SUBTRACTING A BEST RANK-1 APPROXIMATION MAY INCREASE TENSOR RANK

Alwin Stegeman

Heymans Institute for Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands
phone: +31 50 363 6193, fax: +31 50 363 6304,
a.w.stegeman@rug.nl

Pierre Comon

Laboratoire I3S, UMR6070 CNRS, UNS,
2000, route des Lucioles, BP 121, 06903 Sophia Antipolis Cedex, France
phone: + 33 4 92 94 27 17, fax: +33 4 92 94 28 98,
pcomon@unice.fr

ABSTRACT

Is has been shown that a best rank-R approximation of an order-k tensor may not exist when $R \geq 2$ and $k \geq 3$. This poses a serious problem to data analysts using Candecomp/Parafac and related models. It has been observed numerically that, generally, this issue cannot be solved by consecutively computing and substracting best rank-1 approximations. The reason for this is that subtracting a best rank-1 approximation generally does not decrease tensor rank. In this paper, we provide a mathematical treatment of this property for real-valued $2 \times 2 \times 2$ tensors, with symmetric tensors as a special case. Regardless of the symmetry, we show that for generic $2 \times 2 \times 2$ tensors (which have rank 2 or 3), subtracting a best rank-1 approximation will result in a tensor that has rank 3 and lies on the boundary between the rank-2 and rank-3 sets. Hence, for a typical tensor of rank 2, subtracting a best rank-1 approximation has *increased* the tensor rank.

Keywords: tensor rank, low-rank approximation, tensor decomposition, multi-way, Candecomp, Parafac.
AMS subject classifications: 15A03, 15A22, 15A69, 49M27, 62H25.

1. INTRODUCTION

Tensors of order d are defined on the outer product of d linear spaces, $\mathscr{S}_\ell$, $1 \leq \ell \leq d$. Once bases of spaces $\mathscr{S}_\ell$ are fixed, they can be represented by d-way arrays. For simplicity, tensors are usually assimilated with their array representation. We assume throughout the following notation: bold italic uppercase for tensors *e.g.* $\boldsymbol{X}$, bold uppercase for matrices *e.g.* $\mathbf{T}$, bold lowercase for vectors *e.g.* $\mathbf{a}$, calligraphic for sets *e.g.* $\mathscr{S}$, and plain font for scalars *e.g.* X_{ijk}, T_{ij} or a_i, will be distinguished thanks to their font.

Let $\boldsymbol{X}$ be a 3rd order tensor defined on the tensor product $\mathscr{S}_1 \otimes \mathscr{S}_2 \otimes \mathscr{S}_3$. If a change of bases is performed in the spaces $\mathscr{S}_1, \mathscr{S}_2, \mathscr{S}_3$ by invertible matrices $\mathbf{S}, \mathbf{T}, \mathbf{U}$, then the tensor representation $\boldsymbol{X}$ is transformed into

$$\widetilde{\boldsymbol{X}} \stackrel{\text{def}}{=} (\mathbf{S}, \mathbf{T}, \mathbf{U}) \cdot \boldsymbol{X} \tag{1}$$

whose coordinates are given by $\widetilde{X}_{ijk} = \sum_{pqr} S_{ip} T_{jq} U_{kr} X_{pqr}$. This is known as the *multi-lnearity property* enjoyed by tensors. Matrices, which can be associated with linear operators, are tensors of order 2.

The rank of a tensor $\boldsymbol{X}$ is defined as the smallest number of outer product tensors whose sum equals $\boldsymbol{X}$, i.e. the smallest R such that

$$\boldsymbol{X} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r \,. \tag{2}$$

Hence a rank-1 tensor $\boldsymbol{X}$ is the outer product of vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and has entries $X_{ijk} = a_i b_j c_k$. The decomposition of a tensor into a sum of outer products of vectors and the corresponding notion of tensor rank were first introduced and studied by [14] [15].

The multilinear rank of a 3rd order tensor is a triplet (r_1, r_2, r_3), where r_i denotes the rank of the set of mode-i vectors. A mode-i vector is obtained by varying the ith index and keeping other indices fixed.

Usefulness. Tensors play a wider and wider role in numerous application areas including blind techniques for Telecommunications [21] [10] [8], Arithmetic Complexity [20] [28] [1] [27], or Data Analysis [22]. For instance, Independent Component Analysis was originally introduced for symmetric tensors whose rank did not exceed dimension [4] [2]. Now, it has become possible to estimate more factors than the dimension [13] [16] [9]. In some applications, tensors may be symmetric only in some modes [7], or may not be symmetric nor have equal dimensions [3] [22] [19]. In most of these applications, the decomposition of a tensor into a sum of rank-1 terms is relevant, since tensors entering the models to fit have a reduced rank.

