Alternating minimization method for low-rank approximation of matrices and tensors in the Chebyshev norm

Stanislav Morozov [1]

[1]Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences

Introduction

Low-rank approximation algorithms are a crucial component in modern computations. Nowadays, most of the methods tackle the problem of low-rank matrix approximation in the unitary invariant norms. For such norms there exist efficient algorithms, e.g. SVD, that provide the optimal approximation. The quality of the approximation for the unitary invariant norms is related to the decay rate of singular values. However, in some applications, matrices arise that can be successfully approximated by low-rank structures in other norms, independently of the singular values decay rate. We address the problem of building low-rank approximations in the Chebyshev norm.

The best uniform approximation

Let $V \in \mathbb{R}^{n\times r}$, where $n \geq r$ and $a \in \mathbb{R}^n$. The problem

$$\|Vu - a\|_\infty \to \min_{u\in\mathbb{R}^r}$$

is called *the best uniform approximation* problem.

Definition. A matrix $V \in \mathbb{R}^{n\times r}$ with $n \geq r$ is called *Chebyshev* if all its $r \times r$ submatrices are non-singular.

Theorem. Let $V \in \mathbb{R}^{n\times r}$ be a Chebyshev matrix and $a \in \mathbb{R}^n$. Then the solution to the best uniform approximation problem exists, is unique and continuously depends on the matrix V and right-hand side a.

Theorem. Let us consider the problem $\|Vu - a\|_\infty \to \min_{u\in\mathbb{R}^r}$ with a Chebyshev matrix $V \in \mathbb{R}^{n\times r}$ and a vector $a \in \mathbb{R}^n$ that does not belong to the image of V. Let $\hat{u} \in \mathbb{R}^r$. Let us denote the residual by $w = a - V\hat{u}$. Then $\hat{u}$ is the solution of the best uniform approximation problem if and only if there is a set of integers $1 \leq j_1 < j_2 < \cdots < j_{r+1} \leq n$ such that

$$|w_{j_1}| = |w_{j_2}| = \cdots = |w_{j_{r+1}}| = \|w\|_\infty$$

and the signs in the sequence

$$w_{j_1}\Delta_1, w_{j_2}\Delta_2, \ldots, w_{j_{r+1}}\Delta_{r+1}$$

alternate, where $\Delta_k = \det V((j_1, \ldots, j_{k-1}, j_{k+1}, \ldots, j_{r+1}))$.

Algorithm 1 Best uniform approximation algorithm.

Require: Chebyshev matrix $V \in \mathbb{R}^{n\times r}$, right-hand side $a \in \mathbb{R}^n$, initial ordered set $\hat{J}$.
Ensure: Solution $\hat{u} \in \mathbb{R}^r$ to the problem $\|Vu - a\|_\infty \to \min_{u\in\mathbb{R}^r}$, characteristic set $\hat{J}$.

$\hat{V} = V(\hat{J}), \hat{a} = a(\hat{J}), t = 1$
$\hat{Q}, \hat{q}', \hat{R} \leftarrow$ `qr_decomposition`$(\hat{V})$
$\hat{u} \leftarrow$ `uniform_approximation`$(\hat{Q}, \hat{q}', \hat{R}, \hat{a})$
$w = a - V\hat{u}$
while $\|w(\hat{J})\|_\infty < \|w\|_\infty$ **do**
  $\hat{j} \leftarrow \arg\max_{j\in\{1,\ldots,n\}} |w_j|$
  $\hat{k} \leftarrow$ `best_replacement`$(\hat{Q}, \hat{q}', \hat{R}, v^{\hat{j}}, a_{\hat{j}})$
  $\hat{a}_{\hat{k}} = a_{\hat{j}}$
  Replace $\hat{k}$-th row of the matrix $\hat{V}$ with $v^{\hat{j}}$
  Update QR decomposition factors $\hat{Q}, \hat{q}'$ and $\hat{R}$ for the matrix $\hat{V}$
  Replace $\hat{k}$-th element of the ordered set $\hat{J}$ with $\hat{j}$
  $\hat{u} \leftarrow$ `uniform_approximation`$(\hat{Q}, \hat{q}', \hat{R}, \hat{a})$
  $w = a - V\hat{u}$
end while

The complexity of the algorithm is $O(r^3 + Inr)$, where I is the number of iterations and can be estimated as $O(r^{1.5}\log n)$.

