ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report

Nr. 12.2016

March 2016

Structure-preserving low multilinear rank approximation of antisymmetric tensors

Erna Begović Kovač, Daniel Kressner

http://mathicse.epfl.ch

Address: EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Phone: +41 21 69 37648

Structure-preserving low multilinear rank approximation of antisymmetric tensors

Erna Begović Kovač* Daniel Kressner†

March 16, 2016

Abstract

This paper is concerned with low multilinear rank approximations to antisymmetric tensors, that is, multivariate arrays for which the entries change sign when permuting pairs of indices. We show which ranks can be attained by an antisymmetric tensor and discuss the adaption of existing approximation algorithms to preserve antisymmetry, most notably a Jacobi algorithm. Particular attention is paid to the important special case when choosing the rank equal to the order of the tensor. It is shown that this case can be addressed with an unstructured rank-1 approximation. This allows for the straightforward application of the higher-order power method, for which we discuss effective initialization strategies.

1 Introduction

A tensor $\mathcal{A} \in \mathbb{R}^{n\times\cdots\times n}$ of order $d \geq 2$ is called *antisymmetric* if its entries $\mathcal{A}(i_1, i_2, \ldots, i_d)$ change sign when permuting pairs of indices. For example, a tensor of order three with entries is antisymmetric if

$$\mathcal{A}(i_1, i_2, i_3) = -\mathcal{A}(i_2, i_1, i_3) = -\mathcal{A}(i_3, i_2, i_1) = -\mathcal{A}(i_1, i_3, i_2), \qquad i_1, i_2, i_3 = 1, \ldots, n.$$

For order two, the notion of antisymmetric tensors coincides of course with the notion of skew-symmetric matrices.

Antisymmetric tensors play a major role in quantum chemistry, where the Pauli exclusion principle implies that wave functions of fermions are antisymmetric under permutations of variables. This antisymmetry needs to be taken into account when solving the multiparticle Schrödinger equation determining such a wave function; see [11] for a recent overview.

This paper is concerned with finding an approximation $\mathcal{B}$ to a given antisymmetric tensor $\mathcal{A}$ such that $\mathcal{B}$ has a data-sparse representation and is again antisymmetric.

*Faculty of Chemical Engineering and Technology, University of Zagreb, 10000 Zagreb, Croatia. ebegovic@fkit.hr The author has been supported in part by Croatian Science Foundation under the project 3670.

†MATHICSE-ANCHP, EPF Lausanne, 1015 Lausanne, Switzerland. daniel.kressner@epfl.ch

More specifically, we will consider an approximation of multilinear rank r in structure-preserving Tucker decomposition

$$\mathcal{B} = \mathcal{S} \times_1 U \times_2 U \cdots \times_d U, \tag{1.1}$$

where $\mathcal{S} \in \mathbb{R}^{r \times \cdots \times r}$ for some $r \leq n$ is again antisymmetric and $U \in \mathbb{R}^{n \times r}$ has orthonormal columns. This choice is analogous to existing approaches for symmetric tensors, see, e.g., [3, 4, 8]. In this paper, we demonstrate that some existing algorithms for the symmetric case extend to the antisymmetric case. In particular, we study the extension of the Jacobi algorithm by Ishteva, Absil, and Van Dooren [7].

Despite a number of similarities, there are pronounced differences between symmetric and antisymmetric tensors. For example, every (multilinear) rank r can be attained by a symmetric matrix or tensor. In contrast, it is well known that skew-symmetric matrices have even rank. Although this statement does not extend to $d > 2$, we will see that there are still restrictions on the ranks that can be attained by anti-symmetric tensors. In particular, the smallest possible nonzero rank is $r = d$. In this case, the decomposition (1.1) simplifies to

$$\operatorname{anti}\big(\alpha u_1 \otimes u_2 \otimes \cdots \otimes u_d\big), \quad \alpha \in \mathbb{R}, \tag{1.2}$$

with the antisymmetrizer $\mathcal{A} = \operatorname{anti}(\mathcal{X})$ defined by

$$\mathcal{A}(i_1, \ldots, i_d) := \frac{1}{d!} \sum_{\pi \in S_d} \operatorname{sign}(\pi) \mathcal{X}\big(\pi(i_1), \pi(i_2), \ldots, \pi(i_d)\big), \tag{1.3}$$

where S_d denotes the symmetric group on $\{1, \ldots, d\}$. This corresponds to the notion of Slater determinants that feature prominently in the Hartree-Fock method from quantum mechanics. The expression (1.2) suggests the more general decomposition $\operatorname{anti}(\mathcal{X})$ for (non-symmetric) tensor $\mathcal{X}$ of low *tensor* rank. This corresponds to a short sum of Slater determinants used, e.g., in the Multi-Configuration Self-Consistent Field method. Such a low-rank model for antisymmetric tensors has been studied in the literature. In particular, Beylkin, Mohlenkamp, and Pérez [2, 1] have developed an alternating least-squares algorithm for approximating a given antisymmetric tensor $\mathcal{A}$ by $\operatorname{anti}(\mathcal{X})$. The algorithm employs Löwdin's rule to avoid having to deal with the exponentially many terms in the sum (1.3). One contribution of this paper is a much simpler approach for (1.2), that is, when $\mathcal{X}$ has rank 1: The best choice of $\mathcal{X}$ is given by a scalar multiple of the best (non-symmetric) rank-1 approximation of $\mathcal{A}$.

The rest of this paper is organized as follows. In Section 2, we study the multilinear rank of an antisymmetric tensor and recall the higher-order singular value decomposition. Section 3 is concerned with algorithms that aim at the antisymmetric low multilinear rank approximations, the higher-order iterations method and a variant of the Jacobi method. Section 4 is dedicated to the special case of rank-d approximation.

2 Multilinear rank of antisymmetric tensors

Let us first recall some basic concepts related to the multilinear rank of a tensor; see [9] for details. For any $1 \leq \mu \leq d$, the μth *matricization* of a general tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$

is the $n_\mu \times \prod_{\nu \neq \mu} n_\nu$ matrix $\mathbf{X}_{(\mu)}$ defined by

$$\mathbf{X}_{(\mu)}(i_\mu, j) = \mathcal{X}(i_1, \ldots, i_d), \qquad j = j(i_1, \ldots, i_d) := 1 + \sum_{\substack{\nu=1 \\ \nu \neq \mu}}^{d} (i_\nu - 1) \prod_{\substack{\eta=1 \\ \eta \neq \mu}}^{\nu-1} n_\eta. \tag{2.1}$$

The *multilinear rank* of $\mathcal{X}$ is the tuple $(r_1, r_2, \ldots, r_d)$ defined by $r_\mu = \operatorname{rank}(\mathbf{X}_{(\mu)})$. Note that $\mathbf{X}_{(\mu)}$ is a matrix and hence $r_\mu \leq \min\{n_\mu, \prod_{\nu \neq \mu} n_\nu\}$.

For an antisymmetric tensor, all matricizations are essentially the same.

Lemma 2.1 *Let $\mathcal{A} \in \mathbb{R}^{n \times n \times \cdots \times n}$ be an antisymmetric tensor of order d. Then $\boldsymbol{A}_{(\mu)} = (-1)^{|\mu-\nu|} \boldsymbol{A}_{(\nu)}$ holds for any $1 \leq \mu, \nu \leq d$.*

Proof. Without loss of generality, let $\mu \leq \nu$. According to (2.1), $\mathbf{X}_{(\mu)}(i_\mu, j) = \mathcal{X}(i_1, \ldots, i_d)$ implies $\mathbf{X}_{(\nu)}(i_\mu, j) = \mathcal{X}(i_1, \ldots, i_{\mu-1}, i_{\mu+1}, \ldots, i_\nu, i_\mu, i_{\nu+1}, \ldots, i_d)$. The result follows from the observation that the permutation $(1, \ldots, \mu-1, \mu+1, \ldots, \nu, \mu, \nu+1, \ldots, d)$ has sign $(-1)^{|\mu-\nu|}$. □

Lemma 2.1 implies that the multilinear rank of $\mathcal{A}$ always takes the form $(r, \ldots, r)$ for some $1 \leq r \leq n$. In the following, we will simply refer to r as the multilinear rank of an antisymmetric tensor.

