Journal of Computational and Applied Mathematics 317 (2017) 362–370

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Numerical CP decomposition of some difficult tensors

Petr Tichavský [a,*], Anh-Huy Phan [b], Andrzej Cichocki [b]

[a] Institute of Information Theory and Automation, Prague 182 08, Czech Republic
[b] Brain Science Institute, RIKEN, Wakoshi, Japan

ARTICLE INFO

Article history:
Received 24 February 2016
Received in revised form 24 August 2016

Keywords:
Small matrix multiplication
Canonical polyadic tensor decomposition
Levenberg–Marquardt method

ABSTRACT

In this paper, a numerical method is proposed for canonical polyadic (CP) decomposition of small size tensors. The focus is primarily on decomposition of tensors that correspond to small matrix multiplications. Here, rank of the tensors is equal to the smallest number of scalar multiplications that are necessary to accomplish the matrix multiplication. The proposed method is based on a constrained Levenberg–Marquardt optimization. Numerical results indicate the rank and border ranks of tensors that correspond to multiplication of matrices of the size 2×3 and 3×2, 3×3 and 3×2, 3×3 and 3×3, and 3×4 and 4×3. The ranks are 11, 15, 23 and 29, respectively. In particular, a novel algorithm for computing product of matrices of the sizes 3×4 and 4×3 using 29 multiplications is presented.

1. Introduction

The problem of determining the complexity of matrix multiplication became a well studied topic since the discovery of the Strassen's algorithm [1]. The Strassen's algorithm allows multiplying 2×2 matrices using seven multiplications. A consequence of this algorithm is that $n \times n$ matrices can be multiplied by performing of the order $n^{2.81}$ operations. More recent advances have brought the number of operations needed even closer to the n^2 operations. The current record is $O(n^{2.373})$ operations due to Williams [2].

The problem of the matrix multiplication can be rephrased as a problem of decomposing a particular tensor according to its rank [3]. In short, consider an order-3 tensor $\mathcal{T}$ of the size $I \times J \times K$ having elements $\mathcal{T}_{ijk}$ that admit a canonical polyadic (CP) decomposition

$$\mathcal{T}_{ijk} = \sum_{r=1}^{R} A_{ir} B_{jr} C_{kr} \tag{1}$$

where A_{ir}, B_{jr}, C_{kr} are elements of so-called factor matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, respectively. We shall use the symbolic notations of Kolda [4], $\mathcal{T} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$. Then, the smallest R such that a CP decomposition (1) exists, is called the tensor rank. A border rank $\underline{R}$ is defined as the smallest integer such that the given tensor $\mathcal{T}$ can be approximated to arbitrary precision by tensors of rank $\underline{R}$.

The lowest number of the scalar multiplications needed to compute the matrix product corresponds to the ranks of certain tensors. It is equivalent to solution to the so-called Brent equation [5]. The focus of this paper is not on improving the above asymptotic results of [2], but on numerical decomposition of tensors that correspond to multiplication of small matrices and determining their rank [6]. Although the problem is quite old, only partial results are known so far.

* Corresponding author.
E-mail address: tichavsk@utia.cas.cz (P. Tichavský).

http://dx.doi.org/10.1016/j.cam.2016.12.007

The matrix multiplication tensor for the 2 × 2 matrices is already completely clear [7]. It has rank 7 and border rank 7. For the 3 × 3 case, an algorithm using 23 scalar multiplications was found by Laderman [8]. It means that the rank is at most 23. For multiplying two 4 × 4 matrices, one can use twice the Strassen's algorithm, and therefore the rank is at most 49. Multiplication of 5 × 5 matrices was studied by Makarov [9] with the result of 100 multiplications (rank 100).