Matrix algebra is insufficient. The manipulation of tensors remains difficult, because of major differences between their properties when we go from second order to higher. Several of these differences have already been underlined in the past [6], *e.g.* (i) tensor rank often exceeds dimensions, (ii) tensor rank can be different in real and complex fields, (iii) maximal tensor rank is not generic, and is still unknown in general, (iv) computing the rank of a tensor is very difficult, (v) a tensor may not have a best low-rank approximate [23] [24] [25] [12] [18] [26].

It has been observed numerically in [17, section 7] that a best or "good" rank-R approximation cannot be obtained by consecutively computing and substracting R best rank-1 approximations (which always exist). The reason for this is that subtracting a best rank-1 approximation generally does not decrease tensor rank. Hence, the deflation technique practiced for matrices (via the Singular Value Decomposition)

This work has been partially supported by contract ANR-06-BLAN-0074 "Decotes". A. Stegeman is supported by the Dutch Organisation for Scientific Research (NWO), VENI grant 451-04-102 and VIDI grant 452-08-001.

cannot generally be extended to higher-order tensors. A special case where this deflation technique works is when the tensor is diagonalizable by orthonormal multilinear transformation; see [17, section 7].

In this paper, we provide a mathematical treatment of the (in)validity of a rank-1 deflation procedure for higher-order tensors. We consider $2 \times 2 \times 2$ tensors over the real field, with symmetric tensors as a special case. First, however, we discuss the problem of finding a best rank-1 approximate to a 3rd order tensor. The proofs of our results will be available in a forthcoming full-length version of this paper.

2. BEST RANK-1 APPROXIMATION

Finding the best rank-1 approximate consists of minimizing the criterion

$$\Psi = \frac{1}{2}\,||\boldsymbol{X} - \mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}||^2 \tag{3}$$

with respect to vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. The solution will likely depend on the norm, and we shall restrict to the Frobenius norm: $||\boldsymbol{X}||^2 = \sum_{ijk}|X_{ijk}|^2$. Obviously there is a scale indeterminacy in this problem, and we could impose two of these three vectors to be unit norm. We shall not do that here because the presentation would be slightly longer. Let $\bullet_\ell$ denote the summation over the ℓth index (that is the contraction operator in the ℓth space). For instance, the product $\mathbf{A}\mathbf{B}^T$ between two matrices can be written as $\mathbf{A}\bullet_2\mathbf{B}$; if $\boldsymbol{X}$ is a 3rd order tensor, $\boldsymbol{X}\bullet_1\mathbf{a}$ is a matrix, and $\boldsymbol{X}\bullet_1\mathbf{a}\bullet_2\mathbf{b}$ is a vector. And let us rewrite criterion (3) as:

$$\Psi = \frac{1}{2}\,||\boldsymbol{X}||^2 - \boldsymbol{X}\underset{1}{\bullet}\mathbf{a}\underset{2}{\bullet}\mathbf{b}\underset{3}{\bullet}\mathbf{c} + \frac{1}{2}\,||\mathbf{a}||^2||\mathbf{b}||^2||\mathbf{c}||^2\,. \tag{4}$$

Proceeding as in [5], gradients with respect to the three vectors can be obtained:

$$\begin{aligned} d\Psi_a &= -\boldsymbol{X}\underset{2}{\bullet}\mathbf{b}\underset{3}{\bullet}\mathbf{c} + \mathbf{a}||\mathbf{b}||^2||\mathbf{c}||^2 \\ d\Psi_b &= -\boldsymbol{X}\underset{1}{\bullet}\mathbf{a}\underset{3}{\bullet}\mathbf{c} + \mathbf{b}||\mathbf{a}||^2||\mathbf{c}||^2 \\ d\Psi_c &= -\boldsymbol{X}\underset{1}{\bullet}\mathbf{a}\underset{2}{\bullet}\mathbf{b} + \mathbf{c}||\mathbf{a}||^2||\mathbf{b}||^2 \end{aligned}$$

Concerning the uniqueness of a best rank-1 approximate, one may ask the following question: are there tensors for which the solution defined by $\mathbf{a} = ||\mathbf{b}||^{-2}||\mathbf{c}||^{-2}\boldsymbol{X}\bullet_2\mathbf{b}\bullet_3\mathbf{c}$ and $d\Psi_b = d\Psi_c = 0$ is not unique up to scale? We exhibit in this section a family of such tensors.

