Journal of Computational and Applied Mathematics 308 (2016) 69–82

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Symmetric tensor decomposition by an iterative eigendecomposition algorithm

Kim Batselier *, Ngai Wong

Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

ARTICLE INFO

Article history:
Received 8 July 2015
Received in revised form 5 April 2016

MSC:
15A69
15A18
15A06

Keywords:
Symmetric tensor
Decomposition
Rank-1
Eigendecomposition
Least-squares

ABSTRACT

We present an iterative algorithm, called the symmetric tensor eigen-rank-one iterative decomposition (STEROID), for decomposing a symmetric tensor into a real linear combination of symmetric rank-1 unit-norm outer factors using only eigendecompositions and least-squares fitting. Originally designed for a symmetric tensor with an order being a power of two, STEROID is shown to be applicable to any order through an innovative tensor embedding technique. Numerical examples demonstrate the high efficiency and accuracy of the proposed scheme even for large scale problems. Furthermore, we show how STEROID readily solves a problem in nonlinear block-structured system identification and nonlinear state-space identification.

1. Introduction

Symmetric tensors arise naturally in various engineering problems. They are especially important in the problem of blind identification of under-determined mixtures [1–3]. Applications of this problem are found in areas such as speech, mobile communications, biomedical engineering and chemometrics.

The main contribution of this paper is an algorithm, called the **S**ymmetric **T**ensor **E**igen-**R**ank-**O**ne **I**terative **D**ecomposition (STEROID), that decomposes a real symmetric tensor $\mathcal{A}$ into a linear combination of symmetric unit-norm rank-1 tensors

$$\begin{aligned}\mathcal{A} &= l_1\, x_1 \circ x_1 \circ \cdots \circ x_1 + \cdots + l_R\, x_R \circ x_R \circ \cdots \circ x_R, \\ &= l_1\, x_1^d + \cdots + l_R\, x_R^d,\end{aligned} \tag{1}$$

with $l_1, \ldots, l_R \in \mathbb{R}$ and $x_1, \ldots, x_R \in \mathbb{R}^n$. The reality of the scalar coefficients $l_1, \ldots, l_R$ is of particular importance in the nonlinear system identification algorithm presented in Section 5. The $\circ$ operation refers to the outer product, which we define in Section 1.1. The notation x_i^d $(i = 1, \ldots, R)$ denotes the d-times repeated outer product. In contrast to other iterative methods, STEROID does not require any initial guess and, as shown in Section 4, can handle large symmetric tensors. The minimal $R = R_{\min}$ that satisfies (1) is called the symmetric rank of $\mathcal{A}$. More information on the rank of tensors can be found in [4,5] and specifically for symmetric tensors in [6]. The main idea of the algorithm is to first compute a set of vectors $x_1, \ldots, x_R$ $(R \geq R_{\min})$ through repeated eigendecompositions of symmetric matrices. The coefficients $l_1, \ldots, l_R$ are then

* Corresponding author.
E-mail address: kim.batselier@gmail.com (K. Batselier).

http://dx.doi.org/10.1016/j.cam.2016.05.024

found from solving a least-squares problem. STEROID was originally developed for symmetric tensors with an order that is a power of 2. It is however perfectly possible to extend the applicability of the STEROID algorithm to symmetric tensors of arbitrary order by means of an embedding procedure, which we explain in Section 2.2.

In [7] an algorithm is described that decomposes a symmetric tensor over $\mathbb{C}$ using methods from algebraic geometry. This involves computing the eigenvalues of commuting matrices and as a consequence, the l coefficients obtained from this method are generally complex numbers. Most attention in the literature is spent in solving the low-rank (typically rank-1) approximation problem. This problem can be formulated as follows.

Problem 1. Given a dth-order symmetric tensor $\mathcal{A} \in \mathbb{R}^{n\times\cdots\times n}$ and a multilinear rank r, find an orthogonal $n \times r$ matrix U and a core tensor $\mathcal{S} \in \mathbb{R}^{r\times\cdots\times r}$ that minimizes the Frobenius norm

$$\|\mathcal{A} - \mathcal{S} \times_1 U \times_2 U \times_3 \cdots \times_d U\|_F,$$

where $\times_k$ denotes the kth-mode product.

