MSk - пакет программ для приближения плотных матриц в мозаично-скелетонном формате

Валиахметов Б. И.   Тыртышников Е. Е.

ВМК МГУ, ИВМ РАН

valiahmetovbulat@mail.ru

«Суперкомпьютерные дни в России»
Москва, 2023

Содержание

Структурированные матрицы

В вычислительных задачах могут возникать структурированные матрицы (или приближаться такими):

- малоранговые: $A = UV^*$;
- разреженные: $\#\{A_{ij} \neq 0\} \ll MN$;
- тёплицевы: $A_{ij} = a(i - j)$;..

Преимущества относительно плотных матриц

1. Хранится *малое* количество параметров.
2. Матричные операции производятся *быстро*.

Структурированные матрицы

Проблема

Матрица не может быть представлена полностью одним из форматов, но допускает малопараметрическое представление.

Возможные причины (на примере операторов на сетках):

- несимметричность области;
- разнородность области и взаимодействия;
- нерегулярность сетки.

Мозаичное разбиение

Пусть матрица A задается как функция вычисления элемента:

$$A_{ij} = f(y_i, x_j),$$

где $y_i \in Y$, $x_j \in X$, $X, Y \subset \mathbb{R}^d$.
Идея: иерархически разбивать X и Y на подмножества и пытаться находить структуру в соответствующих блоках. Результат – мозаичное разбиение матрицы.

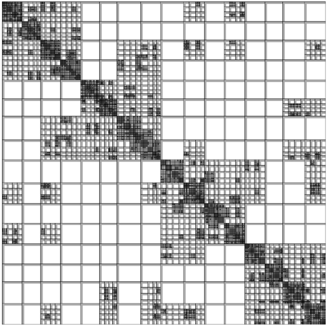


Рис.: Мозаичное разбиение

Мозаично-скелетонный формат

Пусть функция $f(y, x)$ обладает следующим свойством. Если $\hat{X} \subset X$ и $\hat{Y} \subset Y$ *«достаточно далеко»* друг от друга, то матрица $\hat{A} = f(\hat{Y}, \hat{X})$ имеет *малый ранг*: $\operatorname{rank} \hat{A} \ll \min(|\hat{X}|, |\hat{Y}|)$. Тогда можно при достаточно мелком мозаичном разбиеним матрицы A выделить в неё блочно-малоранговую структуру. Такое представление называется **мозаично-скелетонным** форматом.

Что такое *«достаточно далеко»* и какие функции обладают таким свойством?

Асимптотически гладкие функции

Theorem

[a] *Пусть $f(y,x)$ – асимптотически гладкая функция:*

$$\left|\frac{\partial^k f}{\partial y^k}\right| \leq C_f h^k k! \|x-y\|^{g-k}$$

для некоторых постоянных C_f, h и g. Тогда при $\hat{X} \subset [-L/2, L/2]^d$ и $\hat{Y} : \rho(\hat{X}, \hat{Y}) > L$ для произвольного $q \in (0,1)$ существует представление

$$f(\hat{Y}, \hat{X}) = \tilde{A} + R,$$

где $\operatorname{rank} \tilde{A} \leq C(h,d,q)k^d$ *и* $\|R\| \leq q^k L^g \sqrt{|\hat{X}||\hat{Y}|}$.

[a] Tyrtyshnikov, “Mosaic-skeleton approximations”.

Новая реализация

Имеются различные реализации мозачно-скелетонного метода[1].
Частый их недостаток - «монолитность» кода.
Цели новой реализации:

1. Поддержка различных типов данных.
2. Гибкость (расширяемость) с точки зрения построения мозаичной структуры и аппроксимации.
3. Параллельность, в том числе для новых пользовательских типов данных.
4. Балансировка вычисления элементов матрицы между узлами.
5. Собственная реализация.

[1] https://github.com/gchavez2/awesome_hierarchical_matrices

Новая реализация

Результат процецуры аппроксимации – матрица в мозаичном формате. Предоставляемый на выходе интерфейс:

1. матрично-векторное произведение: $y = \mathrm{op}(A)x$, $\mathrm{op}(A) \in \{A, A^T, A^*\}$;
2. вычисление элемента $A(i, j)$;
3. вычисление нормы, ошибки аппроксимации, рангов.

Детали реализации

Мозаика

За разбиение отвечает класс `Oracle`. Его задача – для блока матрицы предоставить информацию: вычислить его или разбить на несколько других.

Предоставляемые реализации:

1. `Elementary` - диагональные блоки рекурсивно разбиваются пополам (структура HODLR);
2. `Distant` - индексы отвечают точкам из $\mathbb{R}^d$, блок разбивается, если «сферы», вмещающие точки строк и столбцов, пересекаются.

Детали реализации

Блоки

`Oracle` обязан для каждого неразбиваемого блока указать его *вид представления данных* и *алгоритм вычисления*.

Предоставляемые реализации блоков и способов вычисления:

1. `Dense` - плотный блок (и одноименный вычислитель) с явным вычислением элементов;
2. `LowRank` - малоранговый блок с вычислителями `SVD` и `Cross`.

Детали реализации

Параллельность

Модель параллелизма основана на очередях и рабочих.

- На одном узле: набор блоков для вычисления, очередь на вычисление элементов и набор рабочих-потоков. Аппроксимация происходит **итерационно** – потоки выполняют порцию работы и запрашивают необходимые элементы матрицы.
- Между узлами: балансировка распределения блоков и балансировка (сообщениями) очередей на вычисление элементов матрицы.

Ядро Гельмгольца

Пример 1: $f(y,x) = \dfrac{\exp(ikr)}{4\pi r}$, $r = y - x$, $k = 1$. Равномерная прямоугольная сетка на единичном квадрате, по 100 делений на сторону: $N = 10000$ $(X = Y)$.

ошибка	сжатие
10^{-2}	7.1%
10^{-4}	9.3%
10^{-6}	12.6%

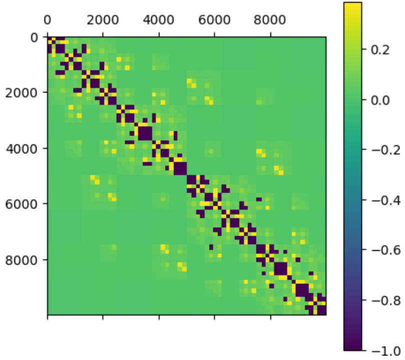


Рис.: Мозаика на квадрате, точность 10^{-4}

Интегральное уравнение

Пример 2: задача дифракции на идеальном проводнике. Элемент матрицы задается интегральным выражением, вычисляемым по квадратурным формулам:

$$\begin{aligned} A_{ij} &= k^2 \int\limits_{\sigma_i} \left(\int\limits_{\sigma_j} F(x-y)\vec{e}_j(y)dy \right) \vec{e}_i(x)dx \\ &- \int\limits_{\sigma_i} \left(\int\limits_{\sigma_j} F(x-y) \operatorname{Div} \vec{e}_j(y)dy \right) \operatorname{Div} \vec{e}_i(x)dx, \end{aligned}$$

где $\vec{e}_i(y)$ - базисная функция RWG на ребре y, $F(r) = \dfrac{e^{ikr}}{4\pi r}$.

Интегральное уравнение

Квазиравномерная треугольная сетка на миндалевидном теле.
Количество уравнений $N = 12936$.

