xc}(\rho) & \leftarrow \text{КШ}, \end{cases}$$

где нелинейная функция V_{xc} (обменно-корреляционный функционал) зависит от электронной плотности $\rho(\mathbf{r}) = 2 \sum_{i=1}^N \phi_i^2(\mathbf{r})$. Отметим, что в уравнении КШ оператор зависит только от функции плотности, что уменьшает сложность вычислений. Полная энергия системы E может быть посчитана по найденным орбиталям ϕ_i в случае уравнений ХФ и по плотности ρ для КШ.

Раздел 3.2 посвящен итерационному процессу. Стандартным итерационным методом решения уравнений КШ/ХФ является итерация самосогласованного поля (self-consistent field, SCF):

$$H(\Phi^{(k)}) \phi_i^{(k+1)} = \lambda_i^{(k+1)} \phi_i^{(k+1)}, \quad i = 1, \dots, N.$$

Для наших целей реализация SCF итерации, вообще говоря, является такой же сложной, как и решение исходной задачи в силу нелинейности представления Таккера. Следовательно, используется более подходящая блочная ите-

рация Грина для уравнения (13) в форме Липпмана-Швингера,¹⁵ которая является более удобной для реализации алгоритма в тензорных форматах.

Перепишем (13) в виде $\phi_i = -2(-\Delta - 2\lambda_i)^{-1} V \phi_i$. Для простоты рассмотрим итерационный процесс для системы с одной орбиталью $\phi \equiv \phi_1$. В этом случае k -й шаг итерации Грина имеет следующий вид

$$\tilde{\phi} = 2(-\Delta - 2\lambda^{(k)})^{-1} V(\Phi^{(k)})\phi^{(k)}, \quad \phi^{(k+1)} = \tilde{\phi}/\|\tilde{\phi}\|.$$

Энергия вычисляется как $\lambda^{(k+1)} = (H(\Phi^{(k+1)})\phi^{(k+1)}, \phi^{(k+1)})$. Гамильтониан $H(\Phi^{(k+1)})$ содержит оператор Лапласа Δ . Ошибка аппроксимации в формате Таккера контролируется только в L_2 норме, поэтому дискретное дифференцирование может усилить ошибку. Для избежания этой проблемы в диссертации предлагается формула для $\lambda^{(k+1)}$ без вычисления производных. В случае нескольких орбиталей алгоритм выглядит следующим образом:

Алгоритм 3.1 Блочная итерация Грина

- 1: Вычислить $\widehat{\Phi} = (\hat{\phi}_1, \dots, \hat{\phi}_N)$: $\hat{\phi}_i = 2(-\Delta - 2\lambda_i^{(k)})^{-1} V(\Phi^{(k)})\phi_i^{(k)}$.
 - 2: Ортогонализировать $\widehat{\Phi}$: $\widetilde{\Phi} = \widehat{\Phi}L^{-T}$, где $L: LL^T = \int_{\mathbb{R}^3} \widehat{\Phi}^T \widehat{\Phi} d\mathbf{r}$.
 - 3: Вычислить матрицу Фока $F = \int_{\mathbb{R}^3} \widetilde{\Phi}^T H(\Phi^{(k+1)}) \widetilde{\Phi} d\mathbf{r}$ с помощью формулы без производных из Утверждения 3.1.
 - 4: Найти новые энергии с помощью диагонализации F : $\Lambda^{(k+1)} = S^{-1}FS$
 - 5: Найти новые орбитали: $\Phi^{(k+1)} = \widetilde{\Phi}S$.
-

В диссертации получена формула для F без вычисления производных:

Утверждение 3.1. В Алгоритме 3.1 матрица Фока F может быть записана в следующем виде без вычисления производных

$$F = \int_{\mathbb{R}^3} (\widetilde{\Phi}^T V(\Phi^{(k+1)})\widetilde{\Phi} - \widetilde{\Phi}^T V(\Phi^{(k)})\Phi^{(k)}L^{-T} + L^{-1}\widehat{\Phi}^T\widehat{\Phi}\Lambda^{(k)}L^{-T}) d\mathbf{r}, \quad (14)$$

где интеграл берется по всем элементам матрицы.

В разделе 3.3 рассматривается дискретизация уравнений. Уравнения дискретизируются на равномерной $n \times n \times n$ прямоугольной сетке, \mathbb{R}^3 заменяется областью $\Omega = [-L, L]^3$, так как орбитали затухают экспоненциально на

¹⁵Multiresolution quantum chemistry: Basic theory and initial applications / R. J. Harrison [и др.] // J. Chem. Phys. 2004. Т. 121, № 23. С. 11587–11598.

