Поиск сильно невырожденных подматриц и их связь со столбцовыми и крестовыми аппроксимациями

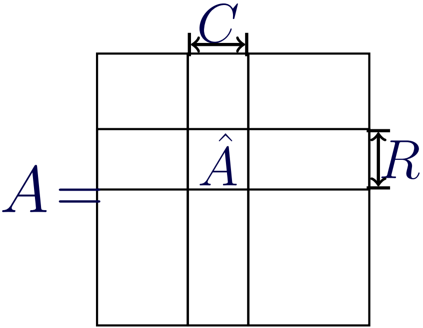
Докладчик: А. И. Осинский

По материалам диссертации на соискание ученой степени доктора физико-математических наук по специальности 1.1.6 Вычислительная математика

“Существование и построение близких к оптимальным столбцовых и крестовых аппроксимаций матриц”

12 апреля 2024 г.
Сколтех

Крестовая аппроксимация



Используется n столбцов C и m строк R для построения аппроксимации ранга $r \leqslant \min(m, n)$.

Крестовая: CGR с произвольным генератором G.
Скелетная: $C\hat{A}^{-1}R$ (соответствует неполному LU).

Столбцовая: CW. Часто применяется $W = C^{+}A$ (соответствует неполному QR).

Область применений

Быстрое $O\left((M+N)\,r^2\right)$ построение аппроксимаций, близких к сингулярному разложению. Выбор репрезентативных строк и столбцов данных.

Очень высокая скорость (позволяет использовать аппроксимацию в итеративных алгоритмах), использование малого числа элементов (важно, если элементы матрицы заданы сложной функцией).

1. Выявляющие ранг LU и QR разложения.
2. Решение задач наименьших квадратов.
3. Решение интегральных уравнений (мозаично-скелетонный метод).
4. Построение тензорных поездов (TT-Cross).
5. Уравнения Смолуховского (система N ОДУ с квадратичной правой частью).
6. Неотрицательные аппроксимации (с помощью переменных проекций).
7. Восстановление матриц (с помощью проекции градиента).

Column subset selection: анализ данных, отбор признаков (например, в популяционной генетике, при тестировании электронных схем, в рекомендательных системах); Drineas, Deshpande и др.

Мотивация

Решение задачи наименьших квадратов:

$$\|Ax - b\|_2 \to \min, \quad A \in \mathbb{C}^{N \times N},$$

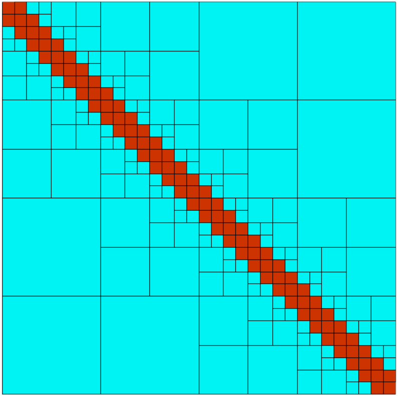
где известно, что у незашумленной ранг $r < N$.

Способы решения:

1. Усеченное сингулярное разложение A (с обнулением малых сингулярных чисел). $O\left(N^3\right)$. Приближенное: $O\left(N^2 r/\varepsilon\right)$.
   Halko, N., Martinsson, P.-G. and Tropp, J. A. (2011). Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. SIAM Review, 53(2), 217–288.
2. Полное ортогональное разложение: неполное QR, $A \approx Q_1 R$, $R = L Q_2$. $O\left(N^2 r\right)$.
   Соответствует столбцовой аппроксимации.
3. Скелетное/неполное LU разложение: $A \approx C \hat{A}^{-1} R$. $O\left(N r^2\right)$.
   При таком времени работы нельзя гарантировать точность алгоритма. Можно исследовать его свойства, свойства аппроксимаций такого вида, свойства найденных подматриц.

Мотивация

Мозаично-скелетонная аппроксимация (иерархические матрицы).



Красные блоки полного ранга, остальные малоранговые.

Возникают при решении интегральных уравнений:

S. Börm, S. Christophersen (2016). Approximation of integral operators by Green quadrature and nested cross approximation. Numerische Mathematik.

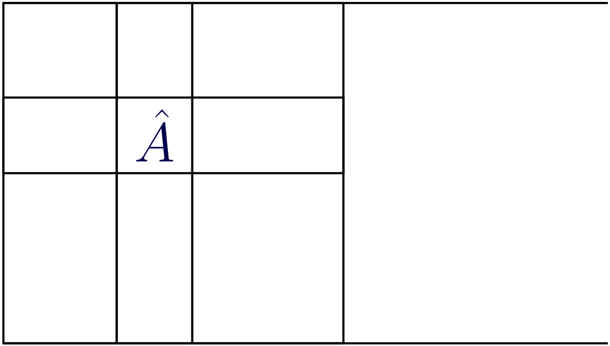
A. Setukha, S. Stavtsev, R. Tret'yakova (2022). Application of Mosaic-Skeleton Approximations of Matrices in the Physical Optics Method for Electromagnetic Scattering Problems. Computational Mathematics and Mathematical Physics.