ошибка	сжатие
10^{-2}	39.4%
10^{-4}	40.4%
10^{-6}	41.9%

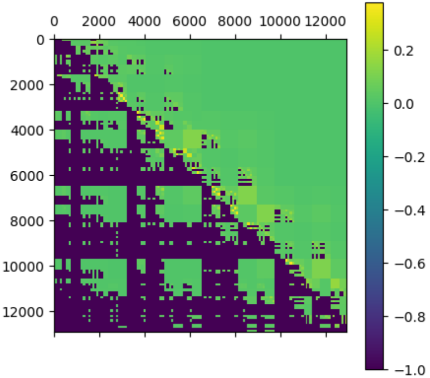


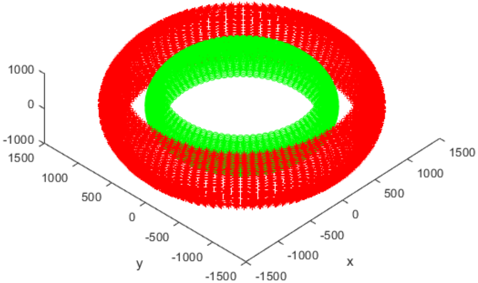
Рис.: Мозаика на Almond, точность 10^{-4}

Магнитный диполь

Пример 3: $f(y,x) = \frac{1}{\rho^3}\left(rr^T - I_3\right)$, где $y - x = \rho r$, $\rho = \|y - x\|$.

Матрица из блоков 3×3, точки расположены квазиравномерно на вложенных скругленных цилиндрах. Размеры матрицы: $M = 15000$, $N = 9600$.

ошибка	сжатие
10^{-3}	4.7%
10^{-4}	7.0%
10^{-6}	11.8%



Магнитный диполь

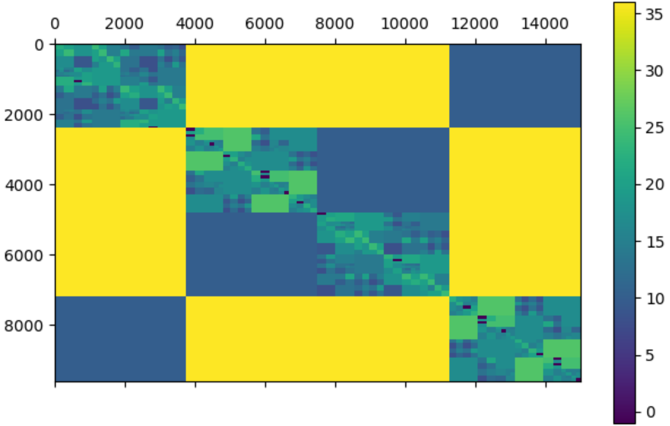


Рис.: Ранги на скругленных цилиндрах, точность 10^{-3}

Параллельность

Функция из Примера 1, равномерная прямоугольная сетка на единичном квадрате, по 300 делений на сторону: $N = 90000$ $(X = Y)$. Тесты проводились на кластере ИВМ РАН[2].

Рис.: На общей памяти

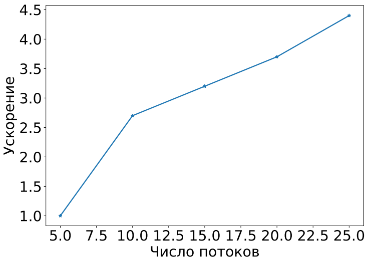
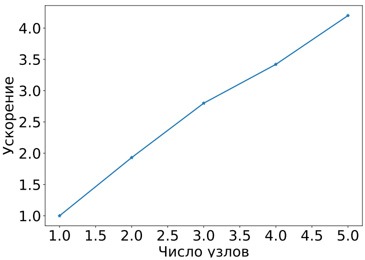


Рис.: На распределенной памяти



[2] cluster2.inm.ras.ru

Итоги

- Выполнена реализация мозаично-скелетонного метода аппроксимации плотных матриц на языке C++ в рамках библиотеки `MSk`.
- Пакет программ имеет несколько заменяемых модулей и поддерживает добавление пользовательских форматов.
- Процедура аппроксимации работает параллельно как на общей, так и на распределенной памяти.

Дальнейшая работа

- Интерфейс к C и Fortran.
- Документация и тестирование.
- Добавление новых типов блоков.
- Улучшение эвристик построения мозаики.
- Улучшение параллельных свойств.
- Синхронная аппроксимация блоков, подход к иерархическим форматам.

Доступ к библиотеке MSk

Ссылка на MSk

https://gitlab.com/bulatral/mosaic-skeleton