

На правах рукописи



Сушникова Дарья Алексеевна

Методы факторизации и решения линейных систем с
блочно-малоранговыми матрицами

05.13.18 — Математическое моделирование, численные методы и комплексы
программ

АВТОРЕФЕРАТ

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

Москва 2017

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте вычислительной математики Российской академии наук.

Научный руководитель:

доцент Сколковского института науки и технологий,
д.ф.-м.н. Оселедец Иван Валерьевич.

Официальные оппоненты:

Ильин Валерий Павлович, доктор физико-математических наук, профессор. Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, главный научный сотрудник.

Гасников Александр Владимирович, доктор физико-математических наук, доцент. Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (государственный университет)», ведущий научный сотрудник.

Ведущая организация:

Федеральное государственное учреждение «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук».

Защита состоится «25» октября 2017 г. в 14:00 на заседании диссертационного совета Д 002.045.01 в Федеральном государственном бюджетном учреждении науки Институте вычислительной математики Российской академии наук, расположенном по адресу: 119333, г. Москва, ул. Губкина, д. 8.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Института вычислительной математики Российской академии наук.

Автореферат разослан «__» _____ 2017 г.

Ученый секретарь

диссертационного совета Д 002.045.01

доктор физико-математических наук

Г. А. Бочаров

Общая характеристика работы

Актуальность темы исследования.

Ряд плотных матриц, возникающих в задачах электростатики, задаче многих тел, а также матрицы, полученные при дискретизации сингулярных и гиперсингулярных интегральных уравнений, обладают блочно-малоранговой структурой. Под блочно-малоранговой структурой мы понимаем структуру блочной матрицы с $M \times M$ блоками, при которой все блоки матрицы кроме $\mathcal{O}(M)$ имеют приближённо малый ранг. За данной структурой лежит следующий общий физический смысл: строки и столбцы таких матриц ассоциированы с некоторыми элементами в пространстве, задана некоторая функция взаимодействия этих элементов, если функция взаимодействия асимптотически гладкая, то взаимодействие разнесённых в пространстве групп элементов можно приблизить малым числом параметров¹ (критерий разделения). Таким образом, блоки, ассоциированные с хорошо разделёнными в пространстве группами элементов, обладают приближенным малым рангом. Другой известный пример блочно-малоранговых матриц связан с матрицами, полученными при дискретизации дифференциальных уравнений. Известно^{2,3}, что если матрица A получена при конечно-элементной дискретизации дифференциального уравнения, удовлетворяющего некоторым ограничениям^{2,4,3}, то обратная к ней приближается блочно-малоранговой матрицей.

Блочно-малоранговые матрицы представляют из себя приближение с хорошей точностью плотных матриц в малопараметрическом формате. Блоки малого ранга представляются в виде произведения матриц меньшего размера. Это позволяет значительно экономить машинную память, так, в отличие от плотной матрицы, которая требует n^2 ячеек памяти, малопараметрическое представление, требует $\mathcal{O}(n(\log^\alpha n)(\log^\beta \varepsilon^{-1}))$ ячеек памяти (в зависимости от типа малопараметрического представления). Данная особенность позволяет хранить при-

¹Tyrtysnikov E. E. Mosaic-skeletonapproximations // *Calcolo*. —1996. —Vol.33, no. 1. — P. 47–57.

²Bebendorf Mario. Why finite element discretizations can be factored by triangular hierarchical matrices // *SIAM J. Numer. Anal.* — 2007. — Vol. 45, no. 4. — P. 1472– 1494.

³Börm Steffen. Approximation of solution operators of elliptic partial differential equations by-and-matrices // *Numerische Mathematik*. — 2010. — Vol. 115, no. 2. — P. 165–193.

⁴Bebendorf Mario, Hackbusch Wolfgang. Existence of H-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞ -coefficients // *Numer. Math.* — 2003. — Vol. 95, no. 1. — P. 1–28.

ближение к плотной матрице, используя память, требуемую для хранения разреженной матрицы такого же размера. Другой характерной особенностью малопараметрического представления является быстрая процедура умножения такой матрицы на вектор ($O(n \log n)$ или $O(n)$ операций в зависимости от вида представления).

Быстрая процедура умножения матрицы на вектор позволяет эффективно применять к решению систем с малопараметрическими матрицами итерационные методы. Однако в случае плохой обусловленности, когда требуется решить систему прямым методом или приближенно, для построения предобуславливателя матрицы в малопараметрическом представлении сталкиваются со значительными трудностями. Одной из основных трудностей является сложный формат хранения малопараметрических матриц: малопараметрические форматы, такие как \mathcal{H} ^{5,6}, \mathcal{H}^2 ^{7,8}, HSS^{9,10} и т.д. рассчитаны на быстрое умножение матрицы на вектор, однако быстрое исключение строк и столбцов для таких матриц является трудоёмкой задачей. Данная работа посвящена методам прямого решения и приближенной факторизации систем с блочно-малоранговыми матрицами в малопараметрическом формате.

Целью диссертационной работы являлась разработка новых приближенных факторизаций блочно-малоранговых матриц и методов их построения.

Научная новизна. Предложен новый метод приближенной факторизации разреженных матриц (метод компрессии и исключения), также предложены два метода разреженной факторизации \mathcal{H}^2 матриц.

Практическая ценность. Предложенный в работе метод приближенной факторизации разреженных матриц может быть использован для приближенного решения, предобуславливания и приближенного вычисления определителя

⁵Hackbusch Wolfgang. A sparse matrix arithmetic based on \mathcal{H} -matrices. part i: Introduction to \mathcal{H} -matrices // Computing. — 1999. — Vol. 62, no. 2. — P. 89–108.