Мотивация

Крестовый тензорный поезд: TT-крестовая интерполяция, численное интегрирование, поиск максимума и т. д.

$2Nr$ случайных столбцов
из $d-1$ измерения

k-я развертка
из N строк
$\times\ r$ предыдущих
индексов столбцов



I. V. Oseledets, E. E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays // Linear Algebra Appl. — 2010. — V. 432, no. 1. – P. 70—88.
D. A. Zheltkov, E. E. Tyrtyshnikov, Global optimization based on TT-decomposition // Russian Journal of Numerical Analysis and Mathematical Modelling. — 2020. — V. 35. – P. 247–261.

Уравнения Смолуховского

$$\frac{dn_k}{dt} = \sum_{i+j=k} C_{ij} n_i n_j - \sum_j C_{kj} n_k n_j.$$

Вычисление N правых частей: $O\left(N^2\right)$.
Если C_{ij} малоранговое: $O\left(Nr \log N\right)$.
Можно использовать, даже если ядро меняется во времени. Крестовая аппроксимация приводит к сложности $O\left(Nr^2 + Nr \log N\right)$.

С. А. Матвеев, Е. Е. Тыртышников, А. П. Смирнов, Н. В. Бриллиантов. Быстрый метод решения уравнений агрегационно-фрагментационной кинетики типа уравнений Смолуховского, Выч. мет. программирование, 15:1 (2014), 1–8.
A. I. Osinsky. Low-rank method for fast solution of generalized Smoluchowski equations. Journal of Computational Physics, 422 (2020), 109764.

Неотрицательные аппроксимации

Переменные проекции на множество матриц с элементами не меньше ε и множество матриц ранга r, пока все элементы не станут неотрицательными.

$$A_{(s+1)} = P_r P_{\geqslant\varepsilon} A_{(s)}.$$

Предложен Andersson and Carlsson, 2013. P_r – крестовая аппроксимация.

$$\varepsilon = \frac{\left\| A_{(1)} - A_{(0)} \right\|_F}{\sqrt{MN}}.$$

Требуемая относительная точность порядка константы (Budzinskiy, 2023), так что достаточно порядка r строк и столбцов.

На решении двухкомпонентной системы Смолуховского:

Метод	Flop на итерацию	Итераций	$\frac{\left\|A - A_{(s)}\right\|_F}{\|A\|_F}$	$\frac{\left\|A - A_{(s)}\right\|_C}{\|A\|_C}$
$\tilde{A} = U\Sigma_r V$, $\varepsilon = 0$	$2{,}3 \cdot 10^{10}$	≈ 700	$2{,}70 \cdot 10^{-2}$	$1{,}48 \cdot 10^{-1}$
$\tilde{A} = CGR$, $\varepsilon > 0$	$6{,}7 \cdot 10^{6}$	15	$2{,}91 \cdot 10^{-2}$	$1{,}49 \cdot 10^{-1}$

Восстановление матриц

$$M = \begin{cases} 1, & \text{для известных элементов,} \\ 0, & \text{для неизвестных элементов.} \end{cases}$$

$$\|A - M \odot X\|_F \to \inf_{X,\ \mathrm{rank}\, X \leqslant r}.$$

Singular value projection (SVP):

$$X_{(s+1)} = P_r\left(X_{(s)} + \tau\left(A - M \odot X_{(s)}\right)\right), \tag{1}$$

предложен Jain, Meka and Dhillon, 2010.

P_r – снова крестовая аппроксимация (получаем Approximate SVP).

С учетом точности быстрой крестовой аппроксимации, получаем сложность шага $O\left(Nr^2/\delta + r^3/\delta^3\right)$. Шаги быстрее (градиентной модификации) ALS при $\delta \gg \sqrt{r/N}$.

Точность аппроксимации

Точность аппроксимации сравнивается с точностью неполного сингулярного разложения ранга r (в спектральной норме или норме Фробениуса):

$$\|A - CGR\|_\xi \leqslant f\left(r, m, n, M, N, \|\cdot\|_\xi\right) \min_{Z,\ \mathrm{rank}\ Z = r} \|A - Z\|_\xi .$$

$G = G\left(\hat{A}\right)$, $\hat{A} \in \mathbb{C}^{m \times n}$, $A \in \mathbb{C}^{M \times N}$. $\xi = C, 2, F$.