Low-rank approximation of matrices

Let $A \in \mathbb{R}^{m\times n}$ be a matrix and $r \in \mathbb{N}$. Let us consider the problem

$$\|A - UV^T\|_C \to \min_{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{n\times r}}.$$

This problem is difficult to solve directly, so to tackle it we assume that one of the matrices (U or V) is known. Then we can consider the problem $\|A - UV^T\|_C \to \min_{U\in\mathbb{R}^{m\times r}}$, which can be decomposed to the set of problems of the form $\|a - Vu\|_\infty \to \min_{u\in\mathbb{R}^r}$, where $V \in \mathbb{R}^{n\times r}$ and $a \in \mathbb{R}^n$. Let $V \in \mathbb{R}^{n\times r}$ be a Chebyshev matrix. Then there is a unique map ϕ such that $\phi(A, V)^i = \arg\min_{x\in\mathbb{R}^r} \|a^i - Vx\|_\infty$. Similarly, we define the map ψ, which provides the solution to the problem $\|A - UV^T\|_C \to \min_{V\in\mathbb{R}^{n\times r}}$.

Definition. Let $A \in \mathbb{R}^{m\times n}$. We say that the pair of sequences of Chebyshev matrices $\{U^{(t)} \in \mathbb{R}^{m\times r}\}_{t\in\mathbb{N}}$ and $\{V^{(t)} \in \mathbb{R}^{n\times r}\}_{t\in\mathbb{N}}$ is obtained by the *alternating minimization method* for the matrix A with the initial point $V^{(0)}$, where $V^{(0)} \in \mathbb{R}^{n\times r}$ is a Chebyshev matrix, if

$$\begin{cases} U^{(t)} = \phi(A, V^{(t-1)}), \\ V^{(t)} = \psi(A, U^{(t)}) \end{cases}$$

for all $t \in \mathbb{N}$.

Let matrices $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ be Chebyshev. Let us denote $G = A - UV^T$. We also denote

$$S(A, U, V) = \{(i, j) : |g_{ij}| = \|G\|_C\},$$

$$\mathcal{I}(A, U, V) = \{i : \exists j \text{ such that } (i, j) \in S(A, U, V)\},$$

$$\mathcal{J}(A, U, V) = \{j : \exists i \text{ such that } (i, j) \in S(A, U, V)\}.$$

Definition. Let $A \in \mathbb{R}^{m\times n}$ be a matrix and matrices $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ be Chebyshev. We say that the triple (A, U, V) possesses *a 2-way alternance of rank* r, if there is a non-empty set $\mathcal{A} \subset \{1, \ldots, m\} \times \{1, \ldots, n\}$ such that $\mathcal{A} \subset S(T, U, V)$ and if $(i, j) \in \mathcal{A}$, then there exist a set I of $r+1$ different indices $1 \leq i_1 < i_2 < \cdots < i_{r+1} \leq m$ such that $i \in I$ and a set J of $r+1$ different indices $1 \leq j_1 < j_2 < \cdots < j_{r+1} \leq n$ such that $j \in J$ with the following properties.

1. It holds $(i, j_1), (i, j_2), \ldots, (i, j_{r+1}), (i_1, j), (i_2, j), \ldots, (i_{r+1}, j) \in \mathcal{A}$.
2. The signs in the sequence $g_{ij_1}D_1(\mathcal{V}), g_{ij_2}D_2(\mathcal{V}), \ldots, g_{ij_{r+1}}D_{r+1}(\mathcal{V})$ and the signs in the sequence $g_{i_1 j}D_1(\mathcal{U}), g_{i_2 j}D_2(\mathcal{U}), \ldots, g_{i_{r+1} j}D_{r+1}(\mathcal{U})$ alternate, where $\mathcal{U} = U(I)$ and $\mathcal{V} = V(J)$.

(a) Original image (b) Approximation (c) Alternance set

Theorem. Let $A \in \mathbb{R}^{m\times n}$ be a matrix, $\operatorname{rank} A > r$. Let $\hat{U} \in \mathbb{R}^{m\times r}$ and $\hat{V} \in \mathbb{R}^{n\times r}$ be a solution to the problem $\|A - UV^T\|_C \to \min_{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{n\times r}}$. Let $\hat{V}$ be Chebyshev and the alternating minimization method for the matrix A and the initial point $V^{(0)} = \hat{V}$ be correct. Then the triple $(A, \hat{U}, \hat{V})$ possesses a 2-way alternance of rank r.

Theorem. Let $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank} A > r$ and the matrix $V \in \mathbb{R}^{n\times r}$ be Chebyshev. Let the alternating minimization method for the matrix A and the initial point $V^{(0)} = V$ be correct and the sequences $\{U^{(t)}\}_{t\in\mathbb{N}}$ and $\{V^{(t)}\}_{t\in\mathbb{N}}$ be constructed by the alternating minimization method. Let a limit point Ξ of the sequence Ξ_t, where $\Xi_t = V^{(t)}/\|V^{(t)}\|_C$ be Chebyshev. Then $(A, \phi(A, \Xi), \Xi)$ possesses a 2-way alternance of rank r.