⁶Tyrtushnikov E. E. Mosaic-skeleton approximations // Calcolo. — 1996. — Vol. 33, no. 1. — P. 47–57.

⁷Börm Steffen. Efficient numerical methods for non-local operators: \mathcal{H}^2 -matrix compression, algorithms and analysis. — European Mathematical Society, 2010. — Vol. 14.

⁸Hackbusch W., Khoromskij B.N., Sauter S. On \mathcal{H}^2 -Matrices // H.-J. Bungartz, et al. (eds.), Lectures on Applied Mathematics. — Springer-Verlag, Berlin Heidelberg, 2000. — P. 9–30.

⁹Solov'yev, Sergey. Multifrontal Hierarchically Solver for 3D Discretized Elliptic Equations // International Conference on Finite Difference Methods. — Springer — 2014. — P. 371–378

¹⁰Jianlin Xia, Shivkumar Chandrasekaran, Ming Gu, Xiaoye S Li. Fast algorithms for hierarchically semiseparable matrices // Numer. Linear Algebr. — 2010. — Vol. 17, no. 6. — P. 953–976.

разреженных положительно определённых матриц, в частности, полученных при дискретизации дифференциальных уравнений.

Методы приближенной факторизации блочно-малоранговых матриц, в свою очередь, могут быть применены для приближенного решения и преобуславливания систем с плотными матрицами и для приближенного вычисления определителя плотных матриц в задачах электростатики, аэро- и гидродинамики, а также в прикладной статистике. Данные методы показали свою эффективность в сравнении методами HODLR^{11,12} и IFMM^{13,14}.

На защиту выносятся следующие результаты и положения: Основной результат работы состоит в том, что предложены новые приближенные факторизации блочно-малоранговых матриц и методы их построения, реализован программный комплекс, который применен к нескольким задачам математического моделирования. В частности:

1. Предложена факторизация разреженной симметричной положительно определённой матрицы в виде произведения матриц перестановки, блочно-диагональных унитарных и разреженных нижне-треугольных факторов (метод компрессии и исключения, compress and eliminate method, CE), на основе которой предложен прямой метод решения систем и метод построения преобуславливателя.
2. Предложены два типа разреженной факторизации \mathcal{H}^2 матриц: «расширенная» и «не-расширенная», которые позволяют ускорить решение линейных систем с \mathcal{H}^2 матрицами, в сравнении с итерационным методом BiCGStab¹⁵ и прямыми методами HODLR^{11,12} и IFMM^{13,14}.
3. Разработан комплекс программ реализующих представленные алгоритмы. Для факторизации CE проведено тестирование на системах, полу-

¹¹Ambikasaran Sivaram, Darve Eric. An $O(N \log N)$ Fast Direct Solver for Partial Hierarchically Semi-Separable Matrices // J. Sci. Comput. — 2013. — Vol. 57, no. 3. — P. 477–501.

¹²Ambikasaran Sivaramand, Foreman-Mackey Daniel, Greengard Leslie. A fast direct methods for gaussian processes // arXiv preprint arXiv:1403.6015. — 2014.

¹³Ambikasaran Sivaram, Darve E. The Inverse Fast Multipole Method // arXiv preprint arXiv:1309.1773. — 2014.

¹⁴Coulier Pieter, Pouransari Hadi, Darve Eric. The inverse fast multipole method: using a fast approximate direct solver as a preconditioner for dense linear systems // arXiv preprint arXiv:1508.01835. — 2015.

¹⁵Andreas Frommer, Volker Hannemann, Bertold Nöckel et al. Accelerating Wilson fermion matrix inversions by means of the stabilized biconjugate gradient algorithm // Int. J. Mod. Phys. C. — 1994. — Vol. 5, no. 06. — P. 1073–1088.

ченных при конечно-разностной дискретизации стационарного уравнения диффузии, а также системах, полученных при конечно-элементной дискретизации уравнения Пуассона и уравнения упругости. Проведено сравнение реализации метода CE с прямыми методами CHOLMOD^{16,17} и UMFPACK¹⁸, а также итерационным методом BiCGStab с предобуславливателями ILU0¹⁹, ILUt²⁰ и ILU2²¹. Для метода расширенной разреженной факторизации проведено тестирование на задачах электростатики и гидромеханики, в ходе которого метод показал свою эффективность по времени в сравнении с непредобусловленным методом BiCGStab. Для не-расширенной разреженной факторизации проводилось тестирование на задаче моделирования гауссовских процессов для задачи регрессии. Метод показал свою эффективность по времени в сравнении с методами NODLR и IFMM.

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

- 56-я научная конференция МФТИ, 2013, Москва
- Научная конференция “Ломоносовские чтения”, 2013, Москва
- Fast Direct Solvers for Elliptic PDEs, Dartmouth College, 2014, USA
- Workshop: Low-rank Optimization and Applications, Hausdorff Center for Mathematics, Bonn, 2015, Germany

¹⁶Yanqing Chen, Timothy A Davis, William W Hager, Sivasankaran Rajamanickam Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate // ACM T. Math. Software. — 2008. — Vol. 35, no. 3. — P. 22.

¹⁷Davis Timothy A, Hager William W. Dynamic supernodes in sparse Cholesky update/downdate and triangular solves // ACM T. Math. Software. — 2009. — Vol. 35, no. 4. — P. 27.

¹⁸Davis T. A. Algorithm 832: UMFPACK V4. 3—an unsymmetric-pattern multifrontal method // ACM T. Math. Software. — 2004. — T. 30. — no. 2. — P. 196-199.