Квадратная матрица и подматрица: $f\left(r, n, N, \|\cdot\|_\xi\right)$.

Можно ли оценить f так, чтобы она была близка к нижним оценкам в соответствующей норме?

Определения

Объем:

$$\mathcal{V}\left(\hat{A}\right) = \sqrt{\max\left(\det\left(AA^*\right), \det\left(A^*A\right)\right)}.$$

Квадратная подматрица, максимизация объема (Горейнов, Тыртышников, Замарашкин, 1997).

Прямоугольная подматрица:

$C\hat{A}^+R$ аппроксимация (Михалев, Оселедец, 2014). Один из размеров подматрицы больше r.

Проективный объем:

$$\mathcal{V}_r\left(\hat{A}\right) = \prod_{k=1}^{r} \sigma_k\left(A\right).$$

$C\hat{A}_r^+R$ аппроксимация. Оба размера подматрицы больше r.

$\hat{A}_r$ – сокращенное сингулярное разложение:

$\sigma_{r+1}\left(A_r\right) = \ldots = \sigma_n\left(\hat{A}_r\right) = 0.$

Локальная максимальность

Если $\mathcal{V}\left(\tilde{A}\right) \leqslant \rho\mathcal{V}\left(\hat{A}\right)$, $\rho > 1$ для любой матрицы $\tilde{A}$, отличающейся от $\hat{A}$ одной строкой и/или столбцом, то подматрица $\hat{A}$ обладает ρ-*локально максимальным* объемом (Gu, Eisenstat, 1996).

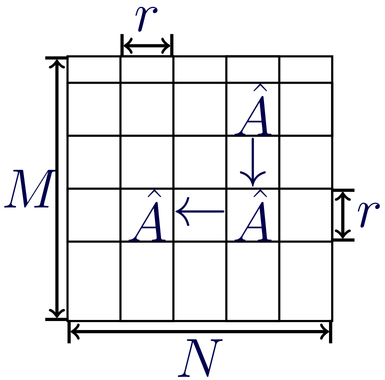
Если $\mathcal{V}_r\left(\tilde{A}\right) \leqslant \rho\mathcal{V}_r\left(\hat{A}\right)$, то $\hat{A}$ обладает ρ-*локально максимальным* проективным объемом.

Смысл: приближенной поиск максимального объема – NP-сложная задача (Civril, Magdon-Ismail, 2013). Поиск ρ-локально максимального объема может быть осуществлен за полиномиальное время.

Поиск локально максимального объема

maxvol

(Горейнов и др., 2010)



Замен в строках/столбцах:
$O\left(r/\log\rho\right)$
(предыдущая оценка
$O\left(r\log_\rho(MN)\right)$, Boutsidis, 2011)

За $O\left(N^2 r\log_\rho r\right)$ можно гарантировать оценки для одновременных замен, но это дорого.

Есть аналоги для прямоугольных подматриц и подматриц локально максимального проективного объема (упрощенно).

Скелетная аппроксимация

Теорема 1 (Горейнов, Тыртышников, 2011).

Пусть $\hat{A} \in \mathbb{C}^{r\times r}$ – подматрица ρ-локально максимального объема в матрице $A \in \mathbb{C}^{M\times N}$ ранга не ниже r. Тогда

$$\|A - C\hat{A}^{-1}R\|_C \leqslant \rho\,(r+1)^2 \min_{Z,\ \mathrm{rank}\ Z=r} \|A - Z\|_C\,.$$

Для неполного LU разложения с полным выбором ведущего элемента коэффициент может быть $\sim 2^r$.

$$A = \begin{bmatrix} 1 & -1 & \cdots & -1 \\ & \ddots & \ddots & \vdots \\ & & 1 & -1 \\ & & & 1 \end{bmatrix} \in \mathbb{R}^{(r+1)\times(r+1)}.$$

Нижняя оценка: $r+1$. Является ли верхняя точной?

Проективный объем

Теорема 2.

Пусть $\hat{A} \in \mathbb{C}^{(2r-1)\times(2r-1)}$ *– подматрица* ρ*-локально максимального* r*-проективного объема в матрице* $A \in \mathbb{C}^{M\times N}$ *ранга не ниже* r*. Тогда*

$$\|A - C\hat{A}_r^+ R\|_C \leqslant 4\rho r \min_{Z,\ \mathrm{rank}\, Z = r} \|A - Z\|_C \,.$$

Основана на том, что в строках R для любого столбца $R_{:,i}$

$$\left\|\hat{A}_r^+ R_{:,i}\right\|_2 \leqslant \sqrt{\frac{r}{n-r+1}}.$$

В терминах правых сингулярных векторов $U \in \mathbb{C}^{r\times N}$, суммируя по всем столбцам

$$\left\|\hat{U}^+\right\|_F \leqslant \sqrt{r\frac{N-r+1}{n-r+1}}.$$

Тензорный случай

Теорема 3.

