Linear Algebra and its Applications 537 (2018) 221–249

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Pseudo-skeleton approximations with better accuracy estimates

A.I. Osinsky [a,c], N.L. Zamarashkin [a,b,c,*]

[a] Institute of Numerical Mathematics RAS, Moscow, Gubkina str., 8, Russia
[b] Moscow State University, Moscow, Leninskie Gory, 1, Russia
[c] Moscow Institute of Physics and Technology, Dolgoprudny, Institutsky per., 9, Russia

ARTICLE INFO

Article history:
Received 29 July 2016
Accepted 29 September 2017
Available online 2 October 2017
Submitted by V. Mehrmann

MSC:
65F30
65F99
65D05

Keywords:
Low rank approximations
Pseudoskeleton approximations
Maximum volume principle

ABSTRACT

We propose a priori accuracy estimates for low-rank matrix approximations that use just a small number of the rows and columns. This number is greater than the approximation rank, unlike the existing methods of pseudo-skeleton approximation. But the estimates are more accurate than previously known ones. This paper generalizes the results of [12,13].

1. Introduction

In the modern computational mathematics, the low-rank approximations of the dense matrices built using a small number of their elements play an important role as effective algorithms for working with very large dense matrices [1–5] and as a part of so called

* Corresponding author.
E-mail addresses: sasha_o@list.ru (A.I. Osinsky), nikolai.zamarashkin@gmail.com (N.L. Zamarashkin).

https://doi.org/10.1016/j.laa.2017.09.032

cross approximation methods of low-parametrical approximations of multidimensional data (see, e.g. [6–9]). The availability of reliable a priori accuracy estimates for low-rank matrix approximations play a crucial role in the justification of the developed algorithms.

The problem of theoretical justification for such methods becomes significantly more complicated while moving from matrices to multidimensional matrices (tensors), i.e. from 2 to 3 and more dimensions. The known accuracy estimates for the low-parametrical approximations of multidimensional data can hardly be considered satisfactory [6–8]. One of the major obstacles to improve them are additional factors in the elementwise estimates for the low-rank approximation of matrices. As a rule, these factors increase with the growth of the approximation rank [12,13].

Consider a matrix $A \in \mathbb{C}^{M \times N}$, where $A = Z + F$ with rank $Z = r$ and $\|F\|_2 \leqslant \varepsilon$. Let us choose m columns $C \in \mathbb{C}^{M \times m}$ and n rows $R \in \mathbb{C}^{n \times N}$ in the matrix A. We are interested in the accuracy estimates for the best or nearly the best rank r approximations of the matrix A by its CGR approximation with some kernel $G \in \mathbb{C}^{n \times m}$ (see, e.g. [11–13]).

Apparently, the first systematic study of such matrix approximations was conducted by Gu, Eisenstat in [10] and Goreinov, Tyrtyshnikov, and Zamarashkin in [11,12]. The main difference is that Gu and Eisenstat supposed the use of all matrix elements in their algorithm, while Goreinov, Tyrtyshnikov, and Zamarashkin constructed the CGR approximation assuming available just a small part of matrix elements. They proved the possibility of using CGR approximations for the particular case $m = n = r$ and obtained several estimates in the spectral matrix norm. In addition, the relation between the quasi-optimal pseudo-skeleton approximations and the maximum volume submatrices was established. By definition, the volume of a square $r \times r$ submatrix $\hat{A}$ is $\mathcal{V}_2(\hat{A}) = |\det(\hat{A})|$.

Developing these ideas Goreinov and Tyrtyshnikov proved the following theorem.

Theorem 1. [13] *Consider a matrix A in the block form*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{r \times r}$ has the maximum volume among all $r \times r$ submatrices in A. Then the following element wise accuracy estimate holds:

$$\left\| A_{22} - A_{21} A_{11}^{-1} A_{12} \right\|_C \leqslant (r+1)\, \sigma_{r+1}(A)\,, \tag{1}$$

where $\sigma_{r+1}(A)$ is the $(r+1)$th singular value of A.

One can easily see that Theorem 1 demonstrates a quasi-optimal accuracy of CGR approximations. But the presence of the factor $r+1$ in (1) devalues the result for deriving practically useful accuracy estimates for multidimensional arrays: in the d-dimensional case a factor like $(r+1)^{d-1}$ arises (see, e.g. [7]) and seems to be improved to