¹⁹Saad Yousef. Iterative methods for sparse linear systems. — Siam, 2003.

²⁰Saad Yousef ILUT: A dual threshold incomplete LU factorization // Numer. Linear Algebr. — 1994. — Vol. 1, no. 4. — P. 387-402.

²¹Kaporin I. E. High quality preconditioning of a general symmetric positive definite matrix based on its UTU + UTR + RTU-decomposition // Numer. Linear Algebr. — 1998. — T. 5. — no. 6. — P. 483-509.

- 4th International Conference on Matrix Methods in Mathematics and Applications, Skolkovo Institute of Science and Technology, 2015, Moscow
- Scalable Hierarchical Algorithms for eXtreme Computing workshop, King Abdullah University of Science and Technology, 2016, Saudi Arabia
- Seminar of Extreme Computing Research Center, King Abdullah University of Science and Technology, 2016, Saudi Arabia
- 59-я научная конференция МФТИ, 2016, Москва
- Семинар имени К.И. Бабенко, ИПМ РАН, 2016, Москва
- Workshop on Fast Direct Solvers, Purdue CCAM, 2016, USA
- Объединённый семинар ИВМиМГ СО РАН и кафедры вычислительной математики НГУ, 2017, Новосибирск
- Сейсмический семинар ИНГиГ СО РАН, 2017, Новосибирск
- Научная конференция “Ломоносов”, 2017, Москва
- Семинар лаборатории ”Математическое моделирование нелинейных процессов в газовых средах”, МФТИ, 2017, Москва

Личный вклад. Результаты, описанные в главе 2, опубликованы в работе [4], эта работа опубликована в соавторстве с И. В. Оселедцем. В работе [4] основная идея метода разработана Д. А. Сушниковой. Также автору диссертации принадлежит программа ЭВМ и численные эксперименты. Оселедцу И.В. принадлежит постановка задачи.

Результаты, описанные в главе 3, опубликованы в работах [1] и [5], эти работы опубликованы в соавторстве с И. В. Оселедцем. В работах [1] и [5] Д. А. Сушниковой принадлежит основная идея метода, программа ЭВМ и численные эксперименты, Оселедцу И.В. принадлежит постановка задачи.

Результаты, описанные в главе 4, опубликованы в работе [2], эта работа опубликована автором самостоятельно.

Объём и структура работы. Работа состоит из 124 страниц и содержит введение, 5 глав, заключение и список литературы. Список литературы включает 89 пунктов.

Основное содержание работы

Во **введении** обоснована актуальность темы диссертационной работы и приведена краткая историческая справка о методах решения линейных систем с блочно-малоранговыми матрицами. Также во введении предложена краткая характеристика диссертации: изложены цель работы, её научная новизна и практическая ценность, сформулированы положения, выносимые на защиту, отмечен личный вклад автора, кроме того, представлен список конференций и научных семинаров, на которых проходила апробация диссертационной работы, и список публикаций автора, содержащих основные результаты.

Первая глава является вводной и содержит обзор блочно-малоранговых методов в приложении к задачам моделирования. Также в ней приведены предварительные сведения об иерархических блочно-малоранговых матрицах. В этой главе вводится ряд ключевых для диссертации понятий, таких как «кластерное дерево», «блочно-малоранговая матрица», «H2 матрица».

Во **второй главе** представлена факторизация разреженных симметричных положительно определённых матриц в виде произведения матриц перестановки, блочно-диагональных унитарных факторов и разреженных нижнетреугольных факторов («Метод компрессии и исключения», Compress and eliminate method, CE), на основе которой предложен прямой метод решения систем и метод построения предобуславливателя. **Раздел 2.1** содержит пошаговое описание метода CE. Основными шагами предложенного алгоритма являются:

1. Построение предварительной перестановки на основе графа разреженности матрицы. (**Раздел 2.1**)
2. Исключение первого уровня матрицы. (**Параграфы 2.1.1 - 2.1.3**)
 - (a) Для каждой блочной строки:
 - i. Малоранговая аппроксимация блоков заполнения.
 - ii. Исключение блочной подстроки с нулевым заполнением.
 - (b) Перестановка и укрупнение блоков остатка
3. Аналогичное исключение K уровней (**Параграф 2.1.4**)
4. Факторизация остатка (**Параграф 2.1.4**)

В параграфе 2.1.4 приведён псевдокод SE алгоритма. Алгоритм компрессии и исключения приводит к приближенной факторизации

$$A \approx QLL^T Q^T, \quad (1)$$

где Q — это унитарная матрица, равная произведению блочно-диагональных матриц и матриц перестановки, L — это блочно-разреженная нижне-треугольная матрица.

В разделе 2.2 проведена оценка сложности SE метода. Параграф 2.2.1 содержит определение блочного шаблона разреженности.

Определение 2.2.1. Для матрицы A с M блочными строками, M блочными столбцами и размером блока $B \times B$ определим $\mathbf{bsp}(A)_{B \times B}^{M \times M}$ (блочный шаблон разреженности) как функцию

$$F_A(i, j) \rightarrow \{0, 1\}, \forall i, j = \overline{1 \dots M},$$

которая возвращает 1, когда блок A_{ij} ненулевой, и 0 в противном случае.

Определение 2.2.2. Через

$$\#\mathbf{bsp}(A)_{B \times B}^{M \times M},$$

обозначим число ненулевых блоков $(B \times B)$ матрицы A .