If we plug the expression of $\mathbf{a}$ back in the equation of stationary values of $\mathbf{b}$, we get that $(\boldsymbol{X}\bullet_3\mathbf{c})\bullet_1(\boldsymbol{X}\bullet_3\mathbf{c})\bullet_2\mathbf{b} = \lambda\,\mathbf{b}$, where $\lambda = ||\mathbf{a}||^2||\mathbf{b}||^2||\mathbf{c}||^4$, which means that $\mathbf{b}$ is an eigenvector of the matrix $(\boldsymbol{X}\bullet_3\mathbf{c})\bullet_1(\boldsymbol{X}\bullet_3\mathbf{c})$. If the latter matrix is proportional to the identity for any $\mathbf{c}$, then any $\mathbf{b}$ is an eigenvector. Analogously, substituting the expression for $\mathbf{a}$ into $d\Psi_c = 0$, we get that $(\boldsymbol{X}\bullet_2\mathbf{b})\bullet_1(\boldsymbol{X}\bullet_2\mathbf{b})\bullet_3\mathbf{c} = \mu\,\mathbf{c}$, where $\mu = ||\mathbf{a}||^2||\mathbf{b}||^4||\mathbf{c}||^2$. If $(\boldsymbol{X}\bullet_2\mathbf{b})\bullet_1(\boldsymbol{X}\bullet_2\mathbf{b})$ is also proportional to the identity for any $\mathbf{b}$, then it follows that any $(\mathbf{a},\mathbf{b},\mathbf{c})$ with $\mathbf{a}$ given by $d\Psi_a = 0$ is a stationary point. Substituting the expression of $\mathbf{a}$ into the criterion (4) then yields a criterion function in $(\mathbf{b},\mathbf{c})$ for which any $(\mathbf{b},\mathbf{c})$ is a stationary point. Hence, the function is constant and any $(\mathbf{b},\mathbf{c})$ is a minimizer. This yields the following proposition

Proposition 1 *If a tensor $\boldsymbol{X}$ is such that the matrix $(\boldsymbol{X}\bullet_3\mathbf{c})$ is orthogonal for any vector $\mathbf{c}$, and $(\boldsymbol{X}\bullet_2\mathbf{b})$ is orthogonal for any vector $\mathbf{b}$, then $\boldsymbol{X}$ has infinitely many best rank-1 approximates.*

In accordance with the usual practice, we shall represent a $p\times p\times 2$ tensor $\boldsymbol{X}$ with two $p\times p$ matrix slices, $\mathbf{X}_1$ and $\mathbf{X}_2$, as $[\mathbf{X}_1\,|\,\mathbf{X}_2]$.

▷ EXAMPLE 1. Let $\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{array}\right]$. Then for any choice of non zero vector $\mathbf{c}$, the matrix $(\boldsymbol{X}\bullet_3\mathbf{c})$, obtained by linear combination of the above two matrix slices, is orthogonal. Also, for any non zero vector $\mathbf{b}$, the matrix $(\boldsymbol{X}\bullet_2\mathbf{b})$ is orthogonal. Hence $\boldsymbol{X}$ has infinitely many rank-1 approximates. ◁

Most tensors have multiple locally best rank-1 approximates, with one of them being better than the others (i.e., a unique best rank-1 approximate), as pointed out in Section 3. Examples will illustrate this fact.

Remark. The tensor in Example 1 has rank 3. Ten Berge *et al.* [29] showed that $\boldsymbol{X}$ has no best rank-2 approximation, the infimum of $||\boldsymbol{X}-\boldsymbol{Y}||^2$ over $\boldsymbol{Y}$ of rank at most 2 being 1. A more general result was obtained by De Silva and Lim [12] who showed that no $2\times2\times2$ tensor of rank 3 has a best rank-2 approximation. Stegeman [23] showed that any sequence of rank-2 approximations $\boldsymbol{Y}^{(n)}$ for which $||\boldsymbol{X}-\boldsymbol{Y}^{(n)}||^2$ converges to the infimum of 1, features diverging components. It is shown in [11] that the stationary points of the $2\times2\times2$ symmetric best rank-1 approximation problem are obtained as the roots of a 3rd degree polynomial.