2.1 Restrictions on the multilinear rank

It is well known that skew-symmetric matrices have even rank. It turns out that this property does not extend to antisymmetric tensors; it is simple to construct tensors of higher order with odd multilinear ranks. However, the following theorem shows that antisymmetry still imposes some (weaker) restrictions on the ranks of antisymmetric tensors that are of small size n relative to d.

Theorem 2.2 *Let $\mathcal{A} \in \mathbb{R}^{n \times n \times \cdots \times n}$ be an antisymmetric tensor of order $d \geq 3$. Then the multilinear rank r of $\mathcal{A}$ satisfies*

(i) $r = 0$ *for* $n < d$;

(ii) $r \leq d$ *for* $n = d$ *or* $n = d + 1$;

(iii) $r \leq n$ *for* $n \geq d + 2$.

There exist tensors $\mathcal{A}$ for which equality is attained in (i)–(iii).

Proof. We will make use of the fact that the entries of $\mathcal{A}$ satisfy

$$\mathcal{A}(i_1, i_2, \ldots, i_d) = 0, \quad \text{if } i_p = i_q, \text{ for some } p \neq q, \ 1 \leq p, q \leq d. \tag{2.2}$$

(i) According to (2.2), all d indices $i_1, \ldots, i_d \in [1, n]$ need to be mutually different for an entry $\mathcal{A}(i_1, i_2, \ldots, i_d)$ to be nonzero. When $n < d$, this is clearly not possible and hence $\mathcal{A} = 0$.

(ii) For $n = d$, the condition $r \leq d$ follows from the size of the matricizations. To show that equality is attained, consider the tensor $\mathcal{A} = d!\,\mathrm{anti}(\mathcal{X})$ where all entries of $\mathcal{X} \in \mathbb{R}^{d\times\cdots\times d}$ are zero except for $\mathcal{X}(1,2,\ldots,d) = 1$. For arbitrary $1 \leq i \leq d$ choose the permutation $p = (i, 1, \ldots, i-1, i+1, \ldots, d)$. By definition, $\mathcal{A}(p) = \mathrm{sign}(p) = (-1)^{i-1}$. By letting $j = j\big(p(2), \ldots, p(d)\big)$, it follows that the jth column of $\mathbf{A}_{(1)}$ equals $(-1)^{i-1}e_i$ with the ith unit vector $e_i \in \mathbb{R}^n$. In particular, $\mathbf{A}_{(1)}$ has d linearly independent columns and is thus of rank d.

Now, let $n = d+1$ and assume, without loss of generality, that $\mathcal{A} \neq 0$. We denote the rows of the matricization $\mathbf{A}_{(1)}$ by $\mathbf{A}_{1,(1)}, \ldots, \mathbf{A}_{d+1,(1)} \in \mathbb{R}^{n^{d-1}}$. This matricization has rank at most d if we can show that these rows are linearly dependent. Let

$$\alpha_k := \mathcal{A}(1, \ldots, k-1, k+1, \ldots, d+1), \quad k = 1, \ldots, d+1.$$

Since $\mathcal{A} \neq 0$, at least one α_k is different from zero. Let us now consider the column of $\mathbf{A}_{(1)}$ corresponding to a fiber $\mathcal{A}(:, i_2, \ldots, i_d)$ for some $i_2, \ldots, i_d \in [1, d+1]$. We may assume that $i_2, \ldots, i_d$ are mutually distinct because otherwise this fiber is zero. For the moment, we also assume that these indices are ordered, that is, $1 \leq i_2 < i_3 < \cdots < i_d \leq d+1$. By the pigeon hole principle, there are two integers $1 \leq k < \ell \leq d+1$ such that $k, l \notin \{i_2, \ldots, i_d\}$. The situation is now as follows:

$$\begin{array}{cccccccccccc}
i_2 & \cdots & i_k & & i_{k+1} & \cdots & i_{l-1} & & i_l & \cdots & i_d \\
\| & \cdots & \| & & \| & \cdots & \| & & \| & \cdots & \| \\
1 & \cdots & k-1 & k & k+1 & \cdots & l-1 & l & l+1 & \cdots & d+1
\end{array}$$

In particular, this implies

$$\begin{aligned}
\mathcal{A}(k, i_2, \ldots, i_d) &= (-1)^{k-1}\mathcal{A}(i_2, \ldots, i_{k-1}, k, i_{k+1}, \ldots, i_d) = (-1)^{k-1}\alpha_\ell, \\
\mathcal{A}(\ell, i_2, \ldots, i_d) &= (-1)^{\ell-2}\mathcal{A}(i_2, \ldots, i_{\ell-1}, \ell, i_{\ell+1}, \ldots, i_d) = (-1)^{\ell-2}\alpha_k.
\end{aligned}$$

Using that $\mathcal{A}(i_1, i_2, \ldots, i_d)$ is only nonzero for mutually distinct indices, we arrive at the linear combination

$$\begin{aligned}
\sum_{i_1=1}^{d+1}(-1)^{i_1}\alpha_{i_1}\mathcal{A}(i_1, i_2, \ldots, i_d) &= (-1)^k\alpha_k\mathcal{A}(k, i_2, \ldots, i_d) + (-1)^\ell\alpha_\ell\mathcal{A}(\ell, i_2, \ldots, i_d) \\
&= (-1)^{2k-1}\alpha_k\alpha_\ell + (-1)^{2\ell-2}\alpha_k\alpha_\ell = -\alpha_k\alpha_\ell + \alpha_k\alpha_\ell = 0.
\end{aligned}$$

Since this relation is not affected by a permutation of $i_2, \ldots, i_d$, it also holds if these indices are not ordered. In summary, we have shown that

$$\sum_{i_1=1}^{d+1}(-1)^{i_1}\alpha_{i_1}\mathbf{A}_{i_1,(1)} = 0$$

and thus the rank of $\mathbf{A}_{(1)}$ is at most d.

For $n = d+1$ equality is attained by the tensor used in the construction for $n = d$ bordered with zeros.

(iii) Let $n \geq d+2$. By the size of the matricization, $r \leq n$. To show that $r = n$ can be attained, let us first define the integer vector $h = (1, 2, \ldots, n, 1, \ldots, d-1)$. We choose the tensor $\mathcal{X} \in \mathbb{R}^{n \times n \times \cdots \times n}$ to be zero except for

$$\mathcal{X}(h_k, h_{k+1}, h_{k+2}, \ldots, h_{k+d-1}) = -d!, \qquad k = 1, 2, \ldots, n.$$

The corresponding sets $\sigma_k = \{h_k, h_{k+1}, h_{k+2}, \ldots, h_{k+d-1}\} \subset \mathbb{N}$ all have cardinality d for $k = 1, 2, \ldots, n$. The set $\{1, 3, 4, \ldots, d\} = \sigma_1 \setminus \{2\}$ is only contained in σ_1. In particular, $n > d \geq 3$ implies that it is not contained in $\sigma_n = \{n, 1, \ldots, d-1\}$. This shows $\mathcal{A}(:, 1, 3, 4, \ldots, d) = e_2$. Analogously, $\mathcal{A}(:, 2, 4, 5, \ldots, d+1) = e_3$ and $\mathcal{A}(:, 3, 5, 6, \ldots, d+2) = e_4$. This construction can be continued until we arrive at the set $\{n-1, 1, \ldots, d-2\} = \sigma_{n-1} \setminus \{n\}$, which is not contained in σ_1 because of $n - 1 > d$, or in any other σ_k except σ_{n-1}. Hence, $\mathcal{A}(:, n-1, 1, \ldots, d-2) = e_n$. Finally, we have $\mathcal{A}(:, n, 2, \ldots, d-2) = e_1$. In summary, we have found n linearly independent columns of $\mathbf{A}_{(1)}$ and, therefore, the multilinear rank of $\mathcal{A}$ is n.