Let $v \in \mathbb{R}^n$ be a Chebyshev vector. Let $\mathcal{S}(v)$ denote the vector with components $\mathcal{S}(v)_i = \operatorname{sign}(v_i)$, $i = 1, \ldots, n$.

Theorem. For almost all matrices $A \in \mathbb{R}^{m\times n}$, if $v_1, v_2 \in \mathbb{R}^n$ are Chebyshev and $\mathcal{S}(v_1) = \mathcal{S}(v_2)$, then

$$\mathcal{S}(\phi(A, v_1)) = \mathcal{S}(\phi(A, v_2))$$

Similarly, if $u_1, u_2 \in \mathbb{R}^m$ and $\mathcal{S}(u_1) = \mathcal{S}(u_2)$, then

$$\mathcal{S}(\psi(A, u_1)) = \mathcal{S}(\psi(A, u_2)).$$

Low-rank approximation of tensors

Let $T \in \mathbb{R}^{m\times n\times k}$. The problem of low-rank tensor approximation in the Chebyshev norm formulates as follows:

$$\left\| T - \sum_{t=1}^{r} u_t \otimes v_t \otimes w_t \right\| \to \min_{u_t, v_t, w_t}.$$

If matrices U and V are known, then W can be found as

$$W = \arg\min_{X\in\mathbb{R}^{k\times r}} \|(U \odot V)X^T - T^{(1,2)}\|_C.$$

It can be easily seen that this problem is broken into the set of independent subproblems

$$w^l = \arg\min_{x\in\mathbb{R}^r} \|(U \odot V)x - \operatorname{vec}(T[:, :, l])\|_\infty,$$

where $U \odot V = \begin{bmatrix} u_1 \otimes v_1 & \ldots & u_r \otimes v_r \end{bmatrix} \in \mathbb{R}^{mn\times r}$ is the Khatri–Rao product of matrices U and V. Then we can define the mapping χ, which gives the optimal solution of W for known U and V. Similarly, we can define ϕ and ψ, which provide the optimal for U and V for known other two matrices.

Numerical evaluation

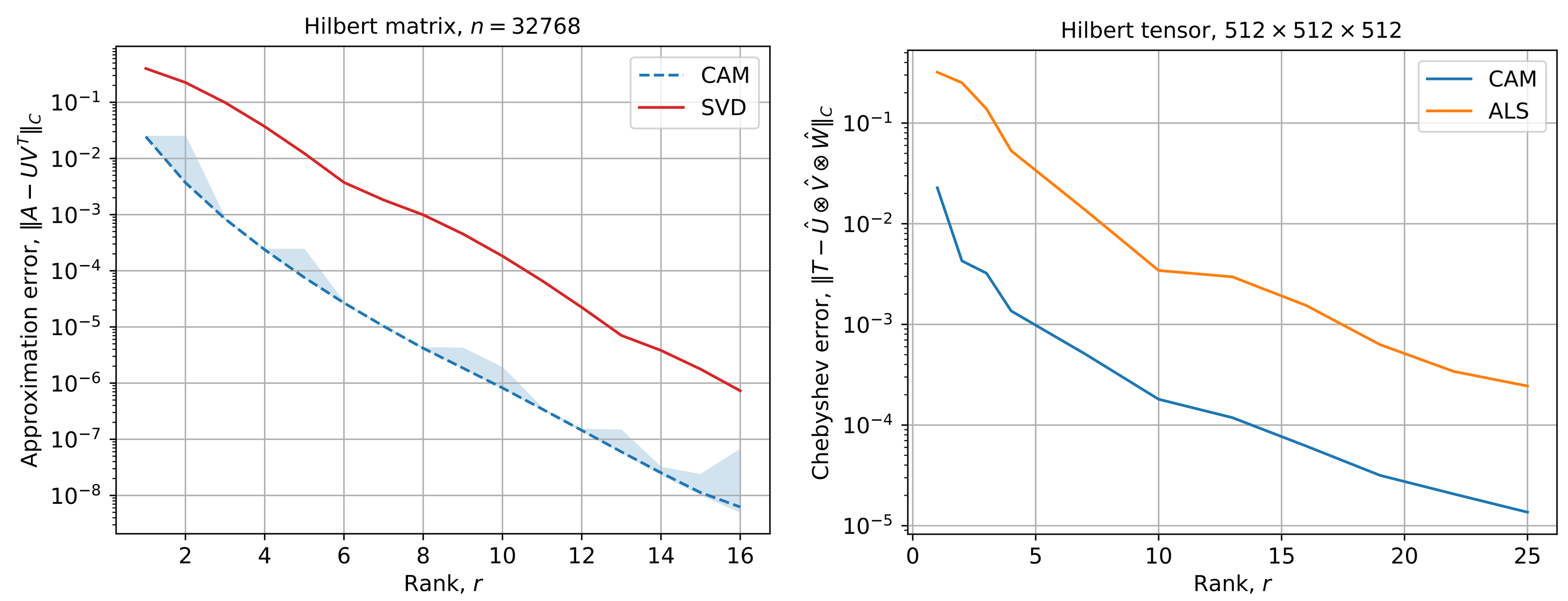

Figure 2. Approximation error for the Hilbert matrix and tensor.