3. BEST RANK-1 SUBTRACTION

From now on, we restrict our discussion to tensors in the real field. De Silva and Lim [12, Section 7] showed that $2\times2\times2$ tensors (over the real field) can be transformed by invertible multilinear matrix multiplications into eight distinct canonical forms. This partitions the space $\mathbb{R}^{2\times2\times2}$ into eight distinct orbits under the action of invertible transformations in each of the 3 modes. Table 1 lists the canonical forms for each orbit as well as their rank and multilinear rank. These quantities are invariant under the transformations defining an orbit. This kind of classification is better known for symmetric tensors or multivariate polynomials [5]. Recall the following result stated by De Silva and Lim [12]:

Lemma 2 *Let $\boldsymbol{X}$ be a $2\times2\times2$ tensor with matrix slices $\mathbf{X}_1$ and $\mathbf{X}_2$.*

(*i*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ or $\mathbf{X}_1\mathbf{X}_2^{-1}$ has real eigenvalues and is diagonalizable, then $\boldsymbol{X}$ is in orbit G_2.*

(*ii*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ or $\mathbf{X}_1\mathbf{X}_2^{-1}$ has two identical real eigenvalues with only one associated eigenvector, then $\boldsymbol{X}$ is in orbit D_3.*

(*iii*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ has complex eigenvalues, then $\boldsymbol{X}$ is in orbit G_3.*

We shall use this lemma to verify the orbit of 2-dimensional 3rd order tensors.

▷ EXAMPLE 2. Consider the tensor

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & -2 \\ 0 & 1 & 1 & 0 \end{array}\right]. \tag{5}$$

Since $\mathbf{X}_2\mathbf{X}_1^{-1}$ has complex eigenvalues, $\boldsymbol{X}$ is in orbit G_3. It can be verified that $\boldsymbol{X}$ has a unique best rank-1 approximation

$$\boldsymbol{Y} = \left[\begin{array}{cc|cc} 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{6}$$

When $\boldsymbol{Y}$ is subtracted from $\boldsymbol{X}$ we end up in orbit D_3, since

$$\boldsymbol{Z} = \boldsymbol{X} - \boldsymbol{Y} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right], \tag{7}$$

can be transformed to the canonical form of orbit D_3 by swapping rows within each slice. $\triangleleft$

	Canonical form	Tensor rank	Multilinear rank
D_0 :	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0	$(0,0,0)$
D_1 :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	1	$(1,1,1)$
D_2 :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	2	$(2,2,1)$
D_2' :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$	2	$(1,2,2)$
D_2'' :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$	2	$(2,1,2)$
G_2 :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	2	$(2,2,2)$
D_3 :	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$	3	$(2,2,2)$
G_3 :	$\left[\begin{array}{cc\|cc} -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$	3	$(2,2,2)$

Table 1: Canonical forms of $2\times2\times2$ tensors for the eight orbits under the action of invertible multilinear matrix multiplications over the real field. The letters D and G stand for "degenerate" (zero volume set in the 8-dimensional space of $2\times2\times2$ tensors) and "typical" (positive volume set), respectively.

For tensors $\boldsymbol{X}$ in the orbits of Table 1, we would like to know in which orbit $\boldsymbol{X}-\boldsymbol{Y}$ is contained, where $\boldsymbol{Y}$ is a best rank-1 approximation of $\boldsymbol{X}$. We have the following result for the degenerate orbits of ranks 1 and 2.

Proposition 3 *Let $\boldsymbol{X}$ be a $2\times2\times2$ tensor, and let $\boldsymbol{Y}$ be a best rank-1 approximation of $\boldsymbol{X}$.*

(i) If $\boldsymbol{X}$ is in orbit D_1, then $\boldsymbol{X}-\boldsymbol{Y}$ is in orbit D_0.

(ii) If $\boldsymbol{X}$ is in orbit D_2, D_2', or D_2'', then $\boldsymbol{X}-\boldsymbol{Y}$ is in orbit D_1.

For $\boldsymbol{X}$ in orbit G_2 or D_3, the tensor $\boldsymbol{X}-\boldsymbol{Y}$ is not restricted to a single orbit.