□

2.2 HOSVD

Given a general tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, the higher-order singular value decomposition (HOSVD) introduced in [3] proceeds by computing the SVDs of the matricizations $\mathbf{X}_{(\mu)}$, $1 \leq \mu \leq d$, and letting $V_\mu \in \mathbb{R}^{n_\mu \times n_\mu}$ contain the left singular vectors. Setting $\mathcal{T} = \mathcal{X} \times_1 V_1^T \times_2 V_2^T \times_3 \cdots \times_d V_d^T$ yields the Tucker decomposition

$$\mathcal{X} = \mathcal{T} \times_1 V_1 \times_2 V_2 \cdots \times_d V_d.$$

The truncated HOSVD for a given multilinear rank $(r_1, \ldots, r_d)$ with $r_\mu \leq n_\mu$ is obtained by setting

$$\mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d, \tag{2.3}$$

with $U_\mu = V_\mu(:, 1 : r_\mu)$ and $\mathcal{S} = \mathcal{T}(1 : r_1, 1 : r_2, \ldots, 1 : r_\mu)$. This gives a quasi-best approximation of $\mathcal{X}$, in the sense that the approximation error in the Frobenius norm, $\|\mathcal{X} - \mathcal{S} \times_1 U_1 \cdots \times_d U_d\|$, is within a factor $\sqrt{d}$ of the error of the best rank-$(r_1, \ldots, r_d)$ approximation. In particular, if $\mathcal{X}$ happens to have multilinear rank $(r_1, \ldots, r_d)$ then the decomposition (2.3) is exact.

We now apply the truncated HOSVD to obtain an approximation of multilinear rank r to an antisymmetric tensor $\mathcal{A}$. By Lemma 2.1, all matrices U_μ in (2.3) can be chosen equal to a fixed matrix U. In turn $\mathcal{S} = \mathcal{A} \times_1 U^T \cdots \times_d U^T$ is again antisymmetric. In summary, the truncated HOSVD described in Algorithm 1 automatically preserves structure and produces a quasi-best antisymmetric approximation.

Algorithm 1 Truncated HOSVD of antisymmetric tensor

Compute matrix $U \in \mathbb{R}^{n \times r}$ containing the leading r left singular vectors of $\mathbf{A}_{(1)}$.
Set $\mathcal{S} = \mathcal{A} \times_1 U^T \cdots \times_d U^T$.
Return approximation $\mathcal{S} \times_1 U \cdots \times_d U$.

Corollary 2.3 *Let $\mathcal{A}$ be an antisymmetric tensor of order d. Then the multilinear rank r of $\mathcal{A}$ satisfies $r = 0$ or $r = d$ or $d + 2 \leq r \leq n$. Any of these ranks can be attained.*

Proof. By the discussion above, an antisymmetric tensor of multilinear rank r can be written as $\mathcal{A} = \mathcal{S} \times_1 U \cdots \times_d U$, where the $r \times \cdots \times r$ tensor $\mathcal{S}$ is again antisymmetric and has multilinear rank r. The statement of the corollary now follows from applying Theorem 2.2 to $\mathcal{S}$. □

Let us inspect the case $r = d$ more closely. Any antisymmetric $d \times \cdots \times d$ tensor of order d takes the form

$$\mathcal{S} = \text{anti}(\alpha e_1 \otimes e_2 \otimes \cdots \otimes e_d), \tag{2.4}$$

for some $\alpha \in \mathbb{R}$; see also the construction in the proof of Theorem 2.2 (i). By letting $U = [u_1, u_2, \ldots, u_d]$, the truncated HOSVD implies that any antisymmetric tensor of order d and multilinear rank d takes the form

$$\mathcal{A} = \text{anti}(\alpha e_1 \otimes e_2 \otimes \cdots \otimes e_d) \times_1 U \times_2 U \cdots \times_d U = \text{anti}(\alpha u_1 \otimes u_2 \otimes \cdots \otimes u_d),$$

verifying the claim (1.2) from the introduction.

3 Low multilinear rank approximation

In this section, we discuss two iterative methods that aim to compute a best antisymmetric multilinear rank-r approximation

$$\min \left\{ \|\mathcal{A} - \mathcal{S} \times_1 U \cdots \times_d U\| : \ \mathcal{S} \in \mathbb{R}^{r \times \cdots \times r} \text{ antisymmetric}, U \in \mathbb{R}^{n \times r} \right\},$$

starting, for example, from the truncated HOSVD of $\mathcal{A}$. Both methods are based on the fact that this minimization problem is equivalent to solving

$$\max \left\{ \|\mathcal{A} \times_1 U^T \cdots \times_d U^T\| : \ U \in \mathbb{R}^{n \times r} \text{ with } U^T U = I_r \right\} \tag{3.1}$$

and setting $\mathcal{S} = \mathcal{A} \times_1 U^T \cdots \times_d U^T$; see, for example, [4].

Remark 3.1 *For symmetric tensors, there is numerical evidence (see, e.g., [7]) that the best (unstructured) approximation of multilinear rank r can usually be chosen symmetric. For $d = 2$ and general r, this follows from the spectral decomposition. For general d and $r = 1$, this property has recently been shown by Friedland [5]. For general d and r, this question remains open.*

For antisymmetric tensors, we will observe the analogous phenomenon below; it appears that the best (unstructured) approximation of multilinear rank r can usually be chosen antisymmetric. For $d = 2$ and even r, this property follows from the real Schur decomposition. For $r = d$ and general d, we will see in Section 4 that it is actually the unstructured rank-1 approximation that gives an antisymmetric multilinear rank-d approximation.

To simplify the presentation, we will consider the case $d = 3$ for the rest of this section; all developments extend in a relatively straightforward manner to general $d > 3$.

3.1 HOOI

The higher-order orthogonal iteration (HOOI) introduced in [10] is a popular approach to the best low multilinear rank approximation of a general tensor. It consists of applying alternating least squares (ALS) to the unstructured variant of the maximization problem (3.1):

$$\max\big\{\|\mathcal{A}\times_1 U_1^T\times_2 U_2^T\times_3 U_3^T\|:\ U_\mu\in\mathbb{R}^{n\times r}\text{ with }U_\mu^TU_\mu=I_r,\mu=1,2,3\big\}.$$

One step of the method optimizes a single factor U_μ while keeping the other two factors fixed. The resulting optimization problem admits a straightforward solution by the SVD; see Algorithm 2.

