Tensor Low Multilinear Rank Approximation by Structured Matrix Low-Rank Approximation*

Mariya Ishteva[1] and Ivan Markovsky[1]

***Abstract*— We present a new connection between higher-order tensors and affinely structured matrices, in the context of low-rank approximation. In particular, we show that the tensor low multilinear rank approximation problem can be reformulated as a structured matrix low-rank approximation, the latter being an extensively studied and well understood problem.**

We first consider symmetric tensors. Although the symmetric tensor problem is at least as difficult as the general unstructured tensor problem, the symmetry allows us to simplify and clearly show the relation to the matrix structured low-rank approximation problem. By imposing linear equality constraints in the optimization problem, the proposed approach is applicable to unstructured tensors, as well as to affinely structured tensors. Therefore, it can be used to find (locally) optimal low multilinear rank approximation with a predefined structure.

An advantage of the proposed approach is that it can deal with more difficult variations of the main problem, including having missing and fixed elements in the given tensor or approximating with respect to a weighted norm. The drawback is its higher computational cost, compared to existing algorithms, partially due to the generality of the approach.

Key words: low-rank, higher-order tensor, multilinear rank, structured matrix, symmetric tensor.

AMS subject classifications: 15A69, 15A23, 15A83.

I. Introduction

The problem of approximating a matrix by a matrix of lower rank has been extensively studied and well-understood. Such approximations are widely used in data mining, machine learning and signal processing as a tool for dimensionality reduction, feature extraction, and classification. The optimal solution can be obtained from the truncated singular value decomposition (SVD).

This paper aims at relating two generalizations of low-rank approximation: one to affinely structured matrices and one to higher-order tensors. Structured matrix approximations [12] are used in system identification, signal processing and computer algebra, among others. The goal of structured low-rank approximation (SLRA) is to approximate a given structured matrix, e.g., symmetric, Hankel or Sylvester matrix, by a low-rank matrix with the same structure. On the other hand, data are often naturally multi-dimensional (multi-way) [15], [11], [10]. For example, term-document matrix in text mining or user-item matrix in recommender systems are naturally extended to term-document-time and user-item-time tensors under the realization that topics and user preferences change in time. An important problem then is approximating tensors by 'low-rank' tensors. Note that the concept of tensor rank is not uniquely defined. A schematic overview of the most common low-rank approximations, corresponding to different rank concepts, is presented in Fig. 1. In this paper, we establish a new connection between tensors and affinely structured matrices, in the context of multilinear rank approximation (denoted ml_rank approximation and also called Tucker-type approximation) [6], [7], and its variations. A connection between symmetric tensors and structured matrices, for the case of decomposing tensors as a sum of rank-1 terms (called canonical decomposition or parafac) [2], [5], has been exploited in [1].

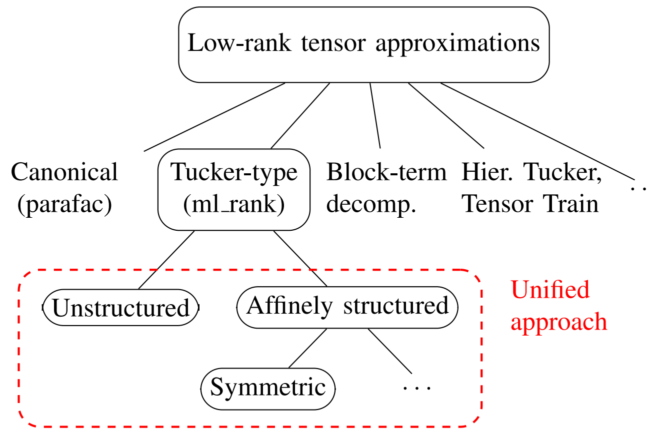

Fig. 1. Overview of tensor low-rank decompositions

The contributions of the paper are, as follows:

- We establish *a new link* between tensors and affinely structured matrices.
- This link allows us to *solve* the tensor low multilinear rank approximation problem by using techniques from structured low-rank approximation. We show how to easily perform symmetric low multilinear approximation. In addition, by imposing a simple constraint on the approximating kernel matrix, our approach is applicable to general (non-symmetric) tensors as well. We are also able to deal with *affinely structured tensors* and find (locally) best low multilinear approximation with the same structure. Moreover, all three variations of the

*The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement number 258581 "Structured low-rank approximation: Theory, algorithms, and applications", the Research Foundation Flanders (FWO-Vlaanderen), the Flemish Government (Methusalem Fund, METH1), and the Belgian Federal Government (Interuniversity Attraction Poles programme VII, Dynamical Systems, Control, and Optimization). MI is an FWO Pegasus Marie Curie Fellow.