$\triangleright$ EXAMPLE 3. For the canonical tensor $\boldsymbol{X}$ of orbit G_2 in Table 1, it can be seen that $\boldsymbol{X}-\boldsymbol{Y}$ is in D_1. On the other hand, consider

$$\boldsymbol{X} = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right]. \tag{8}$$

For this tensor, $\mathbf{X}_2\mathbf{X}_1^{-1}$ has two distinct real eigenvalues. Hence, by Lemma 2, the tensor is in G_2. It can be shown that $\boldsymbol{X}$ has a unique best rank-1 approximation $\boldsymbol{Y}$ such that $\boldsymbol{X}-\boldsymbol{Y}$ equals the canonical tensor of orbit D_3 in Table 1. $\triangleleft$

$\triangleright$ EXAMPLE 4. Next, consider tensors in orbit D_3:

$$\left[\begin{array}{cc|cc} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{array}\right], \left[\begin{array}{cc|cc} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Subtracting the best rank-1 approximation $\boldsymbol{Y}$ from these tensors amounts to replacing the element 2 by zero. Hence, $\boldsymbol{X}-\boldsymbol{Y}$ is in orbit D_2, D_2', and D_2'', respectively. $\triangleleft$

On the other hand, it can be verified numerically or analytically that for $\boldsymbol{X}$ equal to the canonical tensor of orbit D_3 in Table 1, $\boldsymbol{X}-\boldsymbol{Y}$ is also in orbit D_3. Moreover, numerical experiments show that for a generic $\boldsymbol{X}$ in orbit D_3, we have $\boldsymbol{X}-\boldsymbol{Y}$ in orbit D_3 as well. This suggests the following

Conjecture 4 *If $\boldsymbol{X}$ is in D_3 and $\boldsymbol{Y}$ is the best rank-1 approximate of $\boldsymbol{X}$, then almost all tensors $\boldsymbol{X}-\boldsymbol{Y}$ are in D_3.*

Tensors given in Examples 1 or 2 were both in orbit G_3, and we have seen that $\boldsymbol{X}-\boldsymbol{Y}$ is in orbit D_3 in Example 2. For Example 1, this can be proven for any of the infinite best rank-1 approximates $\boldsymbol{Y}$ of $\boldsymbol{X}$. Numerically and analytically, we have not found any $\boldsymbol{X}$ in orbit G_3 for which $\boldsymbol{X}-\boldsymbol{Y}$ is not in orbit D_3.

We have no deterministic result for tensors in orbits G_2 and G_3, but we still have the following result, verified almost everywhere (hence the word "generic"):

Proposition 5 *Let $\boldsymbol{X}$ be a generic $2\times2\times2$ tensor, and $\boldsymbol{Y}$ be a best rank-1 approximation of $\boldsymbol{X}$. Then almost all tensors $\boldsymbol{X}-\boldsymbol{Y}$ are in orbit D_3.*

Hence, for typical tensors in orbit G_2, subtracting a best rank-1 approximate *increases* the rank to 3. For typical tensors in orbit G_3, subtracting a best rank-1 approximate does not affect the rank.

However, some non typical tensors of rank 2 may have a different behavior, as now shown.

Proposition 6 *Let $\boldsymbol{X}$ be a $2\times2\times2$ rank-2 tensor with diagonal slices, and let $\boldsymbol{Y}$ be a best rank-1 approximation of $\boldsymbol{X}$. Then $\boldsymbol{X}-\boldsymbol{Y}$ is in orbit D_1.*

$\triangleright$ EXAMPLE 5. Let $\boldsymbol{X} = \left[\begin{array}{cc|cc} a & 0 & e & 0 \\ 0 & d & 0 & h \end{array}\right]$. Then

$$\boldsymbol{X} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} a \\ e \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} d \\ h \end{pmatrix}$$

Then it can be seen that $\boldsymbol{X} = (\mathbf{I},\mathbf{I},\mathbf{T})\cdot\boldsymbol{I}$, where $\mathbf{I}$ denotes the identity matrix and $\boldsymbol{I}$ the diagonal tensor tensor with ones on its diagonal:

$$\boldsymbol{I} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \text{ and } \mathbf{T} = \begin{bmatrix} a & d \\ e & h \end{bmatrix}$$

This shows that $\boldsymbol{X}$ is is orbit G_2, and Proposition 6 implies that this is an exception to Proposition 5. $\triangleleft$

Proposition 5 states that such exceptions form a set of null measure.

4. SYMMETRIC TENSORS

A tensor is symmetric if its entries are invariant under arbitrary permutations of its indices. There is a bijection between the space of symmetric $I\times I\times I$ tensors and the space of homogeneous polynomials of degree 3 in I variables. A symmetric $I\times I\times I$ tensor $\boldsymbol{X}$ can be associated with the polynomial

$$p(s_1,\ldots,s_I) = \sum_{ijk} x_{ijk}\, s_i s_j s_k\,. \tag{9}$$

The *symmetric rank* of an order-3 symmetric tensor $\boldsymbol{X}$ is the minimal number R such that [6]:

$$\boldsymbol{X} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{a}_r \otimes \mathbf{a}_r \,. \tag{10}$$