Algorithm 2 HOOI for multilinear rank-(r,r,r) approximation

Apply Algorithm 1 to choose initial factors $U_1=U_2=U_3=U$.
repeat
  $\mathcal{X}=\mathcal{A}\times_2 U_2^T\times_3 U_3^T$
  Compute matrix $U_1\in\mathbb{R}^{n\times r}$ containing the leading r left singular vectors of $\mathbf{X}_{(1)}$.
  $\mathcal{Y}=\mathcal{A}\times_1 U_1^T\times_3 U_3^T$
  Compute matrix $U_2\in\mathbb{R}^{n\times r}$ containing the leading r left singular vectors of $\mathbf{Y}_{(2)}$.
  $\mathcal{Z}=\mathcal{A}\times_1 U_1^T\times_2 U_2^T$
  Compute matrix $U_3\in\mathbb{R}^{n\times r}$ containing the leading r left singular vectors of $\mathbf{Z}_{(3)}$.
until convergence
$\mathcal{S}=\mathcal{Z}\times_3 U_3^T$
Return approximation $\mathcal{S}\times_1 U_1\times_2 U_2\times_3 U_3$.

Note that the iterates of Algorithm 2 are *not* antisymmetric. However, similarly as in the symmetric case, we have observed that Algorithm 2 often converges towards an antisymmetric approximation; see Section 3.3 below. To antisymmetrize the output of Algorithm 2, one could set all factors equal to the factor U_μ that maximizes (3.1).

A simple antisymmetric variant of Algorithm 2 consists of setting *all* factors to the factor that has been obtained from the SVD in one step. In the symmetric case, this variant has been observed to suffer from convergence problems [7] and we observed similar difficulties in the antisymmetric case.

3.2 Jacobi algorithm

In contrast to HOOI, the Jacobi algorithm proposed for symmetric tensors in [7] preserves structure, that is, all iterates stay symmetric. In this section, we develop an extension of this algorithm to antisymmetric tensors.

It will be convenient to rewrite the maximization problem (3.1) as

$$\max\big\{f(Q):\ Q\in\mathbb{R}^{n\times n}\text{ with }Q^TQ=I_n\big\},$$

where

$$f(Q)=\|\mathcal{A}\times_1 MQ^T\times_2 MQ^T\times_3 MQ^T\|^2,\qquad M=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}. \tag{3.2}$$

We will denote a Givens rotation acting on rows/columns i and j by

$$R(i,j,\phi) = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} & & i & & j \end{array} \\ \begin{bmatrix} I & & & & \\ & \cos\phi & & -\sin\phi & \\ & & I & & \\ & \sin\phi & & \cos\phi & \\ & & & & I \end{bmatrix} \end{array} \begin{array}{c} \\ \\ i \\ \\ j \end{array}.$$

In the following, (i,j) will be called a pivot pair.

The main idea of the Jacobi algorithm is to repeatedly apply Givens rotations that increase the norm of the $(1:r,1:r,1:r)$ subtensor. For this purpose, it will be sufficient to consider rotations corresponding to the pivot pairs

$$\begin{array}{cccc} (1,r+1), & (1,r+2), & \dots & (1,n) \\ (2,r+1), & (2,r+2), & \dots & (2,n) \\ \vdots & \vdots & & \vdots \\ (r,r+1), & (r,r+2), & \dots & (r,n). \end{array} \tag{3.3}$$

In every iteration of the Jacobi algorithm, we choose a pivot pair that produces a direction of sufficiently strong descent. Letting

$$d_{ij} = \frac{\partial}{\partial\phi} R(i,j,\phi)\Big|_{\phi=0} = \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} & & i & & j \end{array} \\ \begin{bmatrix} I & & & & \\ & 0 & & -1 & \\ & & I & & \\ & 1 & & 0 & \\ & & & & I \end{bmatrix} \end{array} \begin{array}{c} \\ \\ i \\ \\ j \end{array},$$

we can always find pivot pairs (i,j) among (3.3) such that

$$|\langle \mathrm{grad} f(I), d_{ij}\rangle| \geq \epsilon \|\mathrm{grad} f(I)\| \tag{3.4}$$

holds, provided that $0 < \epsilon < 2/n$; see [7, Lemma 5.2].

Once a pivot pair (i,j) satisfying (3.4) is determined, we choose the rotation angle ϕ that maximizes f, i.e., we solve

$$\max\big\{ f\big(R(i,j,\phi)\big) : \ \phi \in [0,\pi] \big\}, \tag{3.5}$$

Because of $1 \leq i \leq r < j \leq n$, the tensor $\mathcal{B} = \mathcal{A} \times_1 R(i,j,\phi)^T \times_2 R(i,j,\phi)^T \times_3 R(i,j,\phi)^T$ differs from $\mathcal{A}$ within the subtensor $(1:r,1:r,1:r)$ only in the three slices $(i,1:r,1:r)$, $(1:r,i,1:r)$, and $(1:r,1:r,i)$. Because of antisymmetry, these slices have identical norms and we can ignore their intersections. Hence, (3.5) becomes equivalent

to maximizing

$$\begin{aligned}\|\mathcal{B}(i,1:r,1:r)\|^2 &= \sum_{p,q=1}^{r} \mathcal{B}(i,p,q)^2 = \sum_{\substack{p,q=1\\ p,q\neq i}}^{r} \mathcal{B}(i,p,q)^2 \\ &= \sum_{\substack{p,q=1\\ p,q\neq i}}^{r} \big(\cos\phi\mathcal{A}(i,p,q) + \sin\phi\mathcal{A}(j,p,q)\big)^2 =: \psi(\phi). \end{aligned} \tag{3.6}$$

Let

$$\alpha_1 = \sum_{\substack{p,q=1\\ p,q\neq i}}^{r} \mathcal{A}(i,p,q)^2,\ \alpha_2 = \sum_{\substack{p,q=1\\ p,q\neq i}}^{r} \mathcal{A}(i,p,q)\mathcal{A}(j,p,q),\ \alpha_3 = \sum_{\substack{p,q=1\\ p,q\neq i}}^{r} \mathcal{A}(j,p,q)^2.$$

Then the derivative of (3.6) takes the form

$$\psi'(\phi) = -2\alpha_1\cos\phi\sin\phi + 2\alpha_2(\cos^2\phi - \sin^2\phi) + 2\alpha_3\cos\phi\sin\phi.$$

In order to find the zeros of this function, we divide it by $\cos^2\phi$ and solve the resulting quadratic equation in $t = \sin\phi/\cos\phi$:

$$\alpha_2 t^2 + (\alpha_1 - \alpha_3)t - \alpha_2 = 0.$$

Among the two solutions to this equation, we choose the one that maximizes (3.6). Algorithm 3 summarizes the described procedure.

Algorithm 3 Jacobi algorithm for antisymmetric multilinear rank-r approximation

Apply Algorithm 1 to choose initial factor $U \in \mathbb{R}^{n\times r}$.
Choose $U_\perp$ such that $Q = [U, U_\perp]$ is orthogonal.
Set $\mathcal{A}_1 = \mathcal{A} \times_1 Q^T \times_2 Q^T \times_3 Q^T$.
repeat
  Choose (i,j) according to (3.3) and (3.4).
  Determine ϕ that maximizes (3.6).
  $Q_{k+1} = Q_k R(i,j,\phi)$
  $\mathcal{A}_{k+1} = \mathcal{A}_k \times_1 R(i,j,\phi)^T \times_2 R(i,j,\phi)^T \times_3 R(i,j,\phi)^T$
until convergence
$U = Q_k(:, 1:r)$
Return approximation $\mathcal{A} \times_1 UU^T \times_2 UU^T \times_3 UU^T$.

For choosing the pivot pair (i,j) in Algorithm 3, we traverse the list (3.3) cyclically. For each pair, the condition (3.4) is checked. If (i,j) does not fulfill this condition, it is skipped and the algorithm continues checking the next pair.