Пусть $E = A - Z$ погрешность некоторого приближения Z ранга r тензора $A \in \mathbb{C}^{N_1 \dots N_d}$. Тогда существует крестовый тензорный поезд $\tilde{A}$ такой, что размеры всех подматриц в нем не больше $n = 2r - 1$, их ранг сокращен до не больше r, и такой, что

$$\left\| A - \tilde{A} \right\|_C \leqslant \frac{\left(4\sqrt{2r-1}\right)^{\lceil \log_2 d \rceil} - 1}{4\sqrt{2r-1} - 1} \cdot 4r \left\| E \right\|_C$$

Коэффициент $\sim \left(4\sqrt{2}r\right)^{\log_2 \sqrt{2d}}$. Достигается ли?

Тривиальная оценка для $n = r$: $(r+1)^d$.

Савостьянов, 2014: $\sim \left(C \cdot r \max\limits_{\hat{A} \in \mathbb{C}^{r \times r}} \mathrm{cond}_C \left(\hat{A} \right) \right)^{\log_2 (2d)}$.

Под E можно подразумевать погрешность тензорного поезда ранга r или канонического разложения ранга r или максимум по разверткам погрешности наилучшей матричной аппроксимации ранга r развертки.

Спектральная норма

По спектральной норме невозможно гарантировать построение аппроксимации столь же высокой точности, что для нормы Фробениуса.

Теорема 4.

$$\sup_{A \in \mathbb{C}^{N \times N}} \min_{C \in \mathbb{C}^{N \times n}} \frac{\|A - CC^+ A\|_2}{\|A - A_r\|_2} = \sup_{\substack{U \in \mathbb{C}^{r \times N}, \\ UU^* = I}} \min_{\substack{\hat{U} \in \mathbb{C}^{r \times n}, \\ \operatorname{rank} \hat{U} = r}} \left\| \hat{U}^+ \right\|_2 = t(r, n, N).$$

$$\sqrt{\frac{N - r + 1}{n - r + 1}} \leqslant t(r, n, N) \leqslant \frac{\sqrt{N}}{\sqrt{n+1} - \sqrt{r}} + \frac{1}{\sqrt{N(n+1)}}.$$

Для $n = r$ (можно ли лучше?):

$$\sqrt{(r+1) \left\lfloor \frac{\min(M, N)}{r+1} \right\rfloor} \leqslant \frac{\|A - CW\|_2}{\|A - A_r\|_2} \leqslant \sqrt{1 + r(\min(M, N) - r)}.$$

Норма Фробениуса

Теорема 5.

$$\sup_{A\in\mathbb{C}^{N\times N}} \min_{C} \frac{\|A-(CC^{+}A)_r\|_F}{\|A-A_r\|_F} \geqslant \sqrt{1+\frac{\sup\limits_{U\in\mathbb{C}^{r\times N}} \min\limits_{\hat{U}\in\mathbb{C}^{r\times N}} \left\|\hat{U}^{+}\right\|_F^2 - r}{N-r}}.$$

Снова встречаем $\hat{U}^{+}$.

Несложный пример приводит к нижней оценке

$$\left\|A-\left(CC^{+}A\right)_r\right\|_F \geqslant \sqrt{1+\frac{r}{n}\cdot\frac{N-r}{N-n}}\,\|A-A_r\|_F\,.$$

Норма Фробениуса

$$
A = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & \ddots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & 1 & \cdots & 1 \\
0 & 0 & 0 & -\varepsilon\sqrt{N-r} & \frac{\varepsilon}{\sqrt{N-r}} & \cdots & \frac{\varepsilon}{\sqrt{N-r}}
\end{bmatrix} \in \mathbb{R}^{(r+1)\times N}.
$$

Здесь $\sigma_{r+1}(A) = \varepsilon\sqrt{N-r+1}$, однако из-за примерно равных весов W (при выборе подматрицы максимального объема) в каждом столбце в последней строке наберется погрешность $\sim \varepsilon\sqrt{N-r}$.

Оценки в среднем

RANDSVD ансамблем матриц $A \sim \mathrm{RANDSVD}(\Sigma)$ называется множество матриц вида $A = W_L \Sigma W_R$, где $\Sigma \in \mathbb{C}^{M\times N}$ фиксирована, а $W_L \in \mathbb{C}^{M\times M}$ и $W_R \in \mathbb{C}^{N\times N}$ – независимые случайные унитарные.

Теорема 6.