The definition of the kth-mode product is given in Section 1.1. Note that the Tucker form $\mathcal{S} \times_1 U \times_2 U \times_3 \cdots \times_d U$ is intrinsically different from (1), since it will also contain terms that are not symmetric. This implies that it is not very meaningful to compare the number of terms from the Tucker form with the number of terms computed by STEROID. Algorithms designed specifically for finding solutions to Problem 1 that consist of a single symmetric term are the symmetric higher-order power method (S-HOPM) [8,9] and the shifted version of S-HOPM (SS-HOPM) [10,11]. General low-rank algorithms are the Quasi-Newton algorithm [12], the Jacobi algorithm [13] and the monotonically convergent algorithm described in [14].

Another common decomposition is the canonical tensor decomposition (CANDECOMP/PARAFAC) [15,16]. This decomposition expresses a tensor as the sum of a finite number of rank-1 tensors. The tensor rank can then be defined as the minimum number of required rank-1 terms. Running a CANDECOMP algorithm such as Alternating Least Squares (ALS) on a symmetric tensor does not guarantee the symmetry of the rank-1 tensors. Other iterative methods [17], using nonlinear optimization methods, are able to guarantee the symmetry of the rank-1 terms. These methods however require the need for an initial guess and the number of computed terms also needs to be decided by the user beforehand. This is the main motivation for the development of the STEROID algorithm. STEROID is an adaptation for symmetric tensors of our earlier developed Tensor Train rank-1 SVD (TTr1SVD) algorithm [18], which in turn was inspired by Tensor Trains [19], and was an independent derivation of PARATREE [20]. In contrast to the iterative methods mentioned above, the STEROID algorithm does not require an initial guess and the total number of terms in the decomposition follows readily from the execution of the algorithm.

The outline of this paper is as follows. First, we define some basic notations in Section 1.1. In Section 2 we fully describe our algorithm by means of a running example, together with the required embedding procedure. Two methods for the reduction of the size of the least-squares problem in the STEROID algorithm are discussed in Section 3. One method exploits the symmetry of the tensor, while the other method exploits the structure of the matrix in the least-squares problem. The algorithm is applied to several examples in Section 4 and compared with the Jacobi algorithm [13], Regalia's iterative method described in [14] and the CANDECOMP-algorithm from the Tensorlab toolbox [17]. In Section 5 we show how STEROID readily solves a problem in nonlinear block-structured system identification [21] and nonlinear state-space identification [22]. In this setting, it is often desired to recover the internal structure of an identified static nonlinear mapping [23–25]. More specifically, it will be shown how STEROID can decouple a set of multivariate polynomials $f_1, \ldots, f_l$ into a collection of univariate polynomials $g_1, \ldots, g_n$, through both an affine and linear transformation.

1.1. Tensor notations and basics

We will adopt the following notational conventions. A dth-order or d-way tensor, assumed real throughout this article, is a multi-dimensional array $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ with elements $\mathcal{A}_{i_1 i_2 \ldots i_d}$ that can be seen as an extension of the matrix format to its general dth-order counterpart. Although the wordings 'order' and 'dimension' seem to be interchangeable in the tensor community, we prefer to call the number of indices i_k $(k = 1, \ldots, d)$ the order of the tensor, while the maximal value n_k $(k = 1, \ldots, d)$ associated with each index the dimension. A cubical tensor is a tensor for which $n_1 = n_2 = \cdots = n_d = n$. The inner product between two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ is defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1, i_2, \ldots, i_d} \mathcal{A}_{i_1 i_2 \ldots i_d} \mathcal{B}_{i_1 i_2 \ldots i_d}.$$

The norm of a tensor is often taken to be the Frobenius norm $\|\mathcal{A}\|_F = \langle \mathcal{A}, \mathcal{A} \rangle^{1/2}$. The kth-mode product of a tensor $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ with a matrix $U \in \mathbb{R}^{p\times n_k}$ is defined by

$$(\mathcal{A} \times_k U)_{i_1\cdots i_{k-1} j i_{k+1}\cdots i_d} = \sum_{i_k=1}^{n_r} U_{j i_k} \mathcal{A}_{i_1\cdots i_k\cdots i_d},$$

such that $\mathcal{A} \times_k U \in \mathbb{R}^{n_1\times\cdots\times n_{k-1}\times p\times n_{k+1}\times\cdots\times n_d}$. A 3rd-order rank-1 tensor $\mathcal{A}$ can always be written as the outer product [16]

$$\mathcal{A} = \lambda\, a \circ b \circ c \quad \text{with components } \mathcal{A}_{i_1 i_2 i_3} = \lambda\, a_{i_1} b_{i_2} c_{i_3}$$