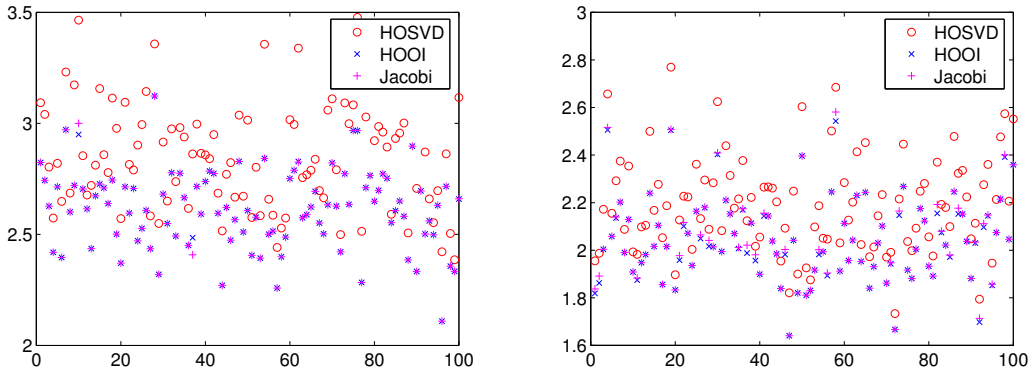
Although observed in practice, it cannot be guaranteed that Algorithm 3 produces the minimum of the function in (3.2). The proof of a weaker convergence result for symmetric tensors [7, Theorem 5.4] directly extends to antisymmetric tensors, resulting in the following statement.

Theorem 3.2 *Let (Q_k) be the sequence of orthogonal matrices generated by Algorithm 3 applied to an antisymmetric tensor $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$. Then every accumulation point of (Q_k) is a stationary point of the function f from* (3.2).

3.3 Numerical Experiments

The algorithms described in this sections have been implemented and tested in Matlab version 7.11.

In our first set of experiments, we study the approximation error obtained by truncated HOSVD, HOOI, and the Jacobi algorithm. The latter two algorithms are iterative; they are considered converged when the norm of the gradient of the objective function is 10^{-10} or below. We have chosen $\epsilon = 1/(10n)$ in the condition (3.4) of the Jacobi algorithm. We tested the algorithms with random tensors generated by applying antisymmetrizer from (1.3) to tensors with uniformly distributed random entries from the interval $[0, 1]$. Figure 1 shows that HOOI and the Jacobi algorithm always improve upon the approximation obtained from the HOSVD. In many cases, HOOI and the Jacobi algorithm result in the same (antisymmetric) approximation. In cases when the error of the Jacobi algorithm is smaller than the one of HOOI, it is observed that the tensor produced by HOOI is not antisymmetric. On the other hand, when the error of HOOI is smaller, the tensor produced by HOOI is antisymmetric. This leads us to conjecture that the best (unstructured) approximation of multilinear rank (r, r, r) to a generic antisymmetric tensor can always been chosen antisymmetric for $r \geq 3$.



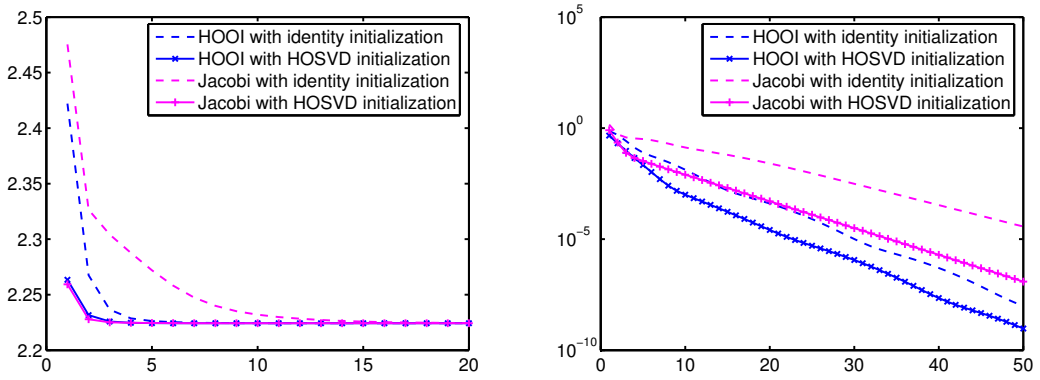
(a) Multilinear rank 3

(b) Multilinear rank 6

Figure 1: Approximation error of low multilinear rank approximation to 100 random antisymmetric $10 \times 10 \times 10$ tensors.

Figure 2 yields insights into the convergence behavior of HOOI and Jacobi algorithm for a representative run with a random antisymmetric tensor. To emphasize the benefits from initializing with the truncated HOSVD we compare with using no initialization, that is, instead of using Algorithm 1 we set $U = \begin{bmatrix} I_r \\ 0 \end{bmatrix}$ and $Q = I_n$ in Algorithms 2 and 3,

respectively. Apart from the approximation error we also show the norm of the gradient of the objective function.



(a) Approximation error

(b) Norm of gradient of objective function

Figure 2: Convergence behavior of HOOI and Jacobi algorithm for multilinear rank-6 approximation of random antisymmetric $10 \times 10 \times 10$ tensor.

We have also considered antisymmetric tensors for which the matrizations exhibit rapid singular value decays. To construct such a tensor, consider the function

$$f(x, y, z) = \exp(-\sqrt{x^2 + 2y^2 + 3z^2})$$

on $[0, 1]^3$. Then we let $\mathcal{X}$ contain its discretization:

$$\mathcal{X}(i_1, i_2, i_3) = f\big(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}\big), \quad i_\mu = 1, \ldots, n,$$

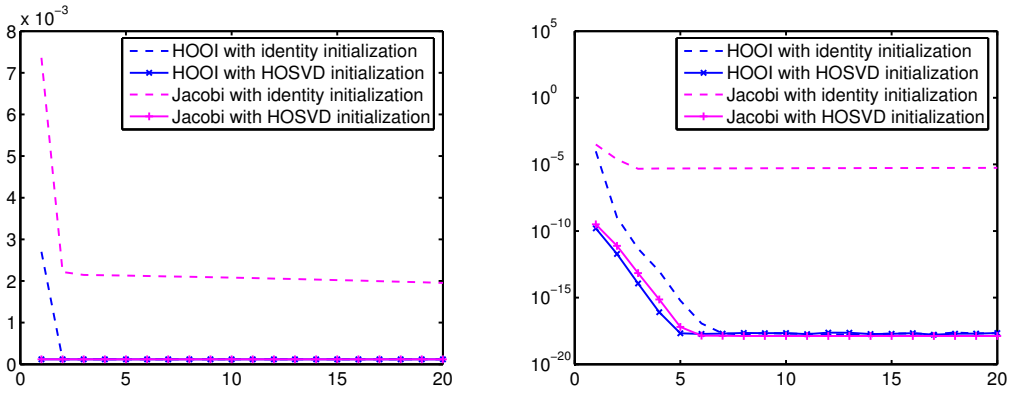
where $\xi_i = (i-1)/(n-1)$, and set $\mathcal{A} = \operatorname{anti}(\mathcal{X})$. Figure 3 shows the obtained results for $n = 20$. It reveals that the HOSVD gives an excellent initial approximation. This is also an example where the Jacobi algorithm with no initialization fails to converge to a global optimum.

4 Multilinear rank-d approximation

Antisymmetric tensors of order d and multilinear rank d have the very particular structure (1.2). As we will discuss in this section, this simplifies the approximation with such tensors significantly.

The following basic lemma plays a key role; it extends the well known fact that $u^T A u = 0$ always holds for a skew-symmetric matrix A.