The orbits of symmetric $2\times 2\times 2$ tensors are given in Table 2.

canonical form		polynomial	sym. rank
D_0 :	$\left[\begin{array}{cc\|cc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	0	0
D_1 :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	s_1^3	1
G_2 :	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	$s_1^3+s_2^3$	2
D_3 :	$\left[\begin{array}{cc\|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$	$3s_1^2s_2$	3
G_3 :	$\left[\begin{array}{cc\|cc} -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$	$-s_1^3+3s_1s_2^2+s_2^3$	3

Table 2: Canonical forms of symmetric $2\times 2\times 2$ tensors and associated polynomials, for the three orbits under the action of invertible multilinear transformations over the real field. The letters D and G stand for "degenerate" (zero volume set in the 4-dimensional space of symmetric $2\times 2\times 2$ tensors) and "typical" (positive volume set), respectively.

The symmetric rank of symmetric tensors of dimension 2 can be obtained from the Sylvester Theorem, at any order. This Theorem is formulated below in the case of third order tensors.

Theorem 7 (Sylvester) *A symmetric $2\times 2\times 2$ tensor with associated polynomial*

$$p(s_1,s_2) = \gamma_3 s_1^3 + 3\,\gamma_2 s_1^2 s_2 + 3\,\gamma_1 s_1 s_2^2 + \gamma_0 s_2^3\,, \tag{11}$$

has a symmetric rank-R decomposition (10) *if and only if there exists a vector $\mathbf{g} = (g_0,\ldots,g_R)^T$ such that*

$$\left[\begin{array}{ccc} \gamma_0 & \cdots & \gamma_R \\ \gamma_1 & \cdots & \gamma_{R+1} \\ \vdots & & \vdots \\ \gamma_{3-R} & \cdots & \gamma_3 \end{array}\right] \mathbf{g} = \mathbf{0}\,, \tag{12}$$

and if the polynomial $q(s_1,s_2) = g_R s_1^R + g_{R-1} s_1^{R-1} s_2 + \cdots + g_1 s_1 s_2^{R-1} + g_0 s_2^R$ has R distinct real roots.

▷ EXAMPLE 6. Using Sylvester's Theorem, one can obtain the following decomposition for the representative of orbit D_3 given in Table 2:

$$6s_1^2s_2 = (s_1+s_2)^3 + (-s_1+s_2)^3 - 2s_2^3. \tag{13}$$

In other words, the associated tensor can be written $\boldsymbol{X} = \mathbf{a}^{\otimes 3} + \mathbf{b}^{\otimes 3} - 2\,\mathbf{c}^{\otimes 3}$, where $\mathbf{a} = [1,\ 1]^T$, $\mathbf{b} = [-1,\ 1]^T$ and $\mathbf{c} = [0,\ 1]^T$. ◁

We have the following analogue of Lemma 2 to verify the orbit of symmetric tensors of dimension 2.

Lemma 8 *Let $\boldsymbol{X}$ be a symmetric $2\times 2\times 2$ tensor with matrix slices $\mathbf{X}_1$ and $\mathbf{X}_2$.*

(*i*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ or $\mathbf{X}_1\mathbf{X}_2^{-1}$ has distinct real eigenvalues, then $\boldsymbol{X}$ is in orbit G_2.*

(*ii*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ or $\mathbf{X}_1\mathbf{X}_2^{-1}$ has two identical real eigenvalues, then $\boldsymbol{X}$ is in orbit D_3.*

(*iii*) *If $\mathbf{X}_2\mathbf{X}_1^{-1}$ has complex eigenvalues, then $\boldsymbol{X}$ is in orbit G_3.*

Next, we present an example of a symmetric $2\times 2\times 2$ tensor in orbit G_3, that has a unique best symmetric rank-1 approximation $\boldsymbol{Y}$, such that $\boldsymbol{X} - \boldsymbol{Y}$ is in orbit D_3.

▷ EXAMPLE 7. Let

$$\boldsymbol{X} = \left[\begin{array}{cc\|cc} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{array}\right]. \tag{14}$$

We have

$$\mathbf{X}_2\mathbf{X}_1^{-1} = \left[\begin{array}{cc} 0 & 1 \\ -1 & 1 \end{array}\right], \tag{15}$$

which has complex eigenvalues. Hence, by Lemma 8 (*iii*) the tensor is in orbit G_3.