This is a conference version of a journal paper in preparation.

[1]Dept. ELEC, Vrije Universiteit Brussel, 1050 Brussels, Belgium {mariya.ishteva, ivan.markovsky}@vub.ac.be

problem are solved *in a unified way*, whereas existing techniques for the symmetric and unstructured tensors in the literature need essentially different solution methods, due to the fact that the symmetric problem is not linear in the factor matrices.

- Finally, we note that the proposed solution approach readily solves the more difficult variations of the main problem, where the given data has *fixed or missing elements*, or where *weighted norm* has to be used as a distance measure between the given data and its approximation.

The drawback of the current approach is its higher computational cost, compared to existing algorithms. This is partially due to the fact that the proposed scheme is more general and is able to solve several problems in the same way.

Outline

The remainder of this paper is organized as follows. In Section II, the low multilinear rank approximation problem is defined and discussed. Next, the structured low-rank approximation problem is briefly presented in Section III. The main link between these two problems is presented in Section IV, in the case of symmetric tensors. The extension to unstructured and affinely structured tensors is briefly discussed in Section V. Preliminary numerical experiments are presented in Section VI. Finally, in Section VII we draw our final conclusions.

Notation

We denote tensors and structure specifications by calligraphic letters ($\mathcal{A}, \widehat{\mathcal{A}}, \ldots$ and $\mathcal{S}$, respectively), matrices by bold-face capitals ($\mathbf{P}, \mathbf{L}, \ldots$), and vectors by lower-case letters ($p, \widehat{p}, \ldots$). The identity matrix is denoted by $\mathbf{I}$ and the zero matrix is denoted by $\mathbf{0}$.

In this paper, we illustrate our ideas in terms of third-order tensors. However, the results are trivially generalizable to tensors of higher orders as well.

II. Low multilinear rank approximation

A. Problem formulation

Third-order tensors are generalizations of matrices to 3-way arrays and find applications in chemometrics, biomedical signal processing, and telecommunications, among others. As a natural generalization of the matrix column- and row-rank, the multilinear rank [6], [7], denoted here as $\mathrm{ml\_rank}(\mathcal{A})$, of a third-order tensor $\mathcal{A}$, is the triplet of the numbers of linearly independent mode-1 vectors (columns), mode-2 vectors (rows), and mode-3 vectors, respectively. It is also useful to define matrix representations of the tensors, stacking each set of vectors in a pre-specified order in a matrix $\mathbf{A}_{(i)}, i = 1, 2, 3$. Note that the entries of the multilinear rank can also be defined as the ranks of these matrices, i.e.,

$$\mathrm{ml\_rank}(\mathcal{A}) = (\mathrm{rank}(\mathbf{A}_1), \mathrm{rank}(\mathbf{A}_2), \mathrm{rank}(\mathbf{A}_3)). \quad (1)$$

Given a third-order tensor $\mathcal{A}$ and a multilinear rank specification (r_1, r_2, r_3), the low multilinear rank approximation problem can be formulated as

$$\min_{\widehat{\mathcal{A}}} \|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2, \quad \text{s. t.} \quad \mathrm{ml\_rank}(\widehat{\mathcal{A}}) \leq (r_1, r_2, r_3), \quad (2)$$

where $\|\cdot\|_F$ stands for the Frobenius norm. If the given tensor has some affine structure, and in particular, if the original tensor is symmetric, it is desirable that this structure is preserved in the approximation as well. This results in adding an additional constraint in (2), i.e.,

$$\min_{\widehat{\mathcal{A}}} \|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2, \quad \text{s. t.} \quad \begin{cases} \mathrm{ml\_rank}(\widehat{\mathcal{A}}) \leq (r_1, r_2, r_3), \\ \widehat{\mathcal{A}} \text{ is structured.} \end{cases} \quad (3)$$