In this paper we present a numerical decomposition of the matrix multiplication tensors. For now, we are not able to improve the known results of Strassen, Laderman and Makarov, instead we show a method of the decomposition with these ranks and numerical results indicating that further improvements are probably not possible. Moreover, the numerical methods allow to guess the border rank of the tensors. As a new result, we have derived a novel algorithm for multiplying two matrices of the size 3 × 4 and 4 × 3 through 29 multiplications.

Traditional numerical tensor decomposition methods include the alternating least squares method (ALS) [10], improved ALS through the enhanced line search (ELS) [11], damped Gauss–Newton method, also known as Levenberg–Marquardt (LM) method [12], and different nonlinear optimization methods, e.g. [13]. For decomposition of the multiplication tensors we have developed a special variant of the constrained LM method. Once an exact representation is found, we propose a method of seeking another representation such that the factor matrices only contain nulls, ones and minus ones.

The rest of the paper is organized as follows. The tensors of the matrix multiplication are introduced in Section 2. The numerical method of their decomposition is presented in Section 3. Section 4 presents numerical results and Section 5 concludes the paper.

2. Tensor of matrix multiplication

Consider two matrices **E** and **F** of the sizes $P \times Q$ and $Q \times S$, respectively, and their matrix product $\mathbf{G} = \mathbf{EF}$ of the size $P \times S$. The operation of the matrix multiplication can be represented by a tensor $\mathcal{T}_{PQS}$ of the size $PQ \times QS \times PS$ which is filled with nulls and ones only, such that

$$\text{vec}(\mathbf{G}) = \mathcal{T}_{PQS} \times_1 \text{vec}(\mathbf{E}^T)^T \times_2 \text{vec}(\mathbf{F}^T)^T \tag{2}$$

regardless of the elements values of **E** and **F**. Here, $\times_i$ denotes a tensor–matrix multiplication along the dimension i, and the operator "vec" stacks all elements of a matrix or tensor in one long column vector.

Note that the number of ones in the tensor $\mathcal{T}_{PQS}$ is PQS; it is the number of scalar multiplications needed for evaluating the matrix product by a conventional matrix multiplication algorithm.

The tensor $\mathcal{T}_{PQS}$ has the elements

$$(\mathcal{T}_{PQS})_{\alpha\beta\gamma} = \delta_{in}\delta_{jk}\delta_{\ell m} \tag{3}$$

where $\alpha = (i-1)Q + j$; $\beta = (k-1)S + \ell$; $\gamma = (m-1)P + n$; $i, n = 1, \dots, P$; $j, k = 1, \dots, Q$; $\ell, m = 1, \dots, S$.

For example,

$$\mathcal{T}_{222} = \left(\begin{array}{cccc|cccc|cccc|cccc} 1&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&1 \end{array}\right). \tag{4}$$

This tensor has the size 4 × 4 × 4, and the vertical lines separate the four frontal slices of the tensor.

A canonical polyadic decomposition of the tensor $\mathcal{T}_{PQS}$ is a representation of the tensor as in (1), $\mathcal{T}_{PQS} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$. For example, a CP decomposition of the tensor $\mathcal{T}_{222}$ in (4) corresponding to the Strassen algorithm [2] is $\mathcal{T}_{222} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$ with

$$\mathbf{A} = \begin{pmatrix} 1&0&1&0&1&-1&0\\ 0&0&0&0&1&0&1\\ 0&1&0&0&0&1&0\\ 1&1&0&1&0&0&-1 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 1&1&0&-1&0&1&0\\ 0&0&1&0&0&1&0\\ 0&0&0&1&0&0&1\\ 1&0&-1&0&1&0&1 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 1&0&0&1&-1&0&1\\ 0&1&0&1&0&0&0\\ 0&0&1&0&1&0&0\\ 1&-1&1&0&0&1&0 \end{pmatrix}.$$

The multiplication tensors have the following properties:

1. Ranks of these tensors exceed the tensors' dimensions.
2. The CP decompositions are not unique.
3. The border ranks of the tensors might be strictly lower than their true ranks.