Next, we compute the best symmetric rank-1 approximation $\boldsymbol{Y}$ to $\boldsymbol{X}$, which has the form

$$\boldsymbol{Y} = \left[\begin{array}{cc\|cc} x_1^3 & x_1^2x_2 & x_1^2x_2 & x_1x_2^2 \\ x_1^2x_2 & x_1x_2^2 & x_1x_2^2 & x_2^3 \end{array}\right]. \tag{16}$$

After some manipulations, it can be shown that the minimum of $||\boldsymbol{X} - \boldsymbol{Y}||^2$ is obtained for $x_1^3 = x_2^3 = 3/4$, that is

$$\boldsymbol{Y} = \frac{1}{4}\left[\begin{array}{cc\|cc} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{array}\right]. \tag{17}$$

By subtraction, we obtain

$$\boldsymbol{Z} = \boldsymbol{X} - \boldsymbol{Y} = \frac{1}{4}\left[\begin{array}{cc\|cc} -3 & 1 & 1 & 1 \\ 1 & 1 & 1 & -3 \end{array}\right], \tag{18}$$

$$\text{and} \qquad \mathbf{Z}_2\mathbf{Z}_1^{-1} = \left[\begin{array}{cc} 0 & 1 \\ -1 & -2 \end{array}\right], \tag{19}$$

which has a double eigenvalue -1. Hence, by Lemma 8 (*ii*) the tensor $\boldsymbol{Z}$ is in orbit D_3. ◁

In our next example, the symmetric $2\times 2\times 2$ tensor is in orbit G_2, and has a unique best symmetric rank-1 approximation $\boldsymbol{Y}$, such that $\boldsymbol{X} - \boldsymbol{Y}$ is in orbit D_3.

▷ EXAMPLE 8. Let

$$\boldsymbol{X} = \left[\begin{array}{cc\|cc} 3 & 1 & 1 & 1 \\ 1 & 1 & 1 & 3 \end{array}\right]. \tag{20}$$

We have

$$\mathbf{X}_2\mathbf{X}_1^{-1} = \left[\begin{array}{cc} 0 & 1 \\ -1 & 4 \end{array}\right], \tag{21}$$

which has real and distinct eigenvalues. Hence, by Lemma 8 (*i*) the tensor is in orbit G_2.

Next, we compute the best symmetric rank-1 approximation $\boldsymbol{Y}$ to $\boldsymbol{X}$. It can be shown that the minimum of $||\boldsymbol{X} - \boldsymbol{Y}||^2$ is obtained for

$$\boldsymbol{Y} = \frac{3}{2}\left[\begin{array}{cc\|cc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}\right]. \tag{22}$$

By subtraction, we obtain

$$\boldsymbol{Z} = \boldsymbol{X} - \boldsymbol{Y} = \frac{1}{2}\left[\begin{array}{cc|cc} 3 & -1 & -1 & -1 \\ -1 & -1 & -1 & 3 \end{array}\right], \qquad (23)$$

$$\text{and} \qquad \mathbf{Z}_2\mathbf{Z}_1^{-1} = \left[\begin{array}{cc} 0 & 1 \\ -1 & -2 \end{array}\right], \qquad (24)$$

which has a double eigenvalue -1. Hence, by Lemma 8 (ii) the tensor $\boldsymbol{Z}$ is in orbit D_3. $\lhd$

Finally, we have the following analogue of Proposition 5.

Proposition 9 *Let $\boldsymbol{X}$ be a generic symmetric $2\times 2\times 2$ tensor, and $\boldsymbol{Y}$ be a best rank-1 approximation of $\boldsymbol{X}$. Then almost all tensors $\boldsymbol{X}-\boldsymbol{Y}$ are in orbit D_3.*

Hence, for typical symmetric $2\times 2\times 2$ tensors with symmetric rank 2, subtracting a best rank-1 approximate *increases* the symmetric rank to 3. For typical symmetric tensors with symmetric rank 3, subtracting a best rank-1 approximate does not affect the symmetric rank.

REFERENCES

[1] P. BÜRGISSER, M. CLAUSEN, and M. A. SHOKROLLAHI. *Algebraic Complexity Theory*, volume 315. Springer, 1997.

[2] J. F. CARDOSO. Blind signal separation: statistical principles. *Proc. of the IEEE*, 90:2009–2025, Oct. 1998. special issue, R.W. Liu and L. Tong eds.

[3] A. CICHOCKI and S.-I. AMARI. *Adaptive Blind Signal and Image Processing*. Wiley, New York, 2002.

[4] P. COMON. Independent Component Analysis. In J.-L. Lacoume, editor, *Higher Order Statistics*, pages 29–38. Elsevier, Amsterdam, London, 1992.