B. Multilinear singular value decomposition

The best (unstructured or symmetric) matrix low-rank approximation can be obtained from the truncated SVD [4, §2.5, §8.6]. The multilinear singular value decomposition (MLSVD) [3], [17], [18] is a generalization of the SVD to higher-order tensors. Every tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ can be decomposed as a product of a tensor $\mathcal{B} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, called core tensor, and three orthogonal matrices $\mathbf{U}^{(i)} \in \mathbb{R}^{n_i \times n_i}$, $i = 1, 2, 3$, i.e.,

$$\mathcal{A} = \mathcal{B} \bullet_1 \mathbf{U}^{(1)} \bullet_2 \mathbf{U}^{(2)} \bullet_3 \mathbf{U}^{(3)},$$

where $\bullet_i$ stands for the tensor-matrix multiplication with respect to mode-i of the tensor. The factor matrices are computed such that the matrix slices of $\mathcal{B}$ in any direction are orthogonal to each other (i.e, their inner product is 0) and have decreasing norm (when increasing the indices). These properties reduce to having a diagonal core matrix if the original tensor is a matrix, i.e., a second-order tensor.

Computing the MLSVD requires computing three SVDs. The columns of the singular matrices $\mathbf{U}^{(i)}, i = 1, 2, 3$ are obtained as the left singular vectors of $\mathbf{A}_{(i)}, i = 1, 2, 3$. The core tensor can then be computed as

$$\mathcal{B} = \mathcal{A} \bullet_1 \mathbf{U}^{(1)^\top} \bullet_2 \mathbf{U}^{(2)^\top} \bullet_3 \mathbf{U}^{(3)^\top}. \quad (4)$$

A striking difference with the matrix case is, however, that the truncated MLSVD results, in general, in a suboptimal solution of (2). This is due to the fact that, in general, tensors of order higher than two, cannot be decomposed with a diagonal core tensor $\mathcal{B}$. Thus, (2) and, in the symmetric case, (3), are solved by iterative algorithms, usually starting from the truncated MLSVD approximation.

III. Structured matrix low-rank approximation

A. Problem formulation

Affinely structured matrices, e.g., Hankel, Toeplitz, or Sylvester matrices, appear naturally in system identification, signal processing, and computer algebra [12]. These matrices have a pattern for the position of their elements and can be defined by a vector of distinct elements $p \in \mathbb{R}^{n_p}$ and a structure specification $\mathcal{S}$,

$$\begin{aligned} \mathcal{S}: \quad \mathbb{R}^{n_p} &\rightarrow \mathbb{R}^{m \times n}, \\ p &\mapsto \mathcal{S}(p), \end{aligned}$$

where m and n are the dimensions of the matrix.

The structured low-rank approximation problem is defined as follows: Given a structure specification $\mathcal{S}$, a structure parameter vector p, and a rank specification r,

$$\min_{\widehat{p}} \|\mathcal{S}(p) - \mathcal{S}(\widehat{p})\|_F^2, \quad \text{s. t.} \quad \operatorname{rank}(\mathcal{S}(\widehat{p})) \le r. \tag{5}$$

To illustrate the meaning of the structure specification $\mathcal{S}$ in the context of third-order tensors, consider the following example. For a given symmetric tensor (tensor invariant under permutation of the indices)[1] $\mathcal{A} \in \mathbb{R}^{3\times3\times3}$, its matrix representation $\mathbf{A}_{(1)}$ is a linearly (and thus also affinely) structured matrix with $p \in \mathbb{R}^{10}$ and

$$\begin{aligned} \mathbf{A}_{(1)} &= \mathcal{S}(p) \\ &= \left[\begin{array}{ccc|ccc|ccc} p_1 & p_2 & p_3 & p_2 & p_4 & p_5 & p_3 & p_5 & p_6 \\ p_2 & p_4 & p_5 & p_4 & p_7 & p_8 & p_5 & p_8 & p_9 \\ p_3 & p_5 & p_6 & p_5 & p_8 & p_9 & p_6 & p_9 & p_{10} \end{array}\right]. \end{aligned} \tag{6}$$

B. Solution approaches

Existing algorithms solve (5) i) by local optimization, ii) by using relaxations, or iii) using heuristics. In this paper, we consider two recent local optimization approaches, namely the kernel approach `slra` [13] and the image approach `penalized slra` [9]. The difference between these two approaches is in the way the rank constraint in (5) is treated. The kernel approach, is based on the fact that the rank of an $m \times n$ matrix and the dimention of its left kernel sum to m. Thus, the rank constraint is reformulated as

$$\operatorname{rank}(\mathcal{S}(\widehat{p})) \le r \iff \mathbf{R}\mathcal{S}(\widehat{p}) = 0, \tag{7}$$

for some full row rank matrix $\mathbf{R} \in \mathbb{R}^{(m-r)\times m}$. Second, the image approach is based on the fact that a matrix has low rank if and only if it can be represented as a product of two matrices with reduced dimension, i.e.,

$$\operatorname{rank}(\mathcal{S}(\widehat{p})) \le r \iff \mathcal{S}(\widehat{p}) = \mathbf{PL}, \tag{8}$$

for some $\mathbf{P} \in \mathbb{R}^{m\times r}$, $\mathbf{L} \in \mathbb{R}^{r\times n}$.

IV. Connecting tensors to affinely structured matrices

For simplicity of the presentation, in this section we consider symmetric tensors and thus problem (3) becomes

$$\min_{\text{sym. } \widehat{\mathcal{A}}} \|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2, \quad \text{s. t.} \quad \text{ml\_rank}(\widehat{\mathcal{A}}) \le (r, r, r). \tag{9}$$

In order to reformulate this problem as a structured low-rank matrix approximation problem, recall first that the Frobenius norm is defined element-wise, and thus

$$\|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2 = \|\mathbf{A}_{(1)} - \widehat{\mathbf{A}}_{(1)}\|_F^2.$$

Second, symmetric tensors are tensors invariant under permutation of the indices, i.e.,

$$\begin{aligned} &\mathcal{A}(i,j,k) = \mathcal{A}(i,k,j) = \mathcal{A}(j,i,k) \\ &= \mathcal{A}(j,k,i) = \mathcal{A}(k,i,j) = \mathcal{A}(k,j,i). \end{aligned}$$

[1] Symmetric tensors are sometimes called supersymmetric.

As a consequence, the matrix representations $\mathbf{A}_{(i)}, i = 1, 2, 3$, of a symmetric tensor $\mathcal{A}$ are equal to each other[2] and thus have equal ranks. We then have

$$\text{ml\_rank}(\mathcal{A}) \le (r, r, r) \quad \iff \quad \operatorname{rank}(\mathbf{A}_{(1)}) \le r.$$

Finally, by considering a structure specification $\mathcal{S}$ corresponding to a matrix representation of a symmetric tensor (as in the example in Section III), i.e., $\widehat{\mathbf{A}}_{(1)} = \mathcal{S}(\widehat{p})$, we can reformulate problem (9) as the structured low-rank approximation problem

$$\boxed{\min_{\widehat{p}} \|\mathbf{A}_{(1)} - \mathcal{S}(\widehat{p})\|_F^2, \quad \text{s. t.} \quad \operatorname{rank}(\mathcal{S}(\widehat{p})) \le r.} \tag{10}$$

Problem (10) is readily solved by existing algorithms [12], [9].

Thus, we established a new connection between tensors and affinely structured matrices, in the context of multilinear rank approximation (2). On one hand, compared to other tensor algorithms in the literature, e.g., [8], the proposed approach has higher computational cost. On the other hand, the matrix structured low-rank approximation problem readily offers additional features, underdeveloped in the tensor world. For example, the proposed approach allows:

- Using the parameter norm $\|p - \widehat{p}\|_2^2$, instead of the Frobenius norm in (10), where p contains the distinct elements of $\mathcal{A}$. This way each *distinct* element of the tensor has equal weight;
- Using the more general weighted norm $(p-\widehat{p})^\top W (p-\widehat{p})$, e.g., if prior knowledge about the importance or the correctness of each (noisy) structure parameter is available;
- Working in the presence of missing or fixed elements [12], [9] in (2).

V. Extension to unstructured and affinely structured tensors

With some modifications, the proposed approach is applicable to unstructured and to affinely structured tensors. The matrix representations of the tensor are not equal to each other anymore, but still contain the same elements. Thus, the structure parameter vector is still the vector of distinct elements, but the structure specification $\mathcal{S}$ changes, namely $\mathcal{S}(p)$ becomes a larger matrix containing all matrix representations of the tensor.

A. Unstructured tensors

The most studied low multilinear rank approximation problem is the one for general unstructured tensors. It is sometimes also utilized in the symmetric case, since practical experience shows that approximating a symmetric tensor by using general purpose algorithms leads to symmetric approximations (although no theoretical proof is known yet).

In the unstructured case, the structure specification $\mathcal{S} : \mathbb{R}^{n^3} \to \mathbb{R}^{3n\times n^2}$ combines the information for all three

[2] This is a generalization of the fact that for any symmetric matrix $\mathbf{M}$, $\mathbf{M}^\top = \mathbf{M}$.

structure specifications, corresponding to each of the matrix representations of the tensor. This is necessary in order to have a single rank constraint, rather than three different ones. Existing algorithms for structured matrix low-rank approximation can then be applied with slight modifications, namely by considering additional fixed zeros in the optimization matrix variables. In the kernel representation, the kernel matrix $\mathbf{R}$ would be a block-diagonal matrix with blocks corresponding to each of the matrix representations of the tensor. In the image representation, the fixed zeros are inherent to one of the factors. For example, it is enough that the $\mathbf{P}$ factor is block-diagonal, with blocks corresponding to each of the matrix representations of the tensor.

These become more difficult problems than the one in the symmetric case (10), but can still be solved by `slra` [13] in the kernel setting and by `penalized slra` [9] in the image representation. Note that the fixed zeros can be considered as linear equality constraints.

B. *Affinely structured tensors*

In [14], in the context of exponential data fitting, low multilinear rank approximations of tensors with Hankel structure are needed. Currently, unstructured approximations are used, but Hankel-structured approximations could potentially lead to better results.

Solving the structured low multilinear rank approximation problem is similar in spirit to solving the unstructured problem. The difference is in the definitions of $\mathcal{S}_i$, namely now

$$\mathcal{S}_i : \mathbb{R}^{n_p} \to \mathbb{R}^{n \times n^2},$$

with $n_p \leq n^3$. If we have a large number of repeated elements in the structure, n_p will be small compared to n^3 in the unstructured case. This would reduce the computational cost but probably not essentially. On the other hand, in this case it is possible to have more local optima due to the increased number of constraints (elements in the approximation being equal to each other).

VI. Numerical illustrations

In this section, we compare the results obtained with the proposed approach, implemented by the methods of [13] and [9], to the results from Tensorlab [16], which is a recent MATLAB toolbox for tensor computations. In the case of symmetric tensors, in addition, we also compare with the state-of-the-art Jacobi algorithm [8].

A. *Symmetric tensor approximation*

Consider the example from [8, §4.2],

$$\mathcal{A}(:,:,1) = \begin{bmatrix} 1.2753 & -0.5811 & -0.0725 \\ -0.5811 & -0.8475 & 0.0379 \\ -0.0725 & 0.0379 & -1.0573 \end{bmatrix},$$

$$\mathcal{A}(:,:,2) = \begin{bmatrix} -0.5811 & -0.8475 & 0.0379 \\ -0.8475 & -1.0771 & -0.6544 \\ 0.0379 & -0.6544 & -0.7375 \end{bmatrix},$$

$$\mathcal{A}(:,:,3) = \begin{bmatrix} -0.0725 & 0.0379 & -1.0573 \\ 0.0379 & -0.6544 & -0.7375 \\ -1.0573 & -0.7375 & 0.1491 \end{bmatrix},$$

where we used MATLAB's notation $\mathcal{A}(:,:,i), i = 1,2,3$, for denoting the i-th frontal slice of $\mathcal{A}$. We have $\mathbf{A}_{(1)} = \mathcal{S}(p)$, with $\mathcal{S}$ as in (6) and

$$p = [\,1.2753 \quad -0.5811 \quad -0.0725 \quad -0.8475 \quad 0.0379 \ldots \\ -1.0573 \quad -1.0771 \quad -0.6544 \quad -0.7375 \quad 0.1491\,].$$

Let $r = 2$, as in the original example, i.e., we are aiming at a rank-$(2,2,2)$ approximation.

We ran Tensorlab's [16] `lmlra` function[3], the `Jacobi` algorithm from [8], and the two variants of the method proposed in this paper ((10) with the kernel (7) and with the image (8) representation). We used the default initializations, which are based on the truncated SVD of $\mathbf{A}_{(1)}$. The results on the relative error $\|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2 / \|\mathcal{A}\|_F^2$ of the approximation are given in Table I (left). As it can be seen, all algorithms

TABLE I

COMPARISON OF ALGORITHMS FOR SOLVING (9) ON THE SYMMETRIC TENSOR APPROXIMATION EXAMPLE FROM [8]. LEFT: RELATIVE APPROXIMATION ERROR $\|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2 / \|\mathcal{A}\|_F^2$ OF `lmlra` [16], `Jacobi` [8], AND THE KERNEL AND IMAGE APPROACHES TO (10). RIGHT: RELATIVE APPROXIMATION ERROR $\|p - \widehat{p}\|_2^2 / \|p\|_2^2$ OF THE KERNEL AND IMAGE APPROACHES FOR THE MODIFIED PROBLEM (9) WITH PARAMETER NORM.

$\|\mathcal{A} - \widehat{\mathcal{A}}\|_F^2 / \|\mathcal{A}\|_F^2$				$\|p - \widehat{p}\|_2^2 / \|p\|_2^2$	
`lmlra`	`Jacobi`	(10), (7)	(10), (8)	(10), (7)	(10), (8)
0.2827	0.2827	0.2827	0.2827	0.1927	0.1927

give the same result[4]. In addition, if we are interested in the parameter norm $\|p - \widehat{p}\|_2^2$ (instead of in the Frobenius norm) in (9), we can apply a weighted version of the proposed method. The relative approximation error $\|p - \widehat{p}\|_2^2 / \|p\|_2^2$ is reported in Table I (right) for both the kernel and the image representation approaches. The other two algorithms (`lmlra` and `Jacobi`), in their current form, cannot be applied in this case.

[3]Although `lmlra` is meant for unstructured tensors only, it can still be applied on symmetric tensors. Practical experience shows that the approximation obtained from `lmlra` is symmetric, although no proof is known.

[4]The numbers differ from the fifth decimal digit on due to numerical errors.

B. Symmetric tensor completion

Let $\mathcal{T} \in \mathbb{R}^{3\times3\times3}$ be a symmetric rank-$(2,2,2)$ tensor. Such tensors can be constructed as a product of a symmetric core tensor $\mathcal{C} \in \mathbb{R}^{2\times2\times2}$ and a matrix $\mathbf{M} \in \mathbb{R}^{3\times2}$, in the following way

$$\mathcal{T} = \mathcal{C} \bullet_1 \mathbf{M} \bullet_2 \mathbf{M} \bullet_3 \mathbf{M}.$$

Let the distinct elements of $\mathcal{T}$ be collected in (or come from) $p_0 \in \mathbb{R}^{10}$ and suppose that the elements of p_0 are affected by additive noise,

$$p = p_0 + 0.1e.$$

In our experiment, the elements of $\mathcal{C}$, $\mathbf{M}$, and e were drawn from the normal distribution with zero mean and unit variance. After the noise has been added, some of the elements of p were removed and only then the vector was given to our algorithm. We removed the third and the sixth elements and ran (10) with (7) (kernel approach) and (10) with (8) (image approach) to obtain a rank-$(2,2,2)$ full tensor from the partially observed noisy tensor. Since the noise has been added to the distinct elements of the tensor, it is more natural to optimize with respect to the parameter norm $\|p - \widehat{p}\|_2^2$. The results from one representative example are reported in Table II. As it can be seen, the kernel and image based algorithms can 'guess' the missing elements with high precision.

TABLE II

RESULTS OF A SYMMETRIC TENSOR LOW-MULITLINEAR RANK COMPLETION EXPERIMENT. THE TRUE STRUCTURE PARAMETER VECTOR IS AFFECTED BY ADDITIVE NOISE AND, ADDITIONALLY, 2 ELEMENTS HAVE BEEN REMOVED. WE PRESENT THE TRUE, THE NOSIY AND THE GIVEN VECTOR, AS WELL AS THE RESULTS FORM THE PROPOSED KERNEL AND IMAGE APPROACH.