Пусть $A \in \mathbb{C}^{M\times N}$ – RANDSVD матрица, $Z = A_r$. Пусть столбцы $Z_C \in \mathbb{C}^{M\times n}$ локально максимального проективного объема в Z. Подматрица $\hat{A}P_r \in \mathbb{C}^{n\times n}$ – локально максимального проективного объема в CP_r, $P_r = Z_C^+ Z_C$. Тогда

$$\mathbb{E}_{W_L, W_R} \left\| A - C \left(\hat{A} P_r \right)^+ R \right\|_F \leqslant \left(1 + \frac{r}{n - r + 1} \right) \| A - A_r \|_F .$$

Реальный коэффициент $1 + \frac{\left(\underset{\hat{U} \in \mathbb{C}^{r\times n}}{\mathbb{E}} \left\| \hat{U}^+ \right\|_F^2 - r \right)}{N - r}$. $\left\| \hat{U}^+ \right\|_F \leqslant \sqrt{r \frac{N - r + 1}{n - r + 1}}$.

Локально максимальный объем предпочтителен, поскольку в самой матрице A поиск всегда остановится (если на каждом шаге объем увеличивать). Вероятность превысить $1 + \frac{cr}{n - r + 1}$ порядка $e^{-c/2}$.

Некоторые полученные оценки

$$\sqrt{\frac{N-r+1}{n-r+1}} \leqslant \frac{\|\boldsymbol{A}-\boldsymbol{C}\boldsymbol{W}\|_2}{\|\boldsymbol{A}-\boldsymbol{A_r}\|_2} \leqslant \frac{\sqrt{N}}{\sqrt{n+1}-\sqrt{r}} + \frac{1}{\sqrt{N(n+1)}}. \quad (M=N).$$

$$\sqrt{1+\frac{r}{n}\cdot\frac{N-r}{N-n}} \leqslant \frac{\|\boldsymbol{A}-\boldsymbol{C}\boldsymbol{W_r}\|_F}{\|\boldsymbol{A}-\boldsymbol{A_r}\|_F} \leqslant \sqrt{1+\frac{r}{n+r+1-\sqrt{1+4r(n+1)}}}.$$

$$\frac{\left\|\boldsymbol{A}-\boldsymbol{C}\left(\hat{\boldsymbol{A}}\boldsymbol{P_r}\right)^{+}\boldsymbol{R}\right\|_F}{\|\boldsymbol{A}-\boldsymbol{A_r}\|_F} \leqslant 1+\frac{r}{n+r+1-\sqrt{1+4r(n+1)}}. \quad (m=n).$$

Предыдущая лучшая оценка сверху (Boutsidis, Woodruff, 2014):
$1+40\frac{r}{n-4r}$.

$$\frac{\sqrt{r}\ln r}{8\ln(2r)} \leqslant \frac{\|\boldsymbol{A}-\boldsymbol{C}\hat{\boldsymbol{A}}^{+}\boldsymbol{R}\|_C}{\min\limits_{\boldsymbol{Z},\ \mathrm{rank}\,\boldsymbol{Z}=r}\|\boldsymbol{A}-\boldsymbol{Z}\|_C} \leqslant 3\sqrt{3r-1}. \quad (m=2r-1, n=3r-1).$$

И др. Можно ли получать нижние напрямую для крестовых?

Некоторые эксперименты

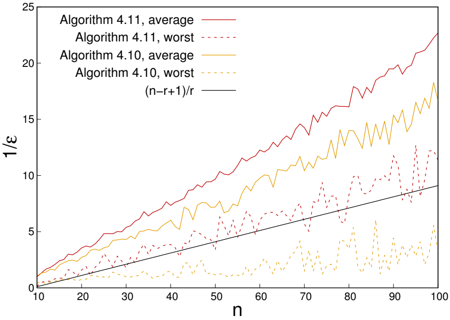


Рис.: Зависимость $1/\varepsilon$ от числа строк и столбцов подматрицы, аппроксимация с помощью алгоритмов поиска подматриц большого проективного объема случайных матриц ранга $r = 10$, $\sigma_k(A) = 1/2^k$, 100 генераций для каждого размера n.

Коэффициент погрешности $1 + \varepsilon \sim 1 + \frac{r}{n-r+1}$.

Поиск сильно невырожденных подматриц

Таблица: Сложность и точность при поиске сильно невырожденных подматриц.