бесконечности. Свертки дискретизируются с помощью галеркинського метода и кусочно-постоянных базисных функций χ_i , $i \in \mathcal{I} \equiv \{0, \dots, n-1\}^3$:

$$\int_{\mathbb{R}^3} \frac{f(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' \approx \sum_{j \in \mathcal{I}} \mathcal{F}_j \mathcal{Q}_{i-j}, \quad \mathcal{Q}_{i-j} = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\chi_i \chi_j}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad \mathcal{F}_i = \int_{\mathbb{R}^3} f \chi_i d\mathbf{r}. \quad (15)$$

Вычисление дискретных сверток из последнего выражения является наиболее затратной частью при вычислении потенциала V на каждой итерации. Для вычисления свертки предлагается новый алгоритм, описанный в главе 4. Нелинейная функция $V_{xc}(\rho)$ вычисляется с помощью метода крестовой аппроксимации. Уравнение Пуассона решается с помощью метода Фурье и метода крестовой аппроксимации для обращения массива из собственных значений дискретного оператора Лапласа. Остальные операции выполняются с помощью явных формул для арифметики в тензорных форматах. В разделе 3.5 приведена оценка сложности одной итерации метода: $N^2 \cdot \mathcal{O}(r^4 + nr^2 + rn \log n)$, где r обозначает максимальный таккеровский ранг орбиталей. На рассматриваемых системах метод сходился приблизительно за 30 итераций.

В разделе 3.6 приведены результаты численных расчетов. Проведено сравнение с базисным подходом, который является стандартным и заключается в приближении решения на подпространстве базисных функций с глобальным носителем. Недостаток базисного подхода заключается в наличии *ошибки набора базисных функций*, контролирование которой является нетривиальной задачей. Предлагаемый в диссертации метод является полностью сеточным, благодаря чему становится возможным контроль ошибки и экстраполяция на последовательности сеток. Отметим, что тензорный подход был использован для решения уравнений ХФ в серии работ Хоромского и Хоромской. Однако в этих работах также используется промежуточный набор глобальных базисных функций.

Результаты численных экспериментов показывают, что предлагаемый метод с использованием экстраполяции является более точным, чем базисный подход с использованием высокоточных базисных наборов. Отметим, что для расчета атомов предлагаемый подход является более медленным. Для молекул время расчета сравнимо с базисным подходом. Отдельно стоит отметить кластеры из регулярно расположенных атомов. В этом случае предлагаемый метод выигрывает у базисного подхода и по скорости, и по точности в силу медленного роста ранга орбиталей с размером кластера.

В **четвертой главе** приводится cross-conv алгоритм приближенного вычисления многомерной свертки (15). Для простоты рассмотрим периодическую свертку двух многомерных массивов \mathcal{F} и \mathcal{G} . Известно, что периодическая свертка может быть представлена в виде нескольких преобразований Фурье и одного поэлементного произведения в “частотной области”:

$$\mathcal{F} * \mathcal{G} = \mathcal{F}^{-1} (\mathcal{F}(\mathcal{F}) \circ \mathcal{F}(\mathcal{G})),$$

где $*$ обозначает свертку, \circ — поэлементное произведение, а \mathcal{F} — многомерное преобразование Фурье. Отметим, что в непериодическом случае выполняется аналогичная формула, но для специально заполненных массивов больших размеров. Можно показать, что многомерное преобразование Фурье не меняет тензорного ранга массива и может быть эффективно выполнено в тензорных форматах. Однако вычисление поэлементного произведения имеет сильную зависимость от ранга.

Предлагаемый cross-conv алгоритм описан в разделе 4.4. Его идея заключается в том, чтобы интерполировать поэлементное произведение с помощью *метода крестовой аппроксимации*, новая эффективная версия которого для формата Таккера в трехмерном случае предлагается в разделе 4.3. Новый метод крестовой аппроксимации строится на основе быстрых формул пересчета параметров разложения с помощью дополнений по Шуру и имеет сложность $\mathcal{O}(nr^2 + r^4)$. Благодаря использованию этого метода, сложность cross-conv алгоритма составляет $\mathcal{O}(nr^2 + r^4 + rn \log n)$.

Для сравнения, алгоритм, базирующийся на сведениях многомерных сверток к одномерным¹⁶, имеет зависимость от ранга $\mathcal{O}(r^8)$, а алгоритм¹⁷ имеет зависимость от ранга $\mathcal{O}(r^6)$. В работе¹⁸ предложен алгоритм вычисления свертки в так называемом квантизованном ТТ (Quantized TT, QTT) формате. Этот алгоритм имеет сложность $\mathcal{O}(d \log^\alpha n)$ и асимптотически превосходит другие методы. В разделе 4.5 проведено численное сравнение с этим методом и показано, что для практически интересных значений n ($n < 10000$ для точности 10^{-5}) cross-conv алгоритм является более быстрым.

¹⁶Khoromskij B. N. Fast and accurate tensor approximation of multivariate convolution with linear scaling in dimension // J. Comp. Appl. Math. 2010. Т. 234, № 11. С. 3122–3139.

¹⁷Savostyanov D. V., Tyrtshnikov E. E. Approximate multiplication of tensor matrices based on the individual filtering of factors // J. Comp. Math. Math. Phys. 2009. Т. 49, № 10. С. 1662–1677.

¹⁸Kazeev V., Khoromskij B., Tyrtshnikov E. Multilevel Toeplitz Matrices Generated by Tensor-Structured Vectors and Convolution with Logarithmic Complexity // SIAM J. Sci. Comput. 2013. Т. 35, № 3. A1511–A1536.

Основные результаты работы

В диссертации разработаны новые тензорные методы решения многомерных частичных задач на собственные значения.

- Для поиска целевого собственного значения предложено обобщение JD метода при ограничении на тензорный ранг решения. Изучены свойства возникающих в этом методе систем линейных уравнений. Показана сходимость регуляризованного метода. Из полученных уравнений для предложенного JD метода получено обобщение метода обратной итерации в малоранговом случае. Показаны преимущества метода в случае, когда возникающие линейные системы решаются неточно.
- Предложен ALS II метод, базирующийся на ALS подходе и методе обратной итерации. Получен результат о локальной сходимости метода через сходимость ALS итерации для минимизации отношения Рэля. Для ALS оптимизации получена новая теория сходимости, явно показывающая связь с мультипликативным методом Шварца.
- Предложена концепция преобуславливания на многообразии для поиска нескольких собственных значений, примененная к методу LOBPCG. С помощью предложенных итераций с высокой точностью рассчитан спектр молекулы ацетонитрила. Показано, что предлагаемый метод превосходит по точности и экономии памяти существующие аналоги для расчета колебательного спектра.
- Рассмотрена задача на собственные значения с нелинейными операторами, возникающими в уравнениях Хартри-Фока и Кона-Шэма. Предложен метод, базирующийся на разложении Таккера, сложность которого линейно зависит от размера сетки по каждому направлению. Проведен точный расчет ряда атомов, молекул и кластеров. Получены результаты, превышающие по точности подход, использующий глобальные базисные функции. Для кластеров с регулярным расположением атомов/молекул предложенный метод является более быстрым, чем базисный подход.
- Для быстрого вычисления многомерной свертки предлагается cross-conv алгоритм, базирующийся на методе крестовой аппроксимации. Показано, что для размеров сетки, интересных на практике, cross-conv алгоритм превосходит по скорости существующие аналоги.

Публикации автора по теме диссертации

1. Rakhuba M. V., Oseledets I. V. Calculating vibrational spectra of molecules using tensor train decomposition // *The Journal of Chemical Physics*. — 2016. — Т. 145. — №. 12. — С. 124101.
2. Rakhuba M. V., Oseledets I. V. Fast multidimensional convolution in low-rank tensor formats via cross approximation // *SIAM Journal on Scientific Computing*. — 2015. — Т. 37. — №. 2. — С. A565-A582.
3. Rakhuba M. V., Oseledets I. V. Grid-based electronic structure calculations: The tensor decomposition approach // *Journal of Computational Physics*. — 2016. — Т. 312. — С. 19-30.
4. Rakhuba M. V., Oseledets I. V. Jacobi-Davidson method on low-rank matrix manifolds // arXiv preprint arXiv:1703.09096. — 2017.
5. Oseledets I. V., Rakhuba M. V. and Uschmajew A. Alternating least squares as moving subspace correction // arXiv preprint: arXiv:1709.07286. — 2017.
6. Rakhuba M. V. Block eigensolvers on low-rank tensor manifolds // *Proceedings of the 88-th Annual Meeting of the International Association of Applied Mathematics and Mechanics, Weimar, 2017*.
7. Рахуба М. В. Малоранговые разложения многомерных массивов и их приложение в расчете колебательного спектра молекул // Сб. тез. конф. “Ломоносов 2017”, Москва, 2017.
8. Рахуба М. В., Оселедец И.В. Методы решения многомерных задач на собственные значения на малоранговых тензорных многообразиях и их приложения в задачах квантовой химии // Тезисы 59-й научной конференции МФТИ, 2016.
9. Rakhuba M. V. Making block eigensolvers really work in higher dimensions // *Proceedings of the 20-th Conference of the International Linear Algebra Society (ILAS), Leuven, 2016*.
10. Рахуба М. В., Оселедец И.В. Быстрый алгоритм вычисления многомерной свертки на основе тензорных аппроксимаций и его применение для расчета электронной структуры молекул // Труды 56-й научной конференции МФТИ, 2013.