True	Noisy	Given	Kernel solution	Image solution
6.2440	6.1010	6.1010	6.0656	6.0618
-7.1795	-7.0727	-7.0727	-7.1897	-7.1902
-1.7843	-1.8658	**NaN**	-1.8810	-1.8702
7.3515	7.4720	7.4720	7.3414	7.3433
0.9617	1.0307	1.0307	0.9483	0.9416
-0.8049	-0.8676	**NaN**	-0.8072	-0.7982
-7.0476	-6.9657	-6.9657	-7.0150	-7.0174
-0.4057	-0.3564	-0.3564	-0.4460	-0.4421
0.5695	0.3586	0.3586	0.4417	0.4362
-0.1993	-0.3468	-0.3468	-0.3087	-0.3036

The numerical experiments presented in this paper aim at confirming the correctness of the proposed method. Comparisons with other algorithms, such as tensor completion algorithms, and examples with unstructured and affinely structured tensors are to be reported in a follow-up paper.

VII. CONCLUSIONS

We have established *a new link* between tensors and affinely structured matrices, allowing us to *solve* the tensor low multilinear rank approximation problem by using techniques from structured low-rank approximation. Three variations of the problem can be solved in this *unified way*, namely the case of symmetric, general unstructured and *affinely structured* tensors, where we find (locally) best low multilinear approximation with the same structure. Moreover, the proposed solution approach readily solves the more difficult variations of the main problem, where the given data has *fixed or missing elements*, or where *weighted norm* has to be used as a distance measure between the given data and its approximation.

The drawback of the current approach is its higher computational cost, compared to existing algorithms. This is partially due to the fact that the proposed scheme is more general and is able to solve several problems in the same way. A topic of further investigation is how the computational cost of proposed approach can be reduced by using the sparse structure of the involved matrices.

REFERENCES

[1] J. Brachat, P. Comon, B. Mourrain, and E. Tsigaridas. Symmetric tensor decomposition. *Linear Algebra and its Applications*, 433(11–12):1851–1872, 2010.

[2] J. Carroll and J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35(3):283–319, 1970.

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

[4] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[5] R. A. Harshman. Foundations of the PARAFAC procedure: Model and conditions for an "explanatory" multi-mode factor analysis. *UCLA Working Papers in Phonetics*, 16(1):1–84, 1970.

[6] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematical Physics*, 6(1):164–189, 1927.

[7] F. L. Hitchcock. Multiple invariants and generalized rank of a p-way matrix or tensor. *Journal of Mathematical Physics*, 7(1):39–79, 1927.

[8] 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.*, 2(34):651–672, 2013.

[9] M. Ishteva, K. Usevich, and I. Markovsky. Regularized structured low-rank approximation with applications. Technical report, Vrije Universiteit Brussel, 2013. Available from `http://arxiv.org/abs/1308.1827`.

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

[11] P. M. Kroonenberg. *Applied Multiway Data Analysis*. Wiley, 2008.

[12] I. Markovsky. *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer, 2012.

[13] I. Markovsky and K. Usevich. Structured low-rank approximation with missing data. *SIAM J. Matrix Anal. Appl.*, 34:814–830, 2013.

[14] J.-M. Papy, L. De Lathauwer, and S. Van Huffel. Exponential data fitting using multilinear algebra: The decimative case. *J. Chemometrics*, 23(7–8):341–351, July/August 2009. Special Issue in Honor of Professor Richard A. Harshman.

[15] A. Smilde, R. Bro, and P. Geladi. *Multi-way Analysis. Applications in the Chemical Sciences*. John Wiley and Sons, Chichester, U.K., 2004.

[16] L. Sorber, M. Van Barel, and L. De Lathauwer. Tensorlab v1.0. Technical report, 2013. Available online, February 2013, `http://esat.kuleuven.be/sista/tensorlab/`.

[17] L. R. Tucker. The extension of factor analysis to three-dimensional matrices. In H. Gulliksen and N. Frederiksen, editors, *Contributions to mathematical psychology*, pages 109–127. Holt, Rinehart & Winston, NY, 1964.

[18] L. R. Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31:279–311, 1966.