A Survey of Tensor Methods

Lieven De Lathauwer
Katholieke Universiteit Leuven
Science and Technology, 8500 Kortrijk, Belgium
and E.E. Dept. (ESAT), B-3001 Leuven, Belgium
Email: Lieven.DeLathauwer@kuleuven-kortrijk.be and Lieven.DeLathauwer@esat.kuleuven.be

***Abstract*— Matrix decompositions have always been at the heart of signal, circuit and system theory. In particular, the Singular Value Decomposition (SVD) has been an important tool. There is currently a shift of paradigm in the algebraic foundations of these fields. Quite recently, Nonnegative Matrix Factorization (NMF) has been shown to outperform SVD at a number of tasks. Increasing research efforts are spent on the study and application of decompositions of higher-order tensors or multi-way arrays. This paper is a partial survey on tensor generalizations of the SVD and their applications. We also touch on Nonnegative Tensor Factorizations.**

I. Introduction

We use $\mathbb{K}$ to denote $\mathbb{R}$ or $\mathbb{C}$ when the difference is not important. We have the following well-known decomposition.

Definition 1. A *Singular Value Decomposition (SVD)* of a matrix $\mathbf{T} \in \mathbb{K}^{I \times J}$ is a decomposition of the form

$$\mathbf{T} = \mathbf{A} \cdot \mathbf{D} \cdot \mathbf{B}^H, \tag{1}$$

in which $\mathbf{A} \in \mathbb{K}^{I \times I}$ and $\mathbf{B} \in \mathbb{K}^{J \times J}$ are orthogonal (unitary) and in which $\mathbf{D} \in \mathbb{R}^{I \times J}$ is diagonal, containing only nonnegative entries, put in nonincreasing order.

The SVD concentrates a remarkable number of powerful properties. Unitarity (orthogonality) of $\mathbf{A}$ and $\mathbf{B}$ guarantee excellent numerical properties. $\mathbf{D}$ is the simplest representation of $\mathbf{T}$ that preserves all geometric information. The decomposition is rank-revealing, i.e., the rank of $\mathbf{T}$ is equal to the number of nonzero entries of $\mathbf{D}$. Consequently, $\mathbf{T}$ can be expressed as the following minimal sum of rank-1 terms:

$$\mathbf{T} = \sum_r d_{rr} \mathbf{a}_r \mathbf{b}_r^H, \tag{2}$$

in which $\mathbf{a}_r$ and $\mathbf{b}_r$ denote the rth column of $\mathbf{A}$ and $\mathbf{B}$, respectively.

It comes as no surprise that the SVD is ubiquitous in signal processing, data mining, system theory, control, scientific computing, etc. [30], [31], [79]. Two of the most well-known applications are Principal Component Analysis [43] and Latent Semantic Analysis [17].

In recent years, researchers have started to realize that many phenomena are inherently multi-way. In such a case, tensors are more natural data representations than matrices — stacking the data in a matrix results in loss of information. Tensor decompositions often have better uniqueness properties than matrix decompositions, which makes that they are often easier to interpret. Multilinear algebra is richer than vector/matrix algebra, which means that more information can be extracted. Roughly speaking, generalizing different properties of the matrix SVD leads to different tensor decompositions. In Sections III, IV and V we will discuss three possible generalizations, and touch on some of their applications. More background material can be found in [3], [44], [49], [50], [71].

Let us turn back to matrix decompositions. Sometimes orthogonality of the left and right singular vectors makes the interpretation of the SVD difficult. Real-valued data are sometimes nonnegative (frequency counts, pixel intensities, spectra, ...). In such cases, the following decomposition may be more appropriate.

Definition 2. A *Nonnegative Matrix Factorization (NMF)* of a matrix $\mathbf{T} \in \mathbb{R}^{I \times J}$ is a decomposition of the form

$$\mathbf{T} = \mathbf{A} \cdot \mathbf{B}^T, \tag{3}$$

in which $\mathbf{A} \in \mathbb{R}^{I \times R}$ and $\mathbf{B} \in \mathbb{R}^{J \times R}$ only contain nonnegative entries.

This decomposition was introduced in [52], [53], [60]. Nonnegative components have the advantage that they cannot partially compensate each other. For this reason, NMF is usually considered as a "sum-of-parts" representation. For more information we refer to [7], [11], [12], [34], [45] and the references therein. Section VI of this paper is a short note on nonnegative versions of the decompositions discussed in Sections III and IV.

II. Multilinear Algebraic Prerequisites

We first introduce some basic multilinear algebraic material.

Definition 3. Consider $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ and $\mathbf{A} \in \mathbb{K}^{J_1 \times I_1}$, $\mathbf{B} \in \mathbb{K}^{J_2 \times I_2}$, $\mathbf{C} \in \mathbb{K}^{J_3 \times I_3}$. Then the *Tucker mode-1 product* $\mathcal{T} \bullet_1 \mathbf{A}$, *mode-2 product* $\mathcal{T} \bullet_2 \mathbf{B}$ *and mode-3 product* $\mathcal{T} \bullet_3 \mathbf{C}$ are defined by

$$\begin{aligned}
(\mathcal{T} \bullet_1 \mathbf{A})_{j_1 i_2 i_3} &= \sum_{i_1=1}^{I_1} t_{i_1 i_2 i_3} a_{j_1 i_1}, \qquad \forall j_1, i_2, i_3, \\
(\mathcal{T} \bullet_2 \mathbf{B})_{i_1 j_2 i_3} &= \sum_{i_2=1}^{I_2} t_{i_1 i_2 i_3} b_{j_2 i_2}, \qquad \forall i_1, j_2, i_3, \\
(\mathcal{T} \bullet_3 \mathbf{C})_{i_1 i_2 j_3} &= \sum_{i_3=1}^{I_3} t_{i_1 i_2 i_3} c_{j_3 i_3}, \qquad \forall i_1, i_2, j_3,
\end{aligned}$$

respectively.

In this notation, (1) is written as $\mathbf{T} = \mathbf{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B}^*$, in which $\cdot^*$ denotes the complex conjugate.

Definition 4. The *outer product* $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ of three vectors $\mathbf{a} \in \mathbb{K}^I$, $\mathbf{b} \in \mathbb{K}^J$ and $\mathbf{c} \in \mathbb{K}^K$ is defined by $t_{ijk} = a_i b_j c_k$, for all values of the indices. We write $\mathcal{T} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c}$.

In this notation, (2) is written as $\mathbf{T} = \sum_r d_{rr}\, \mathbf{a}_r \circ \mathbf{b}_r^*$.

Definition 5. A *mode-n vector* of a tensor $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \times I_3}$ is an I_n-dimensional vector obtained from $\mathcal{T}$ by varying the index i_n and keeping the other indices fixed.

Mode-n vectors generalize "column" and "row vectors".

Definition 6. The *mode-n rank* of a tensor $\mathcal{T}$ is the dimension of the subspace spanned by its mode-n vectors.

978-1-4244-3828-0/09/$25.00 ©2009 IEEE

The mode-n rank of a higher-order tensor is the obvious generalization of the column (row) rank of a matrix.

Definition 7. A third-order tensor has *multilinear rank* (L, M, N) if its mode-1 rank, mode-2 rank and mode-3 rank are equal to L, M and N, respectively.

A rank-$(1, 1, 1)$ tensor is briefly called "rank-1". This implies that a third-order tensor $\mathcal{T}$ has rank 1 if it equals the outer product of 3 nonzero vectors.

An other way to generalize the concept of rank from matrices to tensors is as follows.

Definition 8. The *outer product rank* of a tensor $\mathcal{T}$ is the minimal number of rank-1 tensors that yield $\mathcal{T}$ in a linear combination.

The outer product rank may be estimated by means of the CORCONDIA procedure [8]. In some cases it is equal to the rank of a matrix representation of the tensor [25].

III. The Tucker Decomposition

Definition 9. A *Tucker Decomposition* of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ is a decomposition of $\mathcal{T}$ of the form

$$\mathcal{T} = \mathcal{D} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C}. \tag{4}$$

This decomposition was introduced in [77], [78]. It is not unique. For instance, if $\mathbf{A}$ is post-multiplied by a square nonsingular matrix $\mathbf{F}$, then this can be compensated by replacing $\mathcal{D}$ by $\mathcal{D} \bullet_1 \mathbf{F}^{-1}$. Part of the degrees of freedom can be used to make $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ columnwise orthonormal. One particular constrained version of the Tucker decomposition can be obtained by computing $\mathbf{A}$ as the matrix of left singular vectors of an $(I \times JK)$ matrix in which all the columns of $\mathcal{T}$ are stacked one after the other; $\mathbf{B}$ and $\mathbf{C}$ are obtained by working with the rows and mode-3 vectors, respectively. It is demonstrated in [19] that this particular constrained version of the Tucker decomposition is a striking generalization of the matrix SVD. In this case the Tucker Decomposition is multilinear rank-revealing, i.e., if $\mathcal{T}$ has multilinear rank (L, M, N), then $\mathcal{D} \in \mathbb{K}^{L \times M \times N}$. We also mention that the column space of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ is unique.

In many applications, one wishes to approximate a given tensor by a tensor with prespecified multilinear rank. This is a natural tensor generalization of the best rank-R approximation of matrices. While the optimal approximation of a matrix can be obtained by truncation of its SVD, the optimal tensor approximation cannot in general be obtained by truncation of the Tucker decomposition. However, truncation of the particular constrained version [19] usually yields a pretty good approximation. This estimate may be further refined by means of the algorithms discussed in [20], [36], [41], [42], [46], [50], [67]. One may also start from random initial guesses.

Tucker approximation is useful for dimensionality reduction of large tensor datasets. The actual data analysis can then be carried out in a space of lower dimensions [23]. Very often, Tucker compression precedes the analysis that is discussed in Sections IV and V. Tucker approximation is also important when one wishes to estimate signal subspaces from tensor data. Applications include fuzzy modelling [5], harmonic retrieval [62], [63], image processing [80], [81] and classification [66], to give just a few examples.

IV. The Canonical or Parallel Factor Decomposition

Definition 10. A *Canonical or Parallel Factor Decomposition (CP)* of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ is a decomposition of $\mathcal{T}$ as a linear combination of rank-1 terms:

$$\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r. \tag{5}$$

The fully symmetric variant of PARAFAC, in which $\mathbf{a}_r = \mathbf{b}_r = \mathbf{c}_r$, $r = 1, \ldots, R$, was studied in the nineteenth century in the context of invariant theory [15]. The unsymmetric decomposition was introduced in 1927 [38], [39]. Around 1970, the unsymmetric decomposition was independently reintroduced in Psychometrics [10] and Phonetics [37].

We will focus on the case where (5) involves a minimal number of rank-1 terms. CP is then by definition outer product rank-revealing. To a large extent, the practical importance of CP stems from its uniqueness properties. It is clear that the decomposition can at most be unique up to permutation of the rank-1 terms and up to scaling/counterscaling of the factors in the same term. However, CP is "essentially unique" under quite mild conditions [25], [51], [70], [72], [73]. In particular, it is not necessary to impose (artificial) orthogonality constraints. This makes that CP can often easily be interpreted. Of course, orthogonality can be imposed when useful, but then one gets a different decomposition [14], [22], [47], [54]. Algorithms for the computation of the CP decomposition are discussed in [24], [25], [56], [58], [61], [65], [71], [82] and references therein. It is worth mentioning that, for real-valued tensors, minimization of the residu in (5) may cause CP terms go to infinity, see [32], [74], [75] and references therein.

The success story of CP started in Chemometrics and food industry [71]. CP has also found important applications in signal processing and data analysis [50]. In wireless telecommunications, it provides powerful means for the exploitation of different types of diversity [68], [69]. It also describes the basic structure of higher-order cumulants (or sets of covariance matrices) of multivariate data on which all algebraic methods for independent component analysis are based [6], [9], [14], [21], [22], [40]. CP has further been useful for the analysis of biomedical signals [2], [4], [33], text mining [35], the analysis of social networks [1] and the analysis of web hyperlinks [48], to name just a few applications.

V. Block Term Decompositions

Definition 11. A *decomposition of a tensor $\mathcal{T} \in \mathbb{K}^{I \times J \times K}$ in a sum of multilinear rank-(L, M, N) terms* is a decomposition of $\mathcal{T}$ of the form

$$\mathcal{T} = \sum_{r=1}^{R} \mathcal{D}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r, \tag{6}$$

in which $\mathcal{D}_r \in \mathbb{K}^{L \times M \times N}$ are full multilinear rank-(L, M, N), and in which $\mathbf{A}_r \in \mathbb{K}^{I \times L}$ (with $I \geqslant L$), $\mathbf{B}_r \in \mathbb{K}^{J \times M}$ (with $J \geqslant M$) and $\mathbf{C}_r \in \mathbb{K}^{K \times N}$ (with $K \geqslant N$) are full column rank, $1 \leqslant r \leqslant R$.

This decomposition is an example of a block term decomposition (BTD). The idea is that the given tensor "block" $\mathcal{T}$, of dimensions $(I \times J \times K)$, is decomposed in a sum of "building blocks" of inherently smaller size $(L \times M \times N)$. BTDs were introduced in [28]. The BTD framework unifies the Tucker and CP decompositions. Namely, if there is only one term ($R = 1$), then (6) is a Tucker decomposition of $\mathcal{T}$. On the other hand, if all the blocks have a scalar core ($L = M = N = 1$), then (6) is a CP decomposition. The BTD framework also sheds new light on how the concept of rank can be generalized from matrices to tensors. Namely, one should specify the number of terms (R) *and* their size (L, M, N). In the definition of multilinear rank, one implicitly assumes that $R = 1$. In the definition of outer product rank, one implicitly assumes that $L = M = N = 1$.

In decomposition (6), the order of the terms is arbitrary. Each term can be normalized in the same manner as a Tucker decomposition. It turns out that, apart from the trivial indeterminacies, BTDs are "essentially unique" under some conditions. This is a profound difference between matrices and tensors. Uniqueness conditions are derived in [28], [57]. Algorithms for the computation of BTDs are derived in [29], [57], [58], [59].

The advantage of BTD over CP is that the terms are more general, and hence potentially allow to model more complicated data structures. This is illustrated by the wireless communication application discussed in [16], [26], [59], [68]. In cases without Inter-Symbol-Interference (ISI), the signals transmitted by the different users can be separated by computing a CP decomposition [68]. If ISI is caused by reflections in the far field of the antenna array, a simple BTD can be used ($L = M$ and $N = 1$) [26]. If reflections occur both in the far field and close to the antenna array, then a more complicated BTD can be resorted to [16], [59].

VI. Imposing nonnegativity

Like for matrices, it sometimes makes sense for tensors to assume that the factors in a decomposition are nonnegative. For the Tucker decomposition, this is done in [11], [83]. Nonnegative CP is studied in [13], [45], [55], [64].

VII. Conclusion

Multilinear algebra is a fascinating new field of research, with potentially many applications in signals, circuits and systems. In this tutorial paper we have briefly discussed three tensor generalizations of the matrix SVD. We have also touched on tensor generalizations of the NMF.

Acknowledgment

Research supported by: (1) Research Council K.U.Leuven: GOA-Ambiorics, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT1/08/023, (2) F.W.O.: (a) project G.0321.06, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization", 2007–2011), (4) EU: ERNSI.

References

[1] E. Acar, S.A. Çamtepe, and B. Yener, "Collective sampling and analysis of high-order tensors for chatroom communications," *Proc. IEEE Int. Conf. on Intelligence and Security Informatics*, