Neurocomputing 73 (2010) 3323–3327

Contents lists available at ScienceDirect

Neurocomputing

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

Letters

Numerical optimization of a sum-of-rank-1 decomposition for n-dimensional order-p symmetric tensors

Olexiy Kyrgyzov*, Deniz Erdogmus

Northeastern University, Department of Electrical and Computer Engineering, 380 Huntington Avenue, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 15 October 2009
Received in revised form
14 June 2010
Accepted 28 June 2010
Communicated by S. Choi
Available online 27 July 2010

Keywords:
Symmetric tensor decomposition
Iterative optimization
Basis vector frame

ABSTRACT

In this paper, we present a sum-of-rank-1 type decomposition and its differential model for symmetric tensors and investigate the convergence properties of numerical gradient-based iterative optimization algorithms to obtain this decomposition. The decomposition we propose reinterprets the orthogonality property of the eigenvectors of symmetric matrices as a geometric constraint on the rank-1 matrix bases, which leads to a geometrically constrained eigenvector frame. Relaxing the orthogonality requirement, we developed a set of structured-bases that can be utilized to decompose any symmetric tensor into a similar constrained sum-of-rank-1 decomposition.

1. Introduction

This paper provides a sum-of-rank-1 decomposition method for symmetric tensors and evaluates different iterative optimization techniques to determine this decomposition for tensors of different dimensionalities and orders. A preliminary version of this paper where the decomposition model was proposed and solved for using iterative random search appeared earlier [1]; in this we described a generalized geometric constraint for the basis rank-one tensors in the form of an inner product matrix, which also corresponded to a geometric constraint on the vector frame that would generate these basis tensors for any n-dimensional order-p symmetric tensor. We refrain from calling these basis vectors (and corresponding rank-1 tensors) '*eigen*' because at this time we have not proven any invariance property that would warrant this nomenclature. The proposition is a generalization of the orthogonal coordinate frame interpretation of the eigenvectors of symmetric real-valued matrices, which remain as order-2 special cases.

In addition to the proposed decomposition model, in this paper, we present and compare iterative descent techniques for the optimization of the vector frame rotation parameters and corresponding linear combination coefficients (analogous to eigenvalues). These techniques are Jacobi iteration, gradient descent, Gauss–Newton, and Levenberg–Marquardt methods for Givens rotation angles. This paper contributes to the field of tensor analysis and decomposition that appears increasingly in many fields of applications such as signal [2] and image [3] processing, factor analysis [4,5,6], speech, telecommunications [8], and neuroscience [9]. As tensors could emerge from higher-order statistics such as joint moments and cumulants (e.g. consider the symmetric tensor formed by order-p joint moments of an n-dimensional random vector), we believe, tensor decompositions will play an increasingly more important role in statistical signal processing. A useful and intuitive tensor decomposition definition accompanied by computationally efficient algorithms will be the key to widespread utilization of these multilinear objects in extending subspace techniques.

2. Decomposition of a symmetric tensor

The most widely recognized approaches for sum-of-rank-1 tensor decompositions are as follows: (1) canonical decomposition (CANDECOMP) [12] or alternatively parallel factor analysis (PARAFAC) [4], and (2) the Tucker [5] model. With the CP model, a tensor can be represented as a sum of rank-1 tensors with minimum number of bases in a unique fashion and, by definition, the dimensionality of this basis set is the rank. The CP model does not constrain the geometry of the vectors that yield the rank-1 basis tensors; this is in contrast with the assumption of orthogonal vectors in singular value decomposition (SVD) of matrices. While a matrix might be written as a sum of fewer unconstrained left–right vector products than prescribed by SVD, the orthogonality constraint on the geometry of the vectors has been found to be useful in many applications of SVD. As opposed to the CP, Tucker's proposed decomposition factors tensors as a finite sum assuming orthogonal

* Corresponding author.
E-mail address: kyrgyzov@ece.neu.edu (O. Kyrgyzov).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.06.013

vectors to generate the rank-1 basis tensors similar to SVD, but the result is not necessarily minimal (in fact most of the time the Tucker decomposition will require many more rank-1 bases than the CP). In this paper, we propose an alternative decomposition model that utilizes a set of geometrically constrained vectors to generate the rank-1 bases; the constraint relaxes the orthogonality assumption in the Tucker model, which results in decompositions which turn out to be a one-to-one reparameterization of the symmetric tensor, thus using fewer rank-1 components; however, we cannot also claim that the identified decomposition is minimum rank in the CP-sense.

The decomposition of a real-valued symmetric tensor $\mathbf{A}$ into a sum of rank-1 tensors utilizes basis tensors that are p-way outer products of the same vector (referred to as rank-1 symmetric tensors) [2,8]:

$$\mathbf{A} = \sum_{l=1}^{r} \lambda_l \mathbf{u}_l^{\circ p} \tag{1}$$

where $\mathbf{u}^{\circ p}$ denotes the p-way outer-product of the vector $\mathbf{u}$. Tensor decomposition problem is fundamental to the extension of subspace analysis techniques in signal processing that arise from the study of second order statistics of vector-valued measurements to higher order statistics. Existing examples of such applications include blind source separation. For instance, an exponential multivariate family as a signal model can be factorized using a sum-of-rank-1 tensor decomposition; consider an n-variate order-p polynomial

$$q(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}^{\circ p} = \sum_{l_1=0}^{n} \cdots \sum_{l_p=0}^{n} \mathbf{A}_{l_1,\ldots,l_p} \cdot x_{l_1} \cdot \ldots \cdot x_{l_p}$$

where $x_0 = 1$. If the (symmetric) tensor $\mathbf{A}$ containing these polynomial coefficients is decomposed into the desired form, then the polynomial can be written as $\mathbf{A} \cdot \mathbf{x}^{\circ p} = \sum_{l=1}^{r} \lambda_l (\mathbf{u}_l^T \mathbf{x})^{\circ p}$, and an exponential density $e^{q}(x)$ can be factorized into a product of univariate exponentials. Other applications are reviewed by Kolda and Bader [8] and include finding polynomial factorizations [10,11].

2.1. Order-2 n-dimensional symmetric tensors

A symmetric n-dimensional order-2 tensor is a symmetric matrix. Eigenvector bases for real symmetric matrices are orthogonal, and can always be made into an orthonormal basis. Thus, a real n-dimensional symmetric matrix can be decomposed as

$$\mathbf{A} = \sum_{l=1}^{n} \lambda_l \mathbf{u}_l^{\circ 2} \tag{2}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$ is the matrix where columns form an orthogonal frame in n-dimensional space.

For numerical determination of (2) we can use, for instance, the Jacobi algorithm [7] that tries to find $q = \binom{n}{2} = n(n-1)/2$ rotation angles $\{\theta_k,\ k = 1, \ldots, q\}$, such that we can construct a rotation matrix $\mathbf{R}(\theta_k)$ in plane (i,j) $\{i = 1, \ldots, n-1,\ j = i+1, \ldots, n\}$ with angle θ_k (with a one-to-one correspondence between the indices k and (i,j) in this Givens angle parameterization). This eigendecomposition solution consists of q rotation angles and n eigenvalues. The number of free elements of a symmetric n-dimensional matrix $\mathbf{A}$, $m(n,2) = n(n+1)/2$, equals the sum $n+q$. Consequently, eigendecomposition is simply a reparameterization procedure. Solving for the rotation matrices $\mathbf{R}(\theta_k)$, we can get the orthonormal eigenvectors given by

$$\mathbf{U} = \prod_{k=1}^{q} \mathbf{R}(\theta_k), \quad \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I} \tag{3}$$

Due to the orthonormality of $\mathbf{U}$, the eigenvalues are uniquely identified by the Frobenius inner-product vector between the target matrix and the basis matrices ($r = n$) [8]:

$$\langle \mathbf{u}_l^{\circ 2}, \mathbf{A} \rangle_F = \sum_{i=1}^{r} \lambda_i \langle \mathbf{u}_l^{\circ 2}, \mathbf{u}_i^{\circ 2} \rangle_F = \sum_{i=1}^{r} \lambda_i (\mathbf{u}_l^T \mathbf{u}_i)^2 = \lambda_l \tag{4}$$

2.2. Order-p 2-dimensional symmetric tensors

Let $\mathbf{A}$ be a 2-dimensional order-p real symmetric tensor. The number of free elements (i.e., dimensionality) in such tensors is $m(2,p) = (p+1)$. Rotation of the whole family of 2-dimensional unit vectors is defined just by a single angle θ. For a 1–1 reparameterization, the number of linear combination coefficients, r, plus the number of parameters that characterize rotation of the family of r corresponding vectors, s, should be equal to the number of free elements in the tensor, i.e., $(r+s) = (p+1)$. Alternative approach without restrictions on the vector frame is described in [13].

Incorporating these conditions into the design of the rank-1 sum decomposition on the right-hand side of (1), we obtain that real-symmetric, 2-dimensional order-p tensor $\mathbf{A}$ has the following decomposition ($r = p$):

$$\mathbf{A} = \sum_{l=1}^{r} \lambda_l(\theta) \cdot \mathbf{u}_l^{\circ p}(\theta) \quad \mathbf{u}_l = \begin{bmatrix} \cos(\theta + (l-1)\pi/p) \\ \sin(\theta + (l-1)\pi/p) \end{bmatrix} \tag{5}$$

In this case, a simple line search for θ in the interval $[0, \pi/p)$ is sufficient to optimally fit the decomposition to the tensor with zero error. Employing Gram–Schmidt orthogonalization, the linear combination coefficient vector $\boldsymbol{\lambda}$ is uniquely identified by the inner-product matrix between the basis rank-1 symmetric tensor pairs and the inner-product vector between the target tensor and the basis tensors, i.e., at the optimal decomposition, $\boldsymbol{\lambda}(\theta) = \mathbf{B}^{-1}\mathbf{c}(\theta)$. Here the matrix $\mathbf{B}$ (invariant with respect to θ, since the pairwise angles between the basis vectors leading to the basis rank-1 symmetric tensors are fixed by the frame) and the vector $\mathbf{c}$ are defined elementwise as follows, assuming Frobenius tensor inner product as in (4):

$$B_{ij} = \langle \mathbf{u}_i(\theta), \quad \mathbf{u}_j(\theta) \rangle^p, \quad c_i(\theta) = \langle \mathbf{u}_i^{\circ p}(\theta), \quad \mathbf{A} \rangle_F \tag{6}$$

where $i,j = 1, \ldots, p$. Specifically note that each entry of $\mathbf{B}$ reduces to the following: $B_{ij} = \cos^p((i-j)\pi/p)$. For symmetric matrices, this matrix is simply identity.

2.3. Order-p n-dimensional symmetric tensors

The number of free elements of a symmetric n-dimensional order-p tensor is given by $m(n,p) = \binom{n+p-1}{p}$. Based on the two special cases examined above we conclude that the decomposition of any symmetric tensor should consist of some fixed frame of vectors rotated in n-dimensional space and any angle between pairwise vectors should be constant and depends on order p. As in matrices, we need q rotation angles so we attempt to decompose symmetric n-dimensional order-p tensor as a finite sum of rank-1 tensors as in (1). The number of vectors in this decomposition is $r = \binom{n+p-1}{p} - \binom{n}{2}$. defined value of r is higher or equal to the upper bound of tensor rank in [13].

To obtain the decomposition numerically, we construct a frame of r initial vectors $\mathbf{F}$ and optimize the rotation angles θ such that the Frobenius norm of the error tensor is minimized (to zero). In the spirit of block coordinate descent and fixed point algorithms, for a given candidate frame orientation, the linear combination coefficients are always obtained using (6) and $\boldsymbol{\lambda}(\theta) = \mathbf{B}^{-1}\mathbf{c}(\theta)$. The frame consists of vectors that are recursive

rotations of, for instance, the first column of the n-dimensional identity matrix multiplied by a rotation matrix with an angle of π/p in consecutive dimension index pairs. Specifically, this leads to a system of r vectors, defined as columns of $\mathbf{F}$, denoted by $\mathbf{f}_i$, using the following recursion:

$$\text{for } i=2:r; \quad \text{for } l=\text{mod}(i+n-3,n-1)+1; \quad \mathbf{f}_i=\mathbf{R}_{l,l+1}(\pi/p)\mathbf{f}_{i-1} \tag{7}$$

where $\mathbf{f}_1=\begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}^T$ and $\mathbf{R}_{l,l+1}(\pi/p)$ is a rotation matrix in the plane $(l,l+1)$, $\{l=1,\ldots,n\}$ with angle π/p. Note that the number of vectors in this frame is the same as the number of vectors needed: r, which yields a full-rank $\mathbf{B}$ in (6). This iterative procedure for frame construction gives us a result just for r steps and does not engage an elimination procedure as in our previous work that used an overcomplete frame [1]. The optimization is done iteratively and in a fixed point manner, updating the linear combination coefficients and updating the rotation matrix for the frame of vectors in order to minimize the error tensor Frobenius norm. At each iteration, basis vectors are expressed as (all rotations multiply from left) follows:

$$\mathbf{U}=\left(\prod_{k=1}^{q}\mathbf{R}(\theta_k)\right)\mathbf{F} \tag{8}$$

3. Squared Frobenius norm of decomposition error

In this section we focus on the Frobenius norm of the error (simply referred to as the error from now on) between the actual tensor $\mathbf{A}$ and its current decomposition estimate $\mathbf{T}$: $e=||\mathbf{A}-\mathbf{T}||$. The decomposition of the tensor is achieved by iteratively minimizing the squared error (SE):

$$e^2=\sum_{l_1=1}^{n}\cdots\sum_{l_p=1}^{n}(\mathbf{A}_{l_1\ldots l_p}-\mathbf{T}_{l_1\ldots l_p})^2 \tag{9}$$

Due to the higher order polynomials involved, we do not have a closed form solution for the rotation angles that yield $e(\theta)=0$; therefore we will employ iterative numerical optimization methods. For simpler notation SE in (9) can be reformulated in matrix form using the usual vectorization operator vec(·) that returns a vector whose elements are taken *column*-wise starting from the right-side operand. So a target tensor $\mathbf{T}$ in vector form is $\mathbf{t}=vec(\mathbf{T})$ and the vectorized ith rank-1 symmetric basis tensor obtained from the corresponding basis vector is $\mathbf{x}_i=vec(\mathbf{u}_i^{\circ p})$. Collecting $\mathbf{x}_i$ in the columns of $\mathbf{X}$ and defining $\mathbf{e}=\mathbf{t}-X\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}=\mathbf{B}^{-1}\mathbf{X}^T\mathbf{t}$, SE takes the form $e^2=\mathbf{e}^T\mathbf{e}$. The vectorized error tensor is equivalently

$$\mathbf{e}=(\mathbf{I}-\mathbf{X}\mathbf{B}^{-1}\mathbf{X}^T)\mathbf{t} \tag{10}$$

or with $\mathbf{B}=\mathbf{X}^T\mathbf{X}$ and $\mathbf{X}\mathbf{B}^{-1}\mathbf{X}^T=\mathbf{I}$ (after some algebra):

$$e^2=\mathbf{t}^T(\mathbf{I}-\mathbf{X}\mathbf{B}^{-1}\mathbf{X}^T)\mathbf{t} \tag{11}$$

The gradient of SE with respect to the rotation angles is

$$\nabla e_k^2(\boldsymbol{\theta})=-\mathbf{t}^T([\partial\mathbf{X}/\partial\theta_k]\mathbf{B}^{-1}\mathbf{X}^T+\mathbf{X}\mathbf{B}^{-1}[\partial\mathbf{X}^T/\partial\theta_k])\mathbf{t} \tag{12}$$

The kth column of the Jacobian matrix $\mathbf{J}$ is

$$\mathbf{j}_k(\boldsymbol{\theta})=\frac{\partial\mathbf{e}}{\partial\theta_k}=-([\partial\mathbf{X}/\partial\theta_k]\mathbf{B}^{-1}\mathbf{X}^T+\mathbf{X}\mathbf{B}^{-1}\left[\partial\mathbf{X}^T/\partial\theta_k\right])\mathbf{t} \tag{13}$$

The Hessian matrix for SE consists of the following entries:

$$\mathbf{H}_{kl}(\theta)=-\mathbf{t}^T\left\{\begin{array}{l}\dfrac{\partial^2\mathbf{X}}{\partial\theta_k\partial\theta_l}\mathbf{B}^{-1}\mathbf{X}^T+\dfrac{\partial\mathbf{X}}{\partial\theta_k}\mathbf{B}^{-1}\dfrac{\partial\mathbf{X}^T}{\partial\theta_l}\\ \quad+\dfrac{\partial\mathbf{X}}{\partial\theta_l}\mathbf{B}^{-1}\dfrac{\partial\mathbf{X}^T}{\partial\theta_k}+\mathbf{X}\mathbf{B}^{-1}\dfrac{\partial^2\mathbf{X}^T}{\partial\theta_k\partial\theta_l}\end{array}\right\}\mathbf{t} \tag{14}$$

The ith column of $\mathbf{X}$ is obtained by vectorizing the rotated ith frame vector $\mathbf{f}_i$ according to $\mathbf{u}_i^{\circ p}=(\mathbf{R}(\theta)\mathbf{f}_i)^{\circ p}$. Consequently, we need to evaluate $\partial\mathbf{x}_k/\partial\theta_l=vec(\partial(\mathbf{u}_k^{\circ p})/\partial\theta_l)$ using $\partial\mathbf{R}(\theta)/\partial\theta_i=\mathbf{R}(\theta_1)\mathbf{R}(\theta_2)\cdots[\partial\mathbf{R}(\theta_i)/\partial\theta_i]\cdots\mathbf{R}(\theta_q)$ to obtain $\partial(\mathbf{R}(\theta)\mathbf{f}_k)/\partial\theta_i=[\partial\mathbf{R}(\theta)/\partial\theta_i]\mathbf{f}_k$ and then the chain rule to get

$$\begin{aligned}\partial\mathbf{u}_k^{\circ p}/\partial\theta_i&=\partial((\mathbf{R}(\theta)\mathbf{f}_k)^{\circ p}/\partial\theta_i\\&=[\partial(\mathbf{R}(\theta)\mathbf{f}_k)/\partial\theta_i]\cdot(\mathbf{R}(\boldsymbol{\theta})\mathbf{f}_k)\cdots(\mathbf{R}(\theta)\mathbf{f}_k)+\\&\quad\ldots+(\mathbf{R}(\theta)\mathbf{f}_k)\cdot(\mathbf{R}(\theta)\mathbf{f}_k)\cdots[\partial(\mathbf{R}(\theta)\mathbf{f}_k)/\partial\theta_i]\end{aligned} \tag{15}$$

Similarly, using the chain rule appropriately, second partial derivatives can be obtained to construct the Hessian matrix.

Let $\mathbf{R}(\theta_k)$ be the rotation matrix in plane $\{i,j\}$. Then the sth partial derivative of $\mathbf{R}(\theta_k)$ with respect to angle θ_k is $\partial^s\mathbf{R}(\theta)/\partial^s\theta_k=\mathbf{M}^s\mathbf{R}(\theta)$, where $\mathbf{M}_{mn}=\delta_{mj}\delta_{ni}-\delta_{mi}\delta_{nj}$. In this case we see that all odd and even partial derivatives of rotation matrix $\mathbf{R}(\theta)$ are equal to first and second derivatives, respectively, and are alternating sequences.

4. Optimization algorithms

In this part we briefly describe the algorithms that we use to minimize SE in (11). Optimization algorithms for the widely known Tucker and CP tensor decomposition models are described in [6,8]. The most widely used method for CP and Tucker decompositions is the iterative alternative least square algorithm. This algorithm allows CP and Tucker models to evaluate their decomposition components (each basis vector) in alternating/random order. We cannot use this algorithm because our proposed frame of basis vectors is geometrically constrained and one needs to iterate all vectors simultaneously as they are not independent of each other. We evaluated the performance of four standard iterative-descent algorithms.

Jacobi rotation: Jacobi rotation updates one Givens angle at each iteration using line search or other optimization technique. On the kth iteration, we update angle θ_i using gradient descent where the stepsize α satisfies Wolfe's conditions.

Steepest descent: At each iteration, all Givens angles are updated along the direction of negative gradient.

Gauss–Newton algorithm: Inverse-Hessian provides directional correction in quadratic cost surface regions at the cost of significant increase in computational complexity.

Levenberg–Marquardt algorithm: The Jacobian of the error vector is used to approximate the Hessian to yield convergence rate similar to Gauss–Newton, but at a computational cost comparable to the steepest descent, where $[J(\theta_k)]^{-1}$ is pseudo-inverse of $J(\theta_k)$:.

5. Numerical results

Evaluation of optimization algorithms was performed on rank-1 and full-rank tensors with specific structure as well as in a Monte Carlo fashion using tensors generated randomly. Decompositions were approximated to a prespecified accuracy and the numbers of iterations are compared. First, we demonstrate the nonlinear nature of the optimization problem at hand using two simple symmetric tensors: (i) a tensor of ones, which is a rank-1 tensor since it is the p-way outer product of a vector of ones; (ii) a symmetric full-rank n-dimensional order-p tensor.

A visual illustration of the error surface along vector frame in the cross-section of rotation angles θ_1 and θ_2 for the 3-dimensional order-2 (matrix) and order-3 symmetric tensors of integer entries (type-ii) is provided in Fig. 1. The surface complexity compared to the matrix case is evident. We need to note that an error of decomposition of order-2 any-dimensional or any-order 2-dimensional tensor along any rotation angle θ_i is

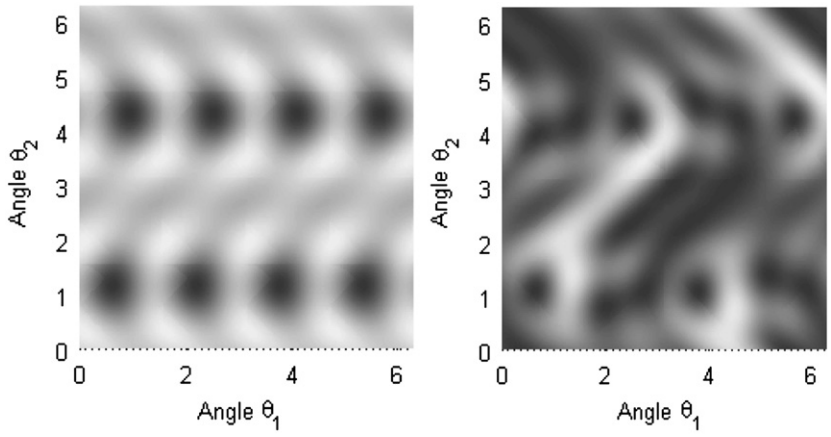

Fig. 1. Error surfaces along angles θ_1 and θ_2 for 3-dimensional (left) order-2 and (right) order-3 tensors.

Table 1
Number of iterations to decompose rank-1 tensors.

	Jacobi rotation			Stp. Desc.			Gauss–Newton			Lev. Marq.		
*n**p*	2	3	4	2	3	4	2	3	4	2	3	4
2	1	1	1	1	1	1	1	1	1	1	1	1
3	8	51	11	9	11	3	11	3	3	7	3	3
4	21	65	63	18	52	28	10	5	5	12	3	4

Table 2
Number of iterations to decompose full rank tensors.

	Jacobi rot.			Stp. Desc.			Gauss–Newton			Lev. Marq.		
*n**p*	2	3	4	2	3	4	2	3	4	2	3	4
2	1	1	1	1	1	1	1	1	1	1	1	1
3	9	T	T	T	T	T	8	15	13	6	7	10
4	T	T	T	T	T	T	24	33	29	11	13	12

periodic (due to the nature of trigonometric function). Tables 1 and 2 contain the number of iterations that the considered algorithms need to achieve an SE less than 10^{-12} (in Matlab, the minimum SE is on the order of 10^{-30}, but optimization to that level of numerical accuracy takes very long). Table 1 shows the results of decomposition of the tensor-of-ones (type-i) and Table 2 shows the results of decomposition for the symmetric tensor of integers (type-ii) with all algorithms starting from the same initial estimates. In these tables, "T" indicates that optimization was terminated at iteration 300, the preset maximum number of iterations and convergence to the desired level had not been achieved. The number of iterations for Jacobi rotations and gradient descent grows exponentially due to the ratio of $\max(\boldsymbol{\lambda})/\min(\boldsymbol{\lambda})$.

Next, we present results from a Monte Carlo experiment in which the decompositions were optimized using iterative random search (similar to stochastic annealing) in order to demonstrate that error levels comparable to the minimum possible numerical accuracy are attainable. In Fig. 2, we show the average normalized SE (per tensor entry) for randomly generated tensors of orders 2–5 for dimensions 2–5. Errors of decomposition grow exponentially due to the number of elements in high order high dimensional tensors growing in a combinatorial fashion combined with fast numerical degradation due to ill-conditioning.

Table 3 summarizes the results of decompositions for 1000 symmetric full rank real-valued tensors with all algorithms starting from random initial estimates for parameters. Experiments reveal that for the given optimality criterion, the Levenberg–Marquardt

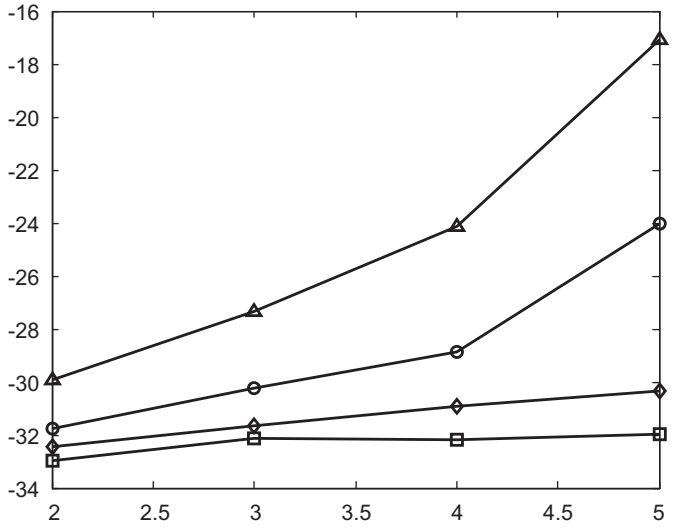

Fig. 2. Monte Carlo average $\log_{10}$ of normalized SE for the proposed eigendecomposition technique for random tensors of orders 2–5 for dimensions 2 (square), 3 (diamond), 4 (circle), 5 (triangle).

Table 3
Median number of iterations when decomposing 1000 random full rank tensors.

*n**p*	Jacobi rot.	Stp. Desc.	Gauss–Newton	Lev. Marq.
3\2	75	53	5	4
3\3	50	71	6	5
3\4	155	118	10	5
4\2	T	295	19	7
4\3	T	T	34	13
4\4	T	T	36	15

algorithm is preferable in convergence speed and computational complexity considerations.

6. Conclusion

In this paper, we propose a geometrically constrained basis vector frame that yields a set of rank-1 symmetric tensor bases that, when rotated appropriately, is able to attain a sum-of-rank-1 decomposition of any order, any dimensional symmetric tensor. The number of variables that parameterize the proposed decomposition is equal to the number of free elements (dimen-

sions) in the symmetric tensor. The proposed approach gives us upper bound of CP-rank for symmetric tensors. We also evaluated standard iterative descent techniques to determine the proposed tensor decomposition solution by optimizing the Givens angles representation of the rotation matrix of the frame of vectors. Due to the relatively complex optimization criterion (pth order polynomials squared), the Levenberg–Marquardt approach has been identified to be preferable considering computational load and convergence speed. Future work will extend the decomposition idea presented here to nonsymmetric arbitrary non-cube tensors. Initial efforts in that direction are promising.

References

[1] O. Kyrgyzov, D. Erdogmus, Geometric structure of sum-of-rank-1 decompositions for n-dimensional order-p symmetric tensors, ISCAS (2008) 1340–1343.

[2] L. De Lathauwer, B. De Moor, J. Vandewalle, A multilinear singular value decomposition, SIAM Journal on Matrix Analysis and Applications 21 (4) (2000) 1253–1278.

[3] L. Minati, D. Aquino, Probing neural connectivity through diffusion tensor imaging (DTI), Cybernetics and Systems (2006) 263–268.

[4] R.A. Harshman, Foundations of the PARAFAC procedure: models and conditions for an explanatory multimodal factor analysis, UCLA Working Papers in Phonetics 16 (1970) 1–84.

[5] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (1966) 279–311.

[6] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-(R1,R2,…,RN) approximation of higher-order tensors, SIAM Journal on Matrix Analysis and Applications 21 (4) (2000) 1324–1342.

[7] H.H. Goldstine, F.J. Murray, J. von Neumann, The Jacobi Method for Real Symmetric Matrices 6 (1) (1959) 59–96.

[8] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Review 51 (3) (2009).

[9] M. De Vos, L. De Lathauwer, B. Vanrumste, S. Van Huffel, W. Van Paesschen, Canonical decomposition of ictal scalp EEG and accurate source localisation: principles and simulation study, Computational Intelligence and Neuroscience (2007) 1–8.

[10] P. Comon, B. Mourrain, Decomposition of quantics in sums of powers of linear forms, Signal Processing 53 (2) (1996) 96–107.

[11] A. Edelman, H. Murakami, Polynomial roots from companion matrix eigenvalues, Mathematics of Computation 64 (1995) 763–776.

[12] J.D. Carroll, J.J. Chang, Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition, Psychometrika 35 (1970) 283–319.

[13] P. Comon, Tensor decompositions, Mathematics in Signal Processing (2002) 1–24.

Olexiy Kyrgyzov was born on 11 October 1979 in Russia. He received the Bachelor and Master Science degrees in Information Management Systems and Technologies from the Dnipropetrovsk National University (DNU), Dnipropetrovsk, Ukraine, in 2001 and 2002, respectively.

Between June 2002 and April 2006, he worked as an instructor at the Radiophysics department and a service engineer at the International department of DNU under the supervision of Prof. Valeriy Dolgov, Prof. Alexandr Akhmetshin, and Prof. Mihail Dyachenko.

Since 2006, he has been working towards his Ph.D. and been involved in investigation as a research assistant in the Computer Science and Electrical Engineering.

Department at Oregon Health and Science University under the supervision of Prof. Deniz Erdogmus. In 2008, together with his dissertation supervisor and other Ph.D. students, he moved to Boston (MA) and was transferred to the Department of Electrical and Computer Engineering at Northeastern university.

His current research interests broadly include nonlinear and statistical signal processing, machine learning, image processing, computer vision, and pattern recognition areas.

Mr. Kyrgyzov serves as a Reviewer for IEEE Transactions on Neural Networks, Journal of Neurocomputing (Elsevier), and IEEE International Symposium on Circuits And Systems.

Deniz Erdogmus graduated with B.S. in EE and the B.S. in Mathematics in 1997, and M.S. in EE in 1999 from the Middle East Technical University, Ankara, Turkey. He received his Ph.D. in ECE from the University of Florida in 2002, where he stayed as a postdoctoral research associate until 2004. He is currently an Assistant Professor in the ECE Department at Northeastern University. Prior to this position he held the same rank jointly at the CSEE and BME Departments of the Oregon Health and Science University. His research focuses on statistical signal processing and machine learning with applications to cognitive signal processing including brain computer interfaces and technologies that collaboratively improve human performance. He has over 60 journal publications and he has been serving as an associate editor for IEEE and Elsevier journals in these areas.