Matrix Equations and Tensor Techniques
IX
Perugia, September 9-10, 2021

On the low-rank approximations in the Chebyshev norm

Stanislav Morozov Nikolai Zamarashkin Eugene Tyrtyshnikov

Marchuk Institute of Numerical Mathematics
of the Russian Academy of Sciences, Moscow

Low-rank matrices and tensors are ubiquitous in science. A lot of methods have been developed using low-rank structures in computational mathematics [1], computational fluid dynamics [2], movie preferences [3] and automated machine learning [4]. To date, most of the methods have been developed to build effective low-rank approximations in the Frobenius norm. In the matrix case, the quality of such approximations depends on the decrease rate of the singular values of the matrix. However, recent results show that low-rank approximations of matrices in other norms can be effective even without decreasing singular values. One fundamental result was proved in [5]:

Theorem. Let $X \in \mathbb{R}^{m \times n}$ with $m \geq n$ and $0 < \varepsilon < 1$. Then, with

$$r = \lceil 72 \log(2n+1)/\varepsilon^2 \rceil \tag{1}$$

we have

$$\inf_{\text{rank}Y \leq r} \|X - Y\|_C \leq \varepsilon \|X\|_2, \quad \text{where} \quad \|X\|_C = \max_{i,j} |X_{ij}| \tag{2}$$

This work is devoted to low-rank approximations in the Chebyshev, that is, elementwise norm (2). For simplicity of presentation, the results are given in the matrix case, although some of them can be generalized to tensors. Let a matrix $A \in \mathbb{R}^{m \times n}$ be given. We call

$$\mu = \inf_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{n \times r}} \left\| A - UV^T \right\|_C, \tag{3}$$

the value of the best Chebyshev approximation of rank r. To date, the problem of constructing low-rank approximations in the Chebyshev norm has been little studied. One of the few works devoted to this topic is [6], in which a method for finding local minima for rank 1 approximations has been proposed. The authors construct an approximation using the alternance method. Let $U^{(0)}$ be given. In the case of rank 1, it is easy to find a matrix (in fact, a vector) $V^{(0)}$, being the solution of the problem

$$\mu = \min_{V \in \mathbb{R}^{n \times r}} \left\| A - U^{(0)} V^T \right\|_C \tag{4}$$

Then the authors find the matrix (vector) $U^{(1)}$ that is optimal for $V^{(0)}$ and repeat the procedure.

The results presented in [6] can be generalized to the case of an arbitrary rank. The key step here is the possibility of solving the problem of the form (4), which becomes challenging for $r > 1$. This problem can be reduced to solving several independent problems of the form

$$\mu = \inf_{v \in \mathbb{R}^r} \|a - Uv\|_\infty, \quad a \in \mathbb{R}^m, \quad U \in \mathbb{R}^{m \times r} \tag{5}$$

Suppose that in the matrix $U \in \mathbb{R}^{m \times r}$ all submatrices of size $r \times r$ are non-singular. Let I_k be a subset of k indices from 1 to m, $I_k = \{i_1, i_2, \dots, i_k\}$. Let us denote a_{I_k} the subvector of the vector a with elements from I_k, and U_{I_k} the submatrix of the matrix U with rows from I_k. It can be shown that there exists a subset I_{r+1} of $r+1$ indices such that the solution of the problem

$$\widehat{\mu} = \inf_{v \in \mathbb{R}^r} \left\|a_{I_{r+1}} - U_{I_{r+1}} v\right\|_\infty \tag{6}$$

coincides with the solution of the problem (5). Such a subset I_{r+1} is called characteristic. The problem of size $r+1$ can be solved exactly in $O(r^4)$ operations ([7]). The above arguments allow us to find the optimal solution to the problem (5) (and therefore (4)) by iterating over all subsets of $r+1$ indices. However, more efficient methods for finding the characteristic set and solving the problem (4) can also be constructed based on generalizations of the Remez algorithm. Such algorithm does not require iterating over all subsets and in practice work in polynomial time.

This work has been supported by Russian Science Foundation Project (21-71-10072).

References

[1] M. Bebendorf, *Hierarchical Matrices: A Means to Efficiently Solve Elliptic Boundary Value Problems*, Lect. Notes Comput. Sci. Eng. 63, (2008)

[2] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher, *More Practical Sketching Algorithms for Low-Rank Matrix Approximation*, Tech. rep. 2018-01, California Institute of Technology, (2018)

[3] R. M. Bell and Y. Koren, *Lessons from the Netflix prize challenge*, ACM SIGKDD Explorations Newsletter, 9, (2007)

[4] C. Yang, Y. Akimoto, D. W. Kim, and M. Udell, *OBOE: Collaborative filtering for automl initialization*, in NIPS Workshop on Automated Machine Learning, (2018)

[5] M. Udell, A. Townsend, *Why are big data matrices approximately low rank?*, SIAM Journal on Mathematics of Data Science, (2019)

[6] V. A. Daugavet, *The uniform approximation of a function of two variables that is given in tabular form by a product of functions of one variable*, U.S.S.R. Comput. Math. Math. Phys., (1971, in Russian)

[7] V. K. Dzyadyk, *On the approximation of functions on sets consisting of a finite number of points*, The theory of approximation of functions and its applications, (1974, in Russian)