Lemma 4.1 *Let $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$ be an antisymmetric tensor of order $d \geq 2$ and $u \in \mathbb{R}^n$. Then $\mathcal{A} \times_\mu u \times_\nu u = 0$ for any $1 \leq \mu < \nu \leq d$.*



(a) Approximation error

(b) Norm of gradient of objective function

Figure 3: Convergence behavior of HOOI and Jacobi algorithm for multilinear rank-7 approximation of function-related tensor.

Proof. Without loss of generality, we may assume that $\mu = d-1$ and $\nu = d$. Then any entry of $\mathcal{B} = \mathcal{A} \times_\mu u \times_\nu u$ satisfies

$$\begin{aligned} \mathcal{B}(i_1, \ldots, i_{d-2}) &= \sum_{j,k=1}^{n} \mathcal{A}(i_1, \ldots, i_{d-2}, j, k) u_j u_k \\ &= -\sum_{j,k=1}^{n} \mathcal{A}(i_1, \ldots, i_{d-2}, k, j) u_j u_k = -\mathcal{B}(i_1, \ldots, i_{d-2}), \end{aligned}$$

which implies $\mathcal{B} = 0$. □

The following theorem establishes an equivalence between the best antisymmetric multilinear rank-d approximation and the best unstructured rank-1 approximation of an antisymmetric tensor.

Theorem 4.2 *Let $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$ be an antisymmetric tensor of order d. Then*

$$\begin{aligned} & \max \left\{ \|\mathcal{A} \times_1 U^T \cdots \times_d U^T\| : \ U \in \mathbb{R}^{n \times d} \text{ with } U^T U = I_d \right\} \\ = \ & d! \max \left\{ |\mathcal{A} \times_1 u_1^T \cdots \times_d u_d^T| : \ [u_1, \ldots, u_d]^T [u_1, \ldots, u_d] = I_d \right\} \qquad (4.1) \\ = \ & d! \max \left\{ |\mathcal{A} \times_1 v_1^T \cdots \times_d v_d^T| : \ \|v_1\| = \cdots = \|v_d\| = 1 \right\}. \qquad (4.2) \end{aligned}$$

Proof. Let $\alpha = |\mathcal{A} \times_1 u_1^T \cdots \times_d u_d^T|$. Using (2.4),

$$\|\mathcal{A} \times_1 U^T \cdots \times_d U^T\|^2 = \|\operatorname{anti}(\alpha e_1 \otimes \cdots \otimes e_d)\|^2 = \alpha^2 \|\operatorname{anti}(e_1 \otimes \cdots \otimes e_d)\|^2 = (\alpha d!)^2,$$

which shows (4.1).

Consider vectors $v_1, \ldots, v_d$ assuming the maximum in (4.2). By the QR decomposition, there is an upper triangular matrix $R \in \mathbb{R}^{d \times d}$ with $|r_{\mu\mu}| \leq 1$ for $\mu = 1, \ldots, d$ and a matrix $[u_1, \ldots, u_d]$ with orthonormal columns such that

$$[v_1, \ldots, v_d] = [u_1, \ldots, u_d] R.$$

Using Lemma 4.1,

$$\begin{aligned}\mathcal{A} \times_1 v_1^T \cdots \times_d v_d^T &= \mathcal{A} \times_1 \sum_{\mu_1=1}^{d} r_{\mu_1,1} u_{\mu_1}^T \cdots \times_d \sum_{\mu_d=d}^{d} r_{\mu_d,d} u_{\mu_d}^T \\ &= \sum_{\mu_1=1}^{d} \cdots \sum_{\mu_d=d}^{d} \mathcal{A} \times_1 r_{\mu_1,1} u_{\mu_1}^T \cdots \times_d r_{\mu_d,d} u_{\mu_d}^T \\ &= \mathcal{A} \times_1 r_{11} u_1^T \cdots \times_d r_{dd} u_d^T = r_{11} \cdots r_{dd} \mathcal{A} \times_1 u_1^T \cdots \times_d u_d^T.\end{aligned}$$

Because of $|r_{\mu\mu}| \leq 1$, this implies $|\mathcal{A} \times_1 v_1^T \cdots \times_d v_d^T| \leq |\mathcal{A} \times_1 u_1^T \cdots \times_d u_d^T|$ and hence (4.2) cannot be larger than (4.1). On the other hand, trivially, (4.1) cannot be larger than (4.2). This shows the equality (4.2). □

4.1 HOPM

Algorithm 4 recalls the higher-order power method (HOPM) proposed in [4] for finding a rank-1 approximation of a tensor $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$.

Algorithm 4 HOPM for rank-1 approximation

Choose initial vectors $u_1, \ldots, u_d \in \mathbb{R}^n$.
repeat
  $v_1 = \mathcal{A} \times_2 u_2^T \times_3 u_3^T \cdots \times_d u_d^T$
  $u_1 = v_1 / \|v_1\|$
  $v_2 = \mathcal{A} \times_1 u_1^T \times_3 u_3^T \cdots \times_d u_d^T$
  $u_2 = v_2 / \|v_2\|$
  $\vdots$
  $v_d = \mathcal{A} \times_1 u_1^T \times_2 u_2^T \cdots \times_{d-1} u_{d-1}^T$
  $u_d = v_d / \|v_d\|$
until convergence
$\alpha = \mathcal{A} \times_1 u_1^T \cdots \times_d u_d^T$
Return approximation $\alpha u_1 \otimes u_2 \otimes \cdots \otimes u_d$.

Assuming that Algorithm 4 converges to a best rank-1 approximation with mutually orthogonal vectors u_μ, Theorem 4.2 allows us to construct the best antisymmetric multilinear rank-d approximation $\mathrm{anti}(\alpha u_1 \otimes \cdots \otimes u_d)$. The following lemma assures mutual orthogonality.

Lemma 4.3 *Let $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$ be an antisymmetric tensor of order $d \leq n$. Then the vectors $u_1, \ldots, u_d$ returned by HOPM form an orthonormal basis, provided that HOPM does not encounter a zero vector.*

Proof. After the first step of HOPM, Lemma 4.1 implies

$$\langle v_1, u_\nu \rangle = \mathcal{A} \times_1 u_\nu^T \times_2 u_2^T \times_3 u_3^T \cdots \times_d u_d^T = 0$$

for any $\nu \neq 1$. Hence, each step of HOPM orthogonalizes one of the vectors $u_\mu = v_\mu/\|v_\mu\|$. In turn, the statement of the lemma holds after at least one sweep of HOPM, even if the initial vectors are not orthogonal. □

Remark 4.4 *To ensure orthogonality numerically, we perform another orthogonalization step after each step of Algorithm 4. In principle, this procedure can also be applied to HOOI, yielding an (unstructured) multilinear rank $(r_1, \ldots, r_d)$ approximation with mutually orthogonal basis matrices U_μ. This can then be turned into an antisymmetric multilinear rank-$(r_1 + \cdots + r_d)$ approximation by setting $U = [U_1, \ldots, U_d]$. We have tested this idea numerically and observed that this often yields a good approximation but the approximation error is usually worse compared to the result of the Jacobi algorithm.*

4.2 Initialization

It remains to discuss a proper initialization strategy for HOPM. For general d, we use the truncated HOSVD from Section 3. For $d = 4$, we propose an antisymmetric variant of the technique proposed by Kofidis and Regalia [8] for symmetric tensors. For this purpose, we define the $(1,2)$-matricization of a 4th-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$ to be the $n_1 n_2 \times n_3 n_4$ matrix $\mathbf{X}_{(1,2)}$ with the entries

$$\mathbf{X}_{(1,2)}(k,\ell) = \mathcal{X}(i_1, i_2, i_3, i_4), \qquad k = j(i_1, i_2), \quad \ell = j(i_3, i_4), \tag{4.3}$$

where the function $j(\cdot)$ is defined as in (2.1).