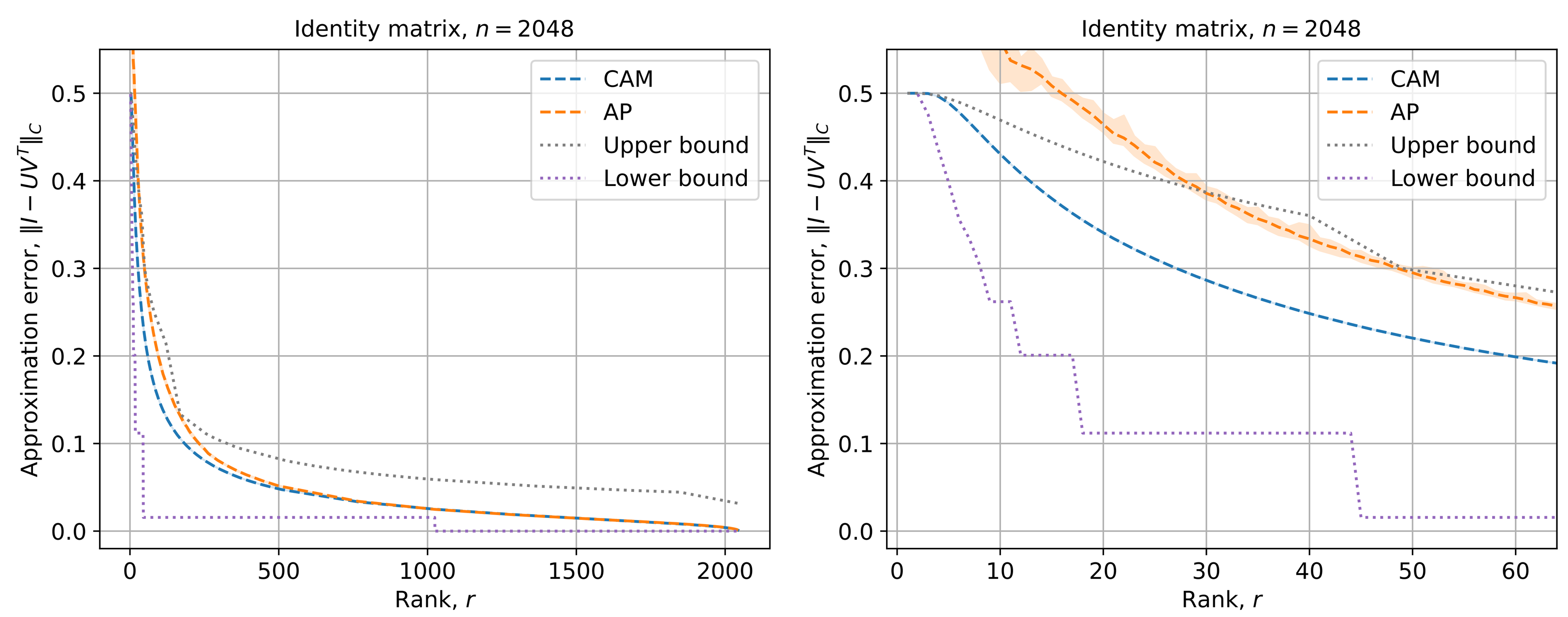

Figure 3. Approximation error of the identity matrix.

References

[1] Stanislav Morozov, Matvey Smirnov, and Nikolai Zamarashkin.
On the optimal rank-1 approximation of matrices in the chebyshev norm.
Linear Algebra and its Applications, 679:4–29, 2023.

[2] Stanislav Morozov, Dmitry Zheltkov, and Alexander Osinsky.
Accelerated alternating minimization algorithm for low-rank approximations in the chebyshev norm.
arXiv preprint arXiv:2410.05247, 2024.

[3] Stanislav Morozov, Dmitry Zheltkov, and Alexander Osinsky.
Refining uniform approximation algorithm for low-rank chebyshev embeddings.
Russian Journal of Numerical Analysis and Mathematical Modelling, 39(5):311–328, 2024.

[4] NL Zamarashkin, SV Morozov, and EE Tyrtyshnikov.
On the best approximation algorithm by low-rank matrices in Chebyshev's norm.
Computational Mathematics and Mathematical Physics, 62(5):701–718, 2022.