[5] P. COMON. Tensor decompositions. In J. G. McWhirter and I. K. Proudler, editors, *Mathematics in Signal Processing V*, pages 1–24. Clarendon Press, Oxford, UK, 2002.

[6] P. COMON, G. GOLUB, L.-H. LIM, and B. MOURRAIN. Symmetric tensors and symmetric tensor rank. *SIAM J. Matrix Ana. Appl.*, 30(3):1254–1279, 2008.

[7] P. COMON and M. RAJIH. Blind identification of under-determined mixtures based on the characteristic function. *Signal Proc.*, 86(9):2271–2281, Sept. 2006.

[8] A. L. F. de ALMEIDA, G. FAVIER, and J. C. M. MOTA. Parafac-based unified tensor modeling for wireless communication systems. *Signal Pro.*, 87(2):337–351, Feb. 2007.

[9] L. de Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM J. Matrix Ana. Appl.*, 28:642–666, 2006.

[10] L. de LATHAUWER and J. CASTAING. Tensor-based techniques for the blind separation of DS-CDMA signals. *Signal Proc.*, 87(2):322–336, Feb. 2007.

[11] L. de Lathauwer, B. de Moor, and J. Vandewalle. On the best rank-1 and rank-$(r_1, r_2, \ldots, r_n)$ approximation of higher-order tensors. *SIAM J. Matrix Ana. Appl.*, 21:1324–1342, 2000.

[12] V. de SILVA and L.-H. LIM. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM J. Matrix Ana. Appl.*, 30(3):1084–1127, 2008.

[13] D. L. DONOHO and X. HUO. Uncertainty principles and ideal atomic decompositions. *IEEE Trans. Inf. Theory*, 47(7):2845–2862, Nov. 2001.

[14] F. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematics and Physics*, 6:164–189, 1927.

[15] F. Hitchcock. Multiple invariants and generalized rank of a p-way matrix or tensor. *Journal of Mathematics and Physics*, 7:39–70, 1927.

[16] T. JIANG and N. SIDIROPOULOS. Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models. *IEEE Trans. Sig. Proc.*, 52(9):2625–2636, Sept. 2004.

[17] E. Kofidis and P. Regalia. On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM J. Matrix Ana. App.*, 23:863–884, 2002.

[18] W. P. KRIJNEN, T. K. DIJKSTRA, and A. STEGEMAN. On the non-existence of optimal solutions and the occurrence of 'degeneracy' in the candecomp/parafac model. *Psychometrika*, 73:431–439, 2008.

[19] P. Kroonenberg. *Applied Multiway Data Analysis*. Wiley Series in Probability and Statistics, 2008.

[20] J. B. KRUSKAL. Three-way arrays: Rank and uniqueness of trilinear decompositions. *Linear Algebra and Appl.*, 18:95–138, 1977.

[21] N. D. SIDIROPOULOS, G. B. GIANNAKIS, and R. BRO. Blind PARAFAC receivers for DS-CDMA systems. *IEEE Trans. on Sig. Proc.*, 48(3):810–823, Mar. 2000.

[22] A. SMILDE, R. BRO, and P. GELADI. *Multi-Way Analysis*. Wiley, 2004.

[23] A. STEGEMAN. Degeneracy in Candecomp/Parafac explained for $p\times p\times 2$ arrays of rank $p+1$ or higher. *Psychometrika*, 71(3):483–501, 2006.

[24] A. STEGEMAN. Degeneracy in Candecomp/Parafac explained for several three-slices arrays with a two-valued typical rank. *Psychometrika*, 72:601–619, 2007.

[25] A. STEGEMAN. Low-rank approximation of generic pxqx2 arrays and diverging components in the Candecomp/Parafac model. *SIAM J. Matrix Ana. Appl.*, 30(3):988–1007, 2008.

[26] A. STEGEMAN and L. de LATHAUWER. A method to avoid diverging components in the Candecomp/Parafac model for generic IxJx2 arrays. *SIAM J. Matrix Ana. Appl.*, 30:1614–1638, 2009.

[27] A. STEGEMAN and N. D. SIDIROPOULOS. On kruskal's condition for the candecomp/parafac decomposition. *Linear Algebra and Applications*, 420:540–552, 2007.

[28] V. STRASSEN. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52:645–685, July 1983.

[29] J. M. F. ten BERGE, H. A. L. KIERS, and J. de LEEUW. Explicit Candecomp/Parafac solutions for a contrived 2x2x2 array of rank three. *Psychome-*

trika, 53:579–584, 1988.