Lemma 4.5 *Let $\mathcal{A} \in \mathbb{R}^{n \times n \times n \times n}$ be antisymmetric. Then the following statements hold:*

1. *$\mathbf{A}_{(1,2)}$ is symmetric.*
2. *Let λ be a nonzero eigenvalue of $\mathbf{A}_{(1,2)}$ with eigenvector $v \in \mathbb{R}^{n^2}$. Then the matricization $\mathbf{V}_{(1)} \in \mathbb{R}^{n \times n}$ of v is skew-symmetric.*
3. *If $\mathcal{A} = \operatorname{anti}(\alpha u_1 \otimes u_2 \otimes u_3 \otimes u_4)$ such that $\alpha \neq 0$ and $[u_1, u_2, u_3, u_4]$ is an orthonormal basis then $\mathbf{A}_{(1,2)}$ has an eigenvalue $\alpha/12$ of multiplicity 3, an eigenvalue $-\alpha/12$ of multiplicity 3, and $n^2 - 6$ zero eigenvalues. Any eigenvector v belonging to a nonzero eigenvalue satisfies* $\operatorname{range}(\mathbf{V}_{(1)}) = \operatorname{span}\{u_1, u_2, u_3, u_4\}$.

Proof. 1. This statement follows directly from the definition (4.3):

$$\mathbf{A}_{(1,2)}(k,\ell) = \mathcal{A}(i_1, i_2, i_3, i_4) = \mathcal{A}(i_3, i_4, i_1, i_2) = \mathbf{A}_{(1,2)}(\ell, k).$$

2. The relation $\mathbf{A}_{(1,2)} v = \lambda v$ implies

$$\begin{aligned}
\mathbf{V}_{(1)}(i_1, i_2) &= v(j(i_1, i_2)) = \frac{1}{\lambda} \sum_{i_3, i_4 = 1}^{n} \mathbf{A}_{(1,2)}(j(i_1, i_2), j(i_3, i_4)) v(j(i_3, i_4)) \\
&= \frac{1}{\lambda} \sum_{i_3, i_4 = 1}^{n} \mathcal{A}(i_1, i_2, i_3, i_4) v(j(i_3, i_4)) \\
&= -\frac{1}{\lambda} \sum_{i_3, i_4 = 1}^{n} \mathcal{A}(i_2, i_1, i_3, i_4) v(j(i_3, i_4)) = -\mathbf{V}_{(1)}(i_2, i_1),
\end{aligned}$$

which shows that $\mathbf{V}_{(1)}$ is skew-symmetric.

3. By the definition of $\mathcal{A}$, $\text{range}(\mathbf{V}_{(1)}) \subset \text{span}\{u_1, u_2, u_3, u_4\}$ and, together with its skew-symmetry, this implies that $\mathbf{V}_{(1)}$ is a linear combination of matrices $u_i u_j^T - u_j u_i^T$ for all $i \neq j$. Let $\pi \in S_d$ and set $\sigma = \text{sign}(\pi)$. Using Lemma 4.1, we have

$$\begin{aligned} & \mathbf{A}_{(1,2)}\big(u_{\pi(1)} \otimes u_{\pi(2)} - u_{\pi(2)} \otimes u_{\pi(1)}\big) \\ = \; & \text{vec}\big(\mathcal{A} \times_1 u_{\pi(1)} \times_2 u_{\pi(2)} - \mathcal{A} \times_1 u_{\pi(2)} \times_2 u_{\pi(1)}\big) \\ = \; & \frac{\alpha}{24}\big(\sigma u_{\pi(3)} \otimes u_{\pi(4)} - \sigma u_{\pi(4)} \otimes u_{\pi(3)} + \sigma u_{\pi(3)} \otimes u_{\pi(4)} - \sigma u_{\pi(4)} \otimes u_{\pi(3)}\big) \\ = \; & \frac{\sigma\alpha}{12}\big(u_{\pi(3)} \otimes u_{\pi(4)} - u_{\pi(4)} \otimes u_{\pi(3)}\big). \end{aligned}$$

In turn, there is an eigenspace of dimension three with orthonormal basis

$$\begin{aligned} &\big(u_1 u_2^T - u_2 u_1^T + u_3 u_4^T - u_4 u_3^T\big)/2, \\ &\big(u_1 u_4^T - u_4 u_1^T + u_2 u_3^T - u_3 u_2^T\big)/2, \\ &\big(u_3 u_1^T - u_1 u_3^T + u_2 u_4^T - u_3 u_2^T\big)/2, \end{aligned} \tag{4.4}$$

belonging to the eigenvalue $\alpha/12$. Due to the orthogonality of u_1, u_2, u_3, u_4 the range of any linear combination of (4.4) equals $\text{span}\{u_1, u_2, u_3, u_4\}$. Analogously, there is an eigenspace of dimension three belonging to the eigenvalue $-\alpha/12$ with the same property. □

Lemma 4.5.3 suggests the initialization strategy described in Algorithm 5.

Algorithm 5 HOPM initialization strategy for antisymmetric tensor of order 4

Compute eigenvector $v \in \mathbb{R}^{n^2}$ belonging to eigenvalue of largest magnitude of $\mathbf{A}_{(1,2)}$.
Form $\mathbf{V}_{(1)} \in \mathbb{R}^{n \times n}$ and compute its SVD.
Return the four leading left singular vectors u_1, u_2, u_3, u_4.

4.3 Numerical Experiments

To investigate the difference between the different initializations, we focus our experiments on antisymmetric tensors of order four. Figure 4 shows the approximation errors returned by HOPM initialized with truncated HOSVD or Algorithm 5, using the random antisymmetric tensors described in Section 3.3. HOPM is considered converged when the norm of the gradient of the objective function reaches 10^{-10} or below. It can be seen that both initialization strategy appear to work equally well in terms of the final approximation error.

Figure 5 shows the convergence behavior for a typical run. It turns out that initializing with Algorithm 5 gives a significant convergence benefit both for the approximation error and the norm of the gradient.

Finally, analogous to Section 3.3, Figure 6 shows results for the $10 \times 10 \times 10 \times 10$ tensor generated by the function

$$f(x, y, z, w) = \exp(-\sqrt{x^2 + 2y^2 + 3z^2 + 4w^2}).$$

In this case, both initialization methods yield excellent approximations.

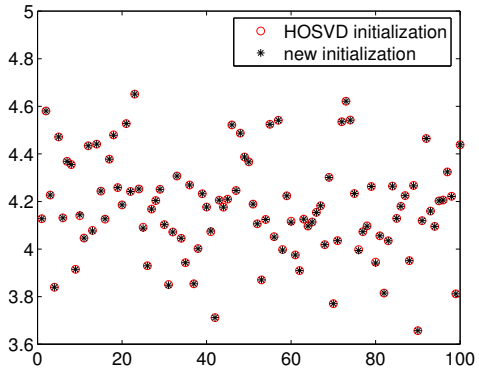


Figure 4: Approximation error of multilinear rank-4 approximation produced by HOPM for 100 random antisymmetric $10 \times 10 \times 10 \times 10$ tensors.