Источник	$\lVert\hat{U}^+\rVert_F^2$	$\lVert\hat{U}^+\rVert_2^2$	Сложность
Leverage ($\delta = 1/2$), $n \geqslant 32r\ln(4r)$	$4rN$	$4N$	$O(Nr + n\log n)$
Min. Vol. Ell. ($\varepsilon = 1$), $n \geqslant 4r\ln\ln r + 42r$	$r + 2(N-n)$	$1 + 2(N-n)$	$O(Nr^2\log\log r)$
Graph spectral sparsification, $n > r$	$r\frac{(\sqrt{N}+\sqrt{n})^2}{(\sqrt{n}-\sqrt{r})^2}$	$\frac{(\sqrt{N}+\sqrt{n})^2}{(\sqrt{n}-\sqrt{r})^2}$	$O(Nnr^2)$
SRRQR, $n = r$	$r + \rho^2 r(N-r)$	$1 + \rho^2 r(N-r)$	$O(Nr^2\log N/\log\rho)$
Жадный (с конца), $n \geqslant r$	$r\frac{N-r+1}{n-r+1}$	$1 + r\frac{N-n}{n-r+1}$	$O\left(N^2 r\log\varepsilon^{-1}\right)$
Volume sampling ($\delta = 1/2$), $n = r$	$\rho^2 r(N-r+1)$	$\rho^2 r(N-r+1)$	$O(Nr^3/\log\rho)$

Avron H., Boutsidis C. Faster Subset Selection for Matrices and Applications // SIAM Journal on Matrix Analysis and Applications. — 2011. — Vol. 34, no. 4.
Woodruff D., Yasuda T. New Subset Selection Algorithms for Low Rank Approximation: Offline and Online // arXiv 2304.09217 (Submitted on 18 Apr 2023).

Поиск сильно невырожденных подматриц

Таблица: Сложность и точность при поиске сильно невырожденных подматриц.

НОВЫЕ ОЦЕНКИ			
Источник	$\lVert\hat{U}^+\rVert_F^2$	$\lVert\hat{U}^+\rVert_2^2$	Сложность
Min. Vol. Ell. ($\varepsilon = 1$), $n > 7{,}45r$	$r + 2(N-n)$	$1 + 2(N-n)$	$O\left(Nr^2 \log\log r\right)$
Теорема 32, $n \geqslant r$	$r\frac{N+2}{\left(\sqrt{n+1}-\sqrt{r}\right)^2}$	$\frac{N+2}{\left(\sqrt{n+1}-\sqrt{r}\right)^2}$	$O(Nnr^2)$
Теорема 33, $n = r$	$r\left(N-r+1\right)$	$1 + r\left(N-r\right)$	$O(Nr^2)$
Теорема 34, $n \geqslant r$	$r\frac{N-r+1}{n-r+1}$	$1 + r\frac{N-n}{n-r+1}$	$O\left(N^2 r\right)$
Der. Vol. Sam., $n \geqslant r$	$r\frac{N-r+1}{n-r+1}$	$1 + r\frac{N-n}{n-r+1}$	$O\left(Nnr^3\right)$
dominant, $n \geqslant r$,	$r + \rho^2 r\frac{N-n}{n-r+1}$	$1 + \rho^2 r\frac{N-n}{n-r+1}$	$O\left(Nr^2 \log\log r + Nnr/\log\rho\right)$

Идея достижения спектральной нормы

Лемма 1.

Пусть для $A \in \mathbb{C}^{r\times r}$, $A = A^*$ *справедливо* $\lambda_{\min}(A) > l$, $\mathrm{tr}\left((A - lI)^{-1}\right) \leqslant 1/\delta_l$. *Пусть* $v \in \mathbb{C}^r$ *– произвольный вектор.* *Тогда* $\lambda_{\min}(A) > l + \delta_l$, *и если* $L_A(v) \geqslant 1/t > 0$, *где*

$$L_A(v) = \frac{\left\|(A - (l+\delta_l)I)^{-1}\, v\right\|_2^2}{\delta_l\, \mathrm{tr}\left((A-(l+\delta_l)I)^{-1}\,(A - lI)^{-1}\right)} - v^*\,(A - (l+\delta_l)I)^{-1}\, v,$$

то

$$\mathrm{tr}\left((A + tvv^* - (l+\delta_l)I)^{-1}\right) \leqslant \mathrm{tr}\left((A - lI)^{-1}\right) \leqslant 1/\delta_l.$$

Batson J. D., Spielman D. A., Srivastava N. Twice-Ramanujan Sparsifiers // Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing. – 2009. – P. 255-262.

Идея достижения спектральной нормы

Формула Шермана-Моррисона для следа:

$$\operatorname{tr}(B + tvv^*)^{-1} = \operatorname{tr} B^{-1} - \frac{v^* B^{-2} v}{1/t + v^* B^{-1} v}.$$

Для $B = A - (l + \delta_l) I$:

$$\operatorname{tr}(A - (l + \delta_l) I + tvv^*)^{-1} =$$
$$= \operatorname{tr}(A - (l + \delta_l) I)^{-1} - \frac{v^* (A - (l + \delta_l) I)^{-2} v}{1/t + v^* (A - (l + \delta_l) I)^{-1} v}. \quad (2)$$

Мы хотим $\operatorname{tr}(A - (l + \delta_l) I + tvv^*)^{-1} \leqslant \operatorname{tr}(A - lI)^{-1}$. Подставляя в это неравенство правую часть (2), получаем критерий.

Идея достижения спектральной нормы

Теорема 7.

$$t(r,n,N) \leqslant \frac{\sqrt{N}}{\sqrt{n+1}-\sqrt{r}} + \sqrt{\frac{1}{N(n+1)}}.$$

Соответствующая подматрица может быть найдена за $O(Nnr^2)$ *операций.*

Выбирается

$$\delta_l = \frac{1}{\frac{N}{1-\sqrt{\frac{r}{n+1}}}+2} \quad \text{и} \quad l_0 = -\frac{\sqrt{r(n+1)}}{\frac{N}{1-\sqrt{\frac{r}{n+1}}}+2}.$$

Этого достаточно (что доказывается случайным равновероятным выбором столбца), чтобы не требовалось $t > 1$ (чтобы достичь $L_A(v) \geqslant 1$).
Далее по индукции $l = l_0 + k\delta_l$, $k = \overline{0, n-1}$.
База: $\operatorname{tr}\left((0 - l_0 I)^{-1}\right) \leqslant 1/\delta_l$.

Идея набора r столбцов

Теорема 8.

Пусть даны ортонормированные строки $\hat{V} \in \mathbb{C}^{r \times N}$. *Тогда за* $O(Nr^2)$ *можно найти подматрицу* $\hat{V} \in \mathbb{C}^{r \times r}$ *матрицы* V *такую, что*

$$\left\| \hat{V}^{-1} \right\|_F \leqslant \sqrt{r\,(N - r + 1)}$$

и

$$\left\| \hat{V}^{-1} \right\|_2 \leqslant \sqrt{1 + r\,(N - r)}.$$

Жадный набор. Критерий:

$$j = \operatorname*{arg\,max}_{j > k} \|V_{2,j}\|_2^2 / \left(1 + \left\| \hat{V}_1^{-1} V_{1,j} \right\|_2^2\right), \quad V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \hat{V}_1 & * \\ 0 & * \end{bmatrix}.$$

Достигаются также на максимальном объеме. Нет лучших оценок для квадратных.

Идея удаления столбцов

Лемма 2.

Для $B = [\, A \; b \,] \in \mathbb{C}^{r\times(n+1)}$, $n \geqslant r$, *выполнено*

$$\|A^+\|_F^2 = \|B^+\|_F^2 + \frac{\left\|(BB^*)^{-1}\, b\right\|_2^2}{1 - \|B^+b\|_2^2}.$$

Симметричное обновление $(BB^*)^{-1}$, его след.

Идея удаления столбцов

Теорема 9.

При последовательном удалении столбцов из произвольных строк $R \in \mathbb{C}^{r \times N}$ до размера $\hat{R} \in \mathbb{C}^{r \times n}$ справедлива оценка

$$\left\| \hat{R} \right\|_F \leqslant \sqrt{\frac{N - r + 1}{n - r + 1}} \left\| R \right\|_F .$$

Общая сложность удаления столбцов $O\left(N\left(N - n\right) r\right)$.

По индукции: удаление одного столбца в среднем (если выбрать столбец случайно равновероятно, отдельно для числителя и знаменателя) $\sqrt{\frac{N-r+1}{(N-1)-r+1}}$.

Близость локально максимального к глобально максимальному

Теорема 10.

Пусть в подматрицу $\hat{A} \in \mathbb{C}^{r \times n}$ *разрешается набирать столбцы* $A \in \mathbb{C}^{r \times N}$ *по нескольку раз, и она обладает* ρ*-локально максимальным объемом.*
Тогда

$$\max_{\tilde{A} \in \mathbb{C}^{r \times n}} \operatorname{vol}\left(\tilde{A}\right) / \operatorname{vol}\left(\hat{A}\right) \leqslant \left(\frac{n}{n-r+1}\left(1+\left(\rho^2-1\right)\frac{n}{r}\right)\right)^{r/2}.$$

Каждый столбец $\hat{A}^{+}\tilde{A}_{:,k}$ по норме не больше $\sqrt{\frac{r+(\rho^2-1)n}{n-r+1}}$.
По норме Фробениуса $\left\|\hat{A}^{+}\tilde{A}\right\|_F \leqslant \sqrt{\frac{nr}{n-r+1}\left(1+\left(\rho^2-1\right)\frac{n}{r}\right)}$.
Неравенство между средним арифметическим (квадратов сингулярных чисел $\hat{A}^{+}\tilde{A}$) и средним геометрическим дает искомое отношение объемов.
При $n = 2r$ экспонента от r.