Также в данном разделе сформулирована и доказана теорема о сложности SE алгоритма в зависимости от разреженности фактора L и блочного шаблона разреженности матрицы A .

Теорема 2.2.3. Алгоритм компрессии и исключения (SE) приводит к приближенной факторизации

$$A \approx QLL^T Q^T,$$

где Q — это унитарная матрица равная произведению блочно-диагональных матриц и матриц перестановки

$$Q = \left(\prod_{j=1}^K \check{P}_j \check{Q}_j \right) P,$$

L — это разреженная блочно нижне-треугольная матрица

$$L = \begin{bmatrix} \begin{bmatrix} L_1 \end{bmatrix} & 0 & & 0 \\ & \begin{bmatrix} L_2 \end{bmatrix} & \dots & \vdots \\ & & & 0 \\ & & & \begin{bmatrix} L_{K+1} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} L_1 & \check{L}_2 & \dots & \check{L}_{K+1} \end{bmatrix},$$

в которой блок L_i содержит

$$\#L_j \leq 2\#\mathbf{bsp}(A_{j-1})_{rJ \times rJ}^{M/J^{j-1} \times M/J^{j-1}}$$

ненулевых блоков, и общее число ненулевых блоков в факторе L не превосходит

$$\#L \leq 2\#\mathbf{bsp}(A_0)_{B \times B}^{M \times M} + 2 \left(\sum_{j=1}^{K-1} \#\mathbf{bsp}(A_0^{2^j})_{B^{J^j} \times B^{J^j}}^{M/J^j \times M/J^j} \right) + \frac{1}{2}(M/J^K)^2 r^2. \quad (2)$$

K — это число уровней в СЕ алгоритме, B — это размер блока, $A_0 = PAP^\top$. Затраты по времени для приближенной факторизации СЕ составляют $\mathcal{O}(B^3(\#L_{K+1} - \#L_1) + B^2\#L)$. Затраты по памяти составляют $\mathcal{O}(B(B - r)\#L + NB)$.

Данная теорема показывает важность изначальной перестановки применённой к матрице A (упомянутой в **разделе 2.1**) для увеличения разреженности фактора L и уменьшения вычислительной сложности метода СЕ. Также в этом разделе приводится теорема о вычислительной сложности решения системы с матрицей разложенной в СЕ.

Теорема 2.2.4. Система $Ax = b$ с вычисленной СЕ факторизацией матрицы A ($A = Q^\top LL^\top Q$) может быть решена за $\mathcal{O}((B - r)B\#L + NB)$ операций.

В **параграфе 2.2.2** приводится метод построения предварительной перестановки P на основе графа \mathcal{G}_{A_0} разреженности матрицы A . В данном разделе формулируется и доказывается утверждение о связи графов разреженности $\mathcal{G}_{A_0}, \mathcal{G}_{A_1}, \dots, \mathcal{G}_{A_K}$ матриц A, A_1, \dots, A_K и числа ненулевых блоков в матрице L .

Утверждение 2.2.5. Число $\#L$ ненулевых блоков в факторе L не превосходит общее число рёбер в графах $\mathcal{G}_{A_0}, \mathcal{G}_{A_1}, \dots, \mathcal{G}_{A_K}$:

$$\#L \leq \sum_{i=0}^K \mathbf{Edge_num}(\mathcal{G}_{A_i}).$$

Также формулируется и доказывается утверждение об линейности алгоритма SE для класса матриц, графы разреженности которых порождены тензорными сетками.

Утверждение 2.2.6. Пусть граф \mathcal{G}_{A_0} определен тензорной сеткой в \mathbb{R}^2 , (точки сетки соответствуют вершинам графа \mathcal{G}_{A_0} , ребра сетки совпадают с ребрами графа \mathcal{G}_{A_0}). Если процедура укрупнения объединяет ближайшие вершины и каждая супер-вершине имеет минимум две вершины в каждом направлении, тогда каждый граф $\mathcal{G}_{A_0}, \mathcal{G}_{A_1}, \dots, \mathcal{G}_{A_K}$ имеет $\mathcal{O}(N)$ вершин.

Из данного утверждения выводится ряд важных следствий.

Следствие 2.2.6. Утверждение 2.2.6 также верно для $\mathbb{R}^d, d > 2$.

Следствие 2.2.7. Пусть матрица A_{trg} имеет граф порожденный тензорной сеткой в \mathbb{R}^d . Факторизация SE матрицы A_{trg} требует $\mathcal{O}(B^3 N + B^2 N K)$ операций и $\mathcal{O}(B(B-r)NK)$ ячеек памяти.

Данное следствие доказывает линейность метода SE для описанного выше класса матриц. Также в этом разделе приводится гипотеза 2.2.8 о расширении утверждения о линейности алгоритма SE на матрицы, граф разреженности которых является графом k -ближайших соседей.

Третья глава содержит два метода разреженной факторизации малопараметрических матриц: «расширенный» и «не-расширенный». В разделе 3.1 описан расширенный метод разреженной факторизации \mathcal{H}^2 матриц. **Параграф 3.1.1** содержит краткое определение и ряд свойств \mathcal{H}^2 матриц, которые будут в дальнейшем использоваться в данной главе. Также в данном разделе описан алгоритм произведения \mathcal{H}^2 матрицы на вектор, являющийся ключевым для расширенной разреженной факторизации.

В параграфе 3.1.2 приводится основная идея разреженной факторизации, а именно, введение дополнительных промежуточных переменных и запись с их

помощью произведения \mathcal{H}^2 матрицы на вектор в матричном виде. После чего такое преобразование, как перенос неизвестных переменных их левой части уравнения в правую, приводит к разреженно-расширенной (sparse extended, SE) матрице.

Параграф 3.1.3 содержит теорему об условии невырожденности SE матрицы и её доказательство.

Теорема 3.1.1. *Если матрица $A \in \mathbb{R}^{N \times N}$ это невырожденная \mathcal{H}^2 матрица, тогда $H = \text{SE}(A) \in \mathbb{R}^{N_H \times N_H}$ также невырожденная, кроме того*

$$N_H < (2k + 1)N,$$

где k — это максимум из количества уровней в кластерных деревьях \mathcal{T}_c и \mathcal{T}_r .

Далее в **параграфе 3.1.4** приводятся три метода решения систем с SE матрицей.

1. (Прямой метод решения) Применим прямой разреженный решатель к $\text{SE}(A)$ и при заданном y вычислим x .
2. (Итерационный метод решения расширенной системы с предобуславливателем) Сконструируем предобуславливатель для итерационного решателя расширенной системы, основанный на блочной структуре матрицы $\text{SE}(A)$.
3. (Итерационный метод решения для системы с исходной матрицей A , используя матрицу $\text{SE}(A)$ в качестве предобуславливателя) Неточное прямое решение системы с SE матрицей может быть использовано в качестве предобуславливателя для итерационного метода решения исходной системы

$$Ax = y.$$

Для получения неточного решения расширенной системы используется несколько шагов итерационного метода. Также на основе расширенной системы можно построить предобуславливатель ILUt.

Раздел 3.2 содержит алгоритм не-расширенной разреженной факторизации \mathcal{H}^2 матрицы. Основная идея не-расширенной разреженной факторизации

(**параграф 3.2.1**) заключается в мультиуровневой процедуре компрессии дальних блоков. Процедура компрессии блоков при помощи сингулярного разложения была ранее описана в алгоритме компрессии и исключения (**раздел 2.1**). Результатом алгоритма будет приближенная разреженная факторизация

$$A = USV^T, \quad (3)$$

где S — это разреженная матрица, которая имеет такой же размер, как матрица A , U и V — это унитарные матрицы, являющиеся произведением матриц перестановки и блочно-диагональных унитарных матриц.

В **параграфе 3.2.2** доказывается разреженность матрицы S .

Утверждение 3.2.4. Если матрица A имеет \mathcal{H}^2 структуру, если алгоритм компрессии имеет L уровней, если размер блока на каждом уровне $l \in \overline{0, \dots, L}$ равен B , если A имеет ближнюю матрицу первого уровня C и если дальние блочные столбцы и строки всегда сжимаются в два раза $r = B/2$, тогда алгоритм компрессии для матрицы A приводит к факторизации

$$A = U^T S V,$$

где U и V — это унитарные матрицы равные произведению блочно-диагональных унитарных матриц компрессии и матриц перестановки

$$U = \left(\prod_{j=0}^K U_j P_j^T \right), \quad V = \left(\prod_{j=0}^K V_j P_j^T \right),$$

а матрица S — это разреженная матрица, которая имеет

$$\#S = \left(4L + 6\left(\frac{1}{2^L} - 1\right) \right) \#bsp(C)_{B \times B}^{M \times M}$$

ненулевых блоков размера $(r \times r)$.

Далее в **параграфе 3.2.3** приводится алгоритм построения разреженной факторизации (3) в том случае, если матрица уже аппроксимирована в \mathcal{H}^2 формате. Тогда факторы S , U и V могут быть быстро построены из параметров \mathcal{H}^2 матрицы.

В **главе четыре** приводится комплекс программ, реализующих методы, описанные в главах 2 и 3.

Раздел 4.1 содержит описание программной реализации метода CE. В **параграфе 4.1.1** приводится пример программы с использованием пакета `ce_solver`, которая реализует $CE(r = 2)$ факторизацию трехмерной матрицы Лапласа. Далее в **параграфе 4.1.2** представлено решение системы, полученной при дискретизации на равномерной кубической сетке трехмерного уравнения диффузии

$$\begin{aligned} \operatorname{div} (k(x) \operatorname{grad} u(x)) &= f(x), \quad x \in \Omega, \\ u|_{\delta\Omega} &= 0, \end{aligned} \tag{4}$$

где $\Omega = [0, 1]^3$, тензор диффузии $k(x) = \operatorname{diag}(x_1^2 + 0.5, x_2^2 + 0.5, x_3^2 + 0.5)$, правая часть $f(x) = 1$. Метод CE рассматривается в качестве прямого метода решения и в качестве предобуславливателя для итерационного метода MINRES.

Далее приводится относительная точность η , которая может быть получена при приближенном прямом решении системы, матрица которой факторизована с помощью $CE(\varepsilon)$, также приводятся затраты по времени и по памяти на $CE(\varepsilon)$ факторизацию. Основная проблема прямого метода решения $CE(\varepsilon)$ заключается в том, что он требует слишком много памяти. Однако $CE(\varepsilon)$ с достаточно большим ε работает быстро и требует немного памяти. Дальше факторизации $CE(\varepsilon)$ и $CE(r)$ рассматриваются в качестве предобуславливателей.

В этом же параграфе представлена сходимость метода MINRES по времени и по итерациям с $CE(\varepsilon)$, $CE(r)$, ILUt предобуславливателями, а также сходимость метода MINRES без предобуславливателя. Отмечается, что число итераций в MINRES без предобуславливателя, также как и число итераций с предобуславливателем ILUt, растёт вместе с ростом сетки, так как обусловленность матрицы ухудшается. $CE(\varepsilon)$ и $CE(r)$ — это хорошие предобуславливатели, в том смысле, что для предобуславливателя $CE(\varepsilon)$ с фиксированным параметром ε число итераций не растёт; для предобуславливателя $CE(r)$ число итераций растёт медленно. Также приводится сравнение общего времени решения для различных параметров ε и r и сравнение памяти и времени, требуемых для факторизации $CE(\varepsilon)$ или $CE(r)$ при различных значениях параметров ε и r .

Далее в этом параграфе приводится сравнение по времени и по памяти следующих методов решения линейных систем с разреженными симметричными положительно определёнными матрицами: $CE(\varepsilon)$ и $CE(r)$ в качестве предобуславливателей для метода MINRES, MINRES с предобуславливателем ILUt и

CHOLMOD. Метод решения с предобуславливателем $CE(r)$ начинает выигрывать по памяти у метода CHOLMOD на $N \approx 5000$ и начинает работать быстрее на $N \approx 10000$. Для максимального N , которое допускает кластер, система решается примерно в 10, быстрее чем CHOLMOD, и требует примерно в 10 меньше памяти.

В параграфе 4.1.3 рассматривается конечно-элементная дискретизация уравнения Пуассона и уравнения упругой деформации. Сначала решается двумерное уравнение Пуассона:

$$\begin{aligned} -\nabla^2 u(x) &= f(x), & x \in \Omega, \\ u(x) &= u_D(x), & x \in \delta\Omega. \end{aligned} \quad (5)$$

где $\Omega = K((0, 0), 1) - K((0.5, 0), 0.5)$, $K(c, r)$ — это круг с центром c и радиусом r . Уравнение Пуассона (5) дискретизируется при помощи метода конечных элементов. Полученная система решается методом BiCGStab из пакета PyAMG с использованием предобуславливателя $CE(r = 2)$, предобуславливателя ILUt с пороговым значением $\tau = 10^{-3}$, предобуславливателя ILU0, предобуславливателя ILU2, и прямого метода решения CHOLMOD. Производится сравнение предобуславливателей $CE(r = 2)$, ILUt, ILU2 и ILU0 для итерационного метода решения BiCGStab. Несмотря на то, что построение предобуславливателя $CE(r = 2)$ требует больше времени, чем построение двух других предобуславливателей, этот предобуславливатель значительно сокращает количество итераций метода BiCGStab.

Приводится сравнение методов решения, приведённых выше, для различных требуемых точностей решения $\epsilon = 10^{-3}, 10^{-6}, 10^{-10}$.

Далее рассматривается трехмерное уравнение упругой деформации для небольших смещений

$$\begin{aligned} -\nabla \sigma &= f, & \in \Omega, \\ \sigma &= \lambda \text{tr}(\varepsilon)I + 2\mu\varepsilon, & x \in \delta\Omega, \\ \varepsilon &= \frac{1}{2} (\nabla u + (\nabla u)^T), \end{aligned} \quad (6)$$

где σ — это тензор напряжений, f — это приложенная сила на единицу объема, λ и μ — это модули упругости Ламе для материала в области Ω ,

$$\Omega = K((0, 0, 0), 1) - K((0.5, 0, 0), 0.48) - K((1, 1, 0), 1) - K((1, -1, 0), 1),$$

$K(c, r)$ — это шар с центром c и радиусом r . Дифференциальное уравнение (6) дискретизируется при помощи метода конечных элементов. Полученная система решается методом BiCGStab с использованием предобуславливателя CE($r = 2$), предобуславливателя ILUt с пороговым значением $\tau = 10^{-3}$, предобуславливателя ILU0, предобуславливателя ILU2 и прямого метода решения CHOLMOD. Производится сравнение предобуславливателей CE($r = 2$), ILUt, ILU2 и ILU0 для итерационного метода решения BiCGStab. Далее приводится сравнение методов решения, приведённых выше, для различных требуемых точностей решения $\epsilon = 10^{-3}, 10^{-6}, 10^{-10}$.

В **параграфе 4.1.4** приводится сравнение скорости решения и затрат по памяти для четырёх тестовых больших симметричных положительно определённых разреженных матриц. Две из них — это матрицы максимального размера из предыдущих примеров, две взяты из коллекций Florida Sparse Matrix Collection и The SuiteSparse Matrix Collection соответственно. Приводятся результаты сравнения следующих методов: BiCGStab с предобуславливателями CE($r = 2$), ILUt, ILU2 и ILU0, а также прямые методы CHOLMOD и UMFPACK. Тестирование проводится для решения систем с точностями $\epsilon = 10^{-3}, 10^{-6}, 10^{-10}$.

Раздел 4.2 посвящен разреженной факторизации блочно-малоранговых матриц. Программная реализация алгоритма расширенной разреженной факторизации приведена в **параграфе 4.2.1**. Предложенные методы сравниваются на примере электростатической задачи

$$\int_{\Omega} \frac{q(y)}{\|x - y\|} dy = f(x), \quad x \in \Omega, \quad (7)$$

где $\Omega = [0, 1]^2$. Функция $f(x)$ задана, требуется найти $q(y)$. Уравнение (7) дискретизировано с использованием метода коллокации с кусочно-постоянными базисными функциями на треугольной сетке Ω_N . Приведено сравнение предложенных методов для различных N . Для Методов 2 и 3 время решения складывается из времени построения предобуславливателя и времени GMRES итераций). Далее представлены затраты по памяти приведённых алгоритмов для различных N . Метод 3 с предобуславливателем SVD-SE оказывается самым эффективным для данной задачи.

Далее рассмотрено гиперсингулярное интегральное уравнение для задачи обтекания плоскости

$$\int_{\Omega} \frac{q(y)}{\|x - y\|^3} dy = f(x), \quad (8)$$

где Ω — это плоская область. Уравнение дискретизировано при помощи метода коллокации с кусочно-постоянными функциями на треугольной сетке, матрица аппроксимирована в \mathcal{H}^2 формате с помощью пакета h2tools. Метод 3 с предобуславливателем SVD-SE оказывается самым эффективным для данной задачи.

Параграф 4.2.2 содержит описание программной реализации метода нерасширенной разреженной факторизации \mathcal{H}^2 матриц. **Листинг 2** содержит пример простой программы, которая реализует данный метод.

Для двух типов \mathcal{H}^2 матриц показывается разреженность фактора S , также показывается линейность построения разреженной факторизации.

Далее рассматривается матрица

$$A_{ij} = 2\delta_{ij} + \exp(-|r_i - r_j|^2),$$

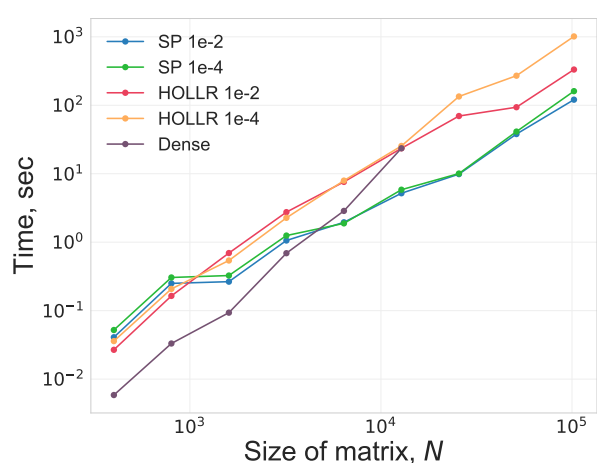
где $r_i \in \mathbb{R}^3$ — это положение i -го элемента. Для данной матрицы требуется построить её факторизацию Холецкого. Сравняются два подхода: метод HODLR и метод аппроксимации в \mathcal{H}^2 формате, разреженной факторизации и факторизации Холецкого разреженной матрицы. Приводится сравнение по времени подхода спарсификации \mathcal{H}^2 с HODLR методом решения в 3D и сравнение общего времени решения.

Также разреженная факторизация рассматривается в качестве предобуславливателя для разреженной матрицы. Для тестов выбрана плотная матрица с обусловленностью, зависящей от параметра. Приводятся результаты работы предобуславливателя для случаев хорошей ($\text{cond}(A) = 10$) и плохой ($\text{cond}(A) = 10^4$) обусловленности матриц и сравнение с результатами IFMM предобуславливателя, которое показывает, что для данной задачи предложенный автором метод предобуславливания систем с \mathcal{H}^2 матрицами оказался более эффективным.

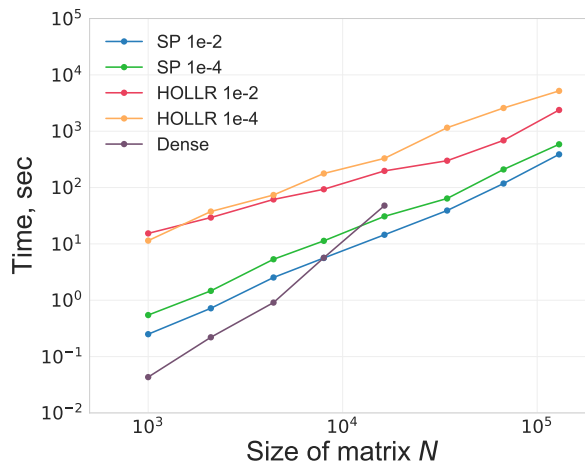
Глава пять содержит приложение разреженной факторизации \mathcal{H}^2 матриц в математической статистике. Рассматривается задача регрессии на основе гауссовских процессов. В **разделе 5.2** на примере одномерной задачи показывается эффективность моделирования коррелированного шума в задаче регрес-

сии при помощи гауссовских процессов. Моделирование выполняется при помощи пакета George.

В **Разделах 5.3 и 5.4** рассматриваются двумерная и трёхмерная задачи регрессии. В процессе моделирования требуется считать логарифм определителя плотной матрицы ковариации K и решать систему с ней. Так как матрица ковариации симметричная и положительно определенная, для этого необходима посчитать её факторизацию Холецкого. По умолчанию, в пакете George матрица факторизуется при помощи пакета NumPy с помощью инструментов для плотных матриц, это приводит к значительным затратам по памяти и по времени. Также в пакет встроен решатель HODLR, он использует блочно-малоранговые техники для быстрой факторизации, однако так как метод HODLR — это аналог одномерных \mathcal{H}^2 матриц, для 2D и 3D задач он не эффективен. Мы предлагаем использовать для аппроксимации разреженное представление \mathcal{H}^2 матрицы, факторизованное с помощью разреженных инструментов, будем называть этот метод «спарсификация \mathcal{H}^2 ». Сравняется общее время вычисления функции правдоподобия для метода HODLR ($\varepsilon = 10^{-2}, 10^{-4}$), метода спарсификации \mathcal{H}^2 ($\varepsilon = 10^{-2}, 10^{-4}$) и наивного плотного метода из пакета NumPy.



(a) Двумерная задача



(b) Трёхмерная задача

Рис. 1: Время вычисления функции правдоподобия

Рисунок 1 показывает, что как в двумерном, так и в трёхмерном случае, для $N > 10^4$ предложенный автором метод разреженной факторизации оказывается

более эффективным, чем метод HODLR и наивный плотный метод. Кроме вычисления логарифма детерминанта, необходимо вычислять и решение системы с матрицей K , однако после построения факторизации обе эти задачи решаются за незначительное время. Распределение параметров при замене точного подсчета функции правдоподобия на приближенное изменяется незначительно. В Таблице 1 приведены коэффициенты детерминации для различных способов подсчета функции правдоподобия для двумерной и трёхмерной задачи.

	R^2 , двумерная задача	R^2 , трехмерная задача
HODLR, $\varepsilon = 1e - 2$	$1 - 1.8 \times 10^{-3}$	$1 - 2.0 \times 10^{-4}$
HODLR, $\varepsilon = 1e - 4$	$1 - 7.9 \times 10^{-5}$	$1 - 1.2 \times 10^{-4}$
Сп. \mathcal{H}^2 , $\varepsilon = 1e - 2$	$1 - 8.5 \times 10^{-5}$	$1 - 4.2 \times 10^{-4}$
Сп. \mathcal{H}^2 , $\varepsilon = 1e - 4$	$1 - 3.4 \times 10^{-5}$	$1 - 4.6 \times 10^{-5}$
Плотн.	$1 - 8.4 \times 10^{-6}$	$1 - 2.6 \times 10^{-6}$

Таблица 1: Коэффициент детерминации для различных методов подсчета правдоподобия.

Заключение.

Диссертация посвящена приближенной факторизации блочно-малоранговых матриц. Основной результат работы состоит в том, что предложены новые приближенные факторизации блочно-малоранговых матриц и методы их построения, реализован программный комплекс, который применен к нескольким задачам математического моделирования. В частности:

1. Предложена факторизация разреженной симметричной положительно определённой матрицы в виде произведения матриц перестановки, блочно-диагональных унитарных и разреженных нижне-треугольных факторов (метод компрессии и исключения, compress and eliminate method, CE), на основе которой предложен прямой метод решения систем и метод построения предобуславливателя.
2. Предложены два типа разреженной факторизации \mathcal{H}^2 матриц: «расширенная» и «не-расширенная», которые позволяют ускорить решение ли-

нейных систем с \mathcal{H}^2 матрицами, в сравнении с итерационным методом BiCGStab и прямыми методами HODLR и IFMM.

3. Разработан комплекс программ реализующих представленные алгоритмы. Для факторизации SE проведено тестирование на системах, полученных при конечно-разностной дискретизации стационарного уравнения диффузии, а также системах, полученных при конечно-элементной дискретизации уравнения Пуассона и уравнения упругости. Проведено сравнение реализации метода SE с прямыми методами CHOLMOD и UMFPACK, а также итерационным методом BiCGStab с предобуславливателями ILU0, ILU_t и ILU2. Для метода расширенной разреженной факторизации проведено тестирование на задачах электростатики и гидромеханики, в ходе которого метод показал свою эффективность по времени в сравнении с непредобусловленным методом BiCGStab. Для нерасширенной разреженной факторизации проводилось тестирование на задаче моделирования гауссовских процессов для задачи регрессии. Метод показал свою эффективность по времени в сравнении с методами HODLR и IFMM.

Одним из ключевых объектов рассмотрения диссертации является разреженная матрица, полученная при дискретизации дифференциального уравнения. На основании идеи о том, что обратная к такой матрице является блочно-малоранговой матрицей, построен метод компрессии и исключения, “на ходу” аппроксимирующий заполнение, которое возникает в матрице при исключении.

Также рассмотрен такой тип блочно-малоранговых матриц, как \mathcal{H}^2 матрицы. Для них предложены два метода приведения к эквивалентным разреженным матрицам. Таким образом, трудная и важная задача приближенной факторизации \mathcal{H}^2 матриц сводится к более простой задаче приближенной факторизации разреженной матрицы.

Все предложенные методы приближенной факторизации могут быть рассмотрены в качестве приближенных методов решения и предобуславливателей для блочно-малоранговых матриц.

Для каждого предложенного в диссертации метода представлена его программная реализация. В качестве приложения предложенных методов, прово-

дится моделирование коррелированного шума на основе гауссовского процесса для задачи регрессии.

Публикации по теме диссертации.

Статьи в журналах из списка ВАК:

1. **Sushnikova D. A., Oseledets I. V.** Preconditioners for hierarchical matrices based on their extended sparse form // Russian Journal of Numerical Analysis and Mathematical Modelling. — 2016. — Т. 31. — с. 29–40.
2. **Сушникова Д. А.** Приложение блочно-малоранговых матриц для задачи регрессии на основе гауссовских процессов //Вычислительные методы и программирование, — 2017. — Т. 18. — С. 214-220.

Статьи в журналах Web of Science:

3. **Ryzhakov G. V. , Mikhalev, A. Y., Sushnikova, D. A., Oseledets, I. V.** Numerical solution of diffraction problems using large matrix compression // Antennas and Propagation (EuCAP), 2015 9th European Conference on. — IEEE, — 2015. – с. 1-3.

Прочие статьи:

4. **Sushnikova D. A., Oseledets I. V.** ” Compress and eliminate” solver for symmetric positive definite sparse matrices //arXiv preprint arXiv:1603.09133. — 2016.
5. **Sushnikova D. A., Oseledets I. V.** Simple non-extensive sparsification of the hierarchical matrices //arXiv preprint arXiv:1705.04601 — 2017.