5 Conclusions

The multilinear rank of an antisymmetric tensor has been analyzed and new algorithms for antisymmetric low multilinear rank approximation have been proposed. The Jacobi algorithm initialized with truncated HOSVD preserves antisymmetry and appears to enjoy excellent global convergence properties. We have shown that a best unstructured rank-1 approximation can always be turned into a best antisymmetric multilinear rank-d approximation. In such a scenario, HOPM initialized either with truncated HOSVD (for $d \neq 4$) or Algorithm 5 (for $d = 4$) is certainly the method of choice. The algorithms discussed in this paper could provide a building block in the design of low-rank tensor algorithms [6] for eigenvalue problems with antisymmetric eigenvectors. In particular, the simplicity of HOPM makes it well suited in the context of truncated iterations and greedy strategies.

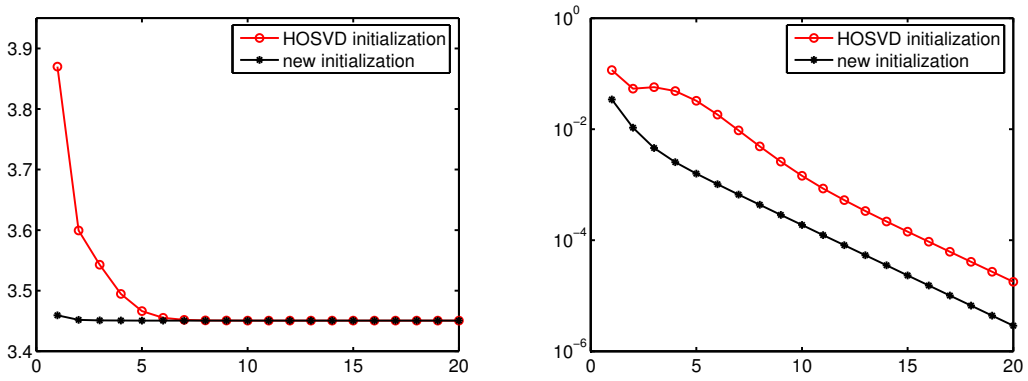
References

[1] G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.*, 26(6):2133–2159, 2005.

[2] G. Beylkin, M. J. Mohlenkamp, and F. Pérez. Approximating a wavefunction as an unconstrained sum of Slater determinants. *J. Math. Phys.*, 49(3):032107, 28, 2008.

[3] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.

[4] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank-$(r_1, \ldots, r_n)$ approximation of higher order tensors. *SIAM J. Matrix Anal. Appl.*, 21(4):1324–1342, 2000.



(a) Approximation error

(b) Norm of gradient of objective function

Figure 5: Convergence behavior of HOPM for multilinear rank-4 approximation of random antisymmetric $10 \times 10 \times 10 \times 10$ tensor.

[5] S. Friedland. Best rank one approximation of real symmetric tensors can be chosen symmetric. *Front. Math. China*, 8(1):19–40, 2013.

[6] L. Grasedyck, D. Kressner, and C. Tobler. A literature survey of low-rank tensor approximation techniques. *GAMM-Mitt.*, 36(1):53–78, 2013.

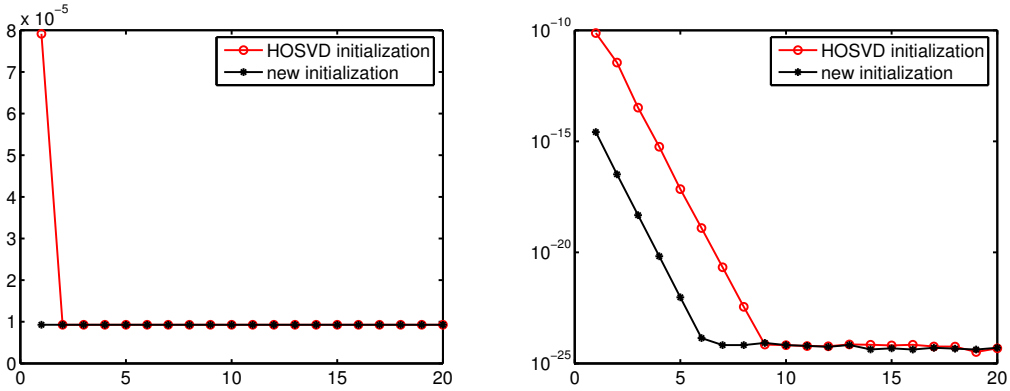
[7] M. Ishteva, P.-A. Absil, and P. Van Dooren. Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors. *SIAM J. Matrix Anal. Appl.*, 34(2):651–672, 2013.

[8] E. Kofidis and P. A. Regalia. On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM J. Matrix Anal. Appl.*, 23(3):863–884, 2002.

[9] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[10] B. N. Sheehan and Y. Saad. Higher order orthogonal iteration of tensors (HOOI) and its relation to PCA and GLRAM. *Proceedings of the 2007 SIAM International Conference on Data Mining*, pages 355–365, 2007.

[11] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza. Tensor product methods and entanglement optimization for ab initio quantum chemistry. *International Journal of Quantum Chemistry*, 115(19):1342–1391, 2015.



(a) Approximation error

(b) Norm of gradient of objective function

Figure 6: Convergence behavior of HOPM algorithm for multilinear rank-4 approximation of function-related tensor

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics

Ecole Polytechnique Fédérale (EPFL)

CH-1015 Lausanne

01.2016 ANDREA MANZONI, FEDERICO NEGRI, ALFIO QUARTERONI:
Dimensionality reduction of parameter-dependent problems through proper orthogonal decomposition

02.2016 SHENFENG ZHU, LUCA DEDÈ, ALFIO QUARTERONI:
Isogeometric analysis and proper orthogonal decomposition for the acoustic wave equation

03.2016 ROBERT LUCE, PETER HILDEBRANDT, UWE KUHLMANN, JÖRG LIESEN,:
Using separable non-negative matrix factorization techniques for the analysis of time-resolved Raman spectra

04.2015 ASSYR ABDULLE, TIMOTHÉE POUCHON:
Effective models for the multidimensional wave equation in heterogeneous media over long time and numerical homogenization

05.2016 ALFIO QUARTERONI, TONI LASSILA, SIMONE ROSSI, RICARDO RUIZ-BAIER:
Integrated heart – Coupling multiscale and multiphysics models for the simulation of the cardiac function

06.2016 M.G.C. NESTOLA, E. FAGGIANO, C. VERGARA, R.M. LANCELLOTTI, S. IPPOLITO, S. FILIPPI, A. QUARTERONI, R. SCROFANI :
Computational comparison of aortic root stresses in presence of stentless and stented aortic valve bio-prostheses

07.2016 M. LANGE, S. PALAMARA, T. LASSILA, C. VERGARA, A. QUARTERONI, A.F. FRANGI:
Improved hybrid/GPU algorithm for solving cardiac electrophysiology problems on Purkinje networks

08.2016 ALFIO QUARTERONI, ALESSANDRO VENEZIANI, CHRISTIAN VERGARA:
Geometric multiscale modeling of the cardiovascular system, between theory and practice

09.2016 ROCCO M. LANCELLOTTI, CHRISTIAN VERGARA, LORENZO VALDETTARO, SANJEEB BOSE, ALFIO QUARTERONI:
Large Eddy simulations for blood fluid-dynamics in real stenotic carotids

10.2016 PAOLO PACCIARINI, PAOLA GERVASIO, ALFIO QUARTERONI:
Spectral based discontinuous Galerkin reduced basis element method for parametrized Stokes problems

11.2016 ANDREA BARTEZZAGHI, LUCA DEDÈ, ALFIO QUARTERONI:
Isogeometric analysis of geometric partial differential equations

12.2016 ERNA BEGOVIĆ KOVAČ, DANIEL KRESSNER:
Structure-preserving low multilinear rank approximation of antisymmetric tensors