Число шагов для локально максимального объема

Следствие 1.

За $O\left(Nr^2 \log\log r\right)$ операций можно достичь подматрицы, объем которой отличается от максимального не более, чем в $6^{r/2}$ раз.

Идея доказательства: набрать $2r$ столбцов, достичь для них $\sqrt{2}$-локально максимального объема, затем удалить.

Этот результат позволяет находить ρ-локальный за $O\left(r/\log\rho\right)$ шагов.

Таким же образом находится хорошая стартовая подматрица для поиска минимального эллипсоида (избегаем $4r\ln\ln r$ лишних столбцов).

Если разрешить брать те же столбцы по несколько раз, оценка будет существенно лучше, и позволит также искать $r \times r$ подматрицы $e^{r/2}$-максимального объема.

Заключение

ОСНОВНЫЕ РЕЗУЛЬТАТЫ:

1. Получены близкие к оптимальным оценки крестовых и столбцовых аппроксимаций (верхние и нижние).
2. Вероятностные оценки точности позволяют оценить эффективность в тех случаях, когда необходимо получить аппроксимацию, используя лишь малую часть входных данных.
3. Предложенные алгоритмы позволяют строить приближение с коэффициентом погрешности по норме Фробениуса $1+\varepsilon$ за $O\left(Nr^2/\varepsilon+(r/\varepsilon)^3\right)$ операций, чего не могут достичь никакие другие известные алгоритмы.
4. Предложенные алгоритмы восстановления матриц и построения неотрицательных аппроксимаций асимптотически быстрее версий, основанных на других методах приближенного вычисления сингулярного разложения.

Все алгоритмы доступны в GitHub: Projective-volume-low-rank

Публикации

1. Замарашкин Н. Л., **Осинский А. И.** Новые оценки точности псевдоскелетных аппроксимаций матриц // Доклады академии наук. 2016. Т. 471, № 3. С. 263-266.
2. **Osinsky A. I.**, Zamarashkin N. L. Pseudo-skeleton approximations with better accuracy estimates // Linear Algebra and its Applications. – 2018. – V. 537, no. 4. – P. 221-249.
3. Замарашкин Н. Л., **Осинский А. И.** О существовании близкой к оптимальной скелетной аппроксимации матрицы во фробениусовой норме // Доклады Академии наук. — 2018. — Т. 479, № 5. — С. 489–492.
4. **Осинский А. И.**, Оценки аппроксимации тензорных поездов по норме Чебышёва // Журнал вычислительной математики и математической физики. — 2019. — Т. 59, № 2. — С. 211–216.
5. Zheltkov D. A., **Osinsky A. I.** Global Optimization Algorithms Using Tensor Trains. In: Lirkov, I., Margenov, S. (eds) Large-Scale Scientific Computing. LSSC 2019. Lecture Notes in Computer Science. — 2020. — V. 11958. Springer, Cham.
6. Замарашкин Н. Л., **Осинский А. И.** О точности крестовых и столбцовых малоранговых maxvol-приближений в среднем // Журнал вычислительной математики и математической физики. — 2021. — Т. 61, № 5. — С. 813–826.
7. Лебедева О. С., **Осинский А. И.**, Петров С. В. Приближенные алгоритмы малоранговой аппроксимации в задаче восполнения матрицы на случайном шаблоне // Журнал вычислительной математики и математической физики. — 2021. — Т. 61, № 5. — С. 827–844.

Публикации

8. Kalinov A., **Osinsky A. I.**, Matveev S. A., Otieno W., Brilliantov N. V. Direct simulation Monte Carlo for new regimes in aggregation-fragmentation kinetics // Journal of Computational Physics. — 2022. — V. 467. — P. 111439.
9. **Osinsky A. I.**, Brilliantov N. V. Exact solutions of temperature-dependent Smoluchowski equations // Journal of Physics A: Mathematical and Theoretical. — 2022. — V. 55, no. 42. — P. 425003.
10. **Osinsky A. I.** Volume-based subset selection // Numerical Linear Algebra with Applications. — 2023. — V. 31, no. 1. — P. e2525.
11. **Осинский А. И.** Нижние оценки точности столбцовых аппроксимаций матриц // Журнал вычислительной математики и математической физики. — 2023. — Т. 63, № 11. — С. 1816.
12. **Osinsky A. I.** Low-rank Monte Carlo for Smoluchowski-class equations // Journal of Computational Physics. — 2024. — V. 506, no. 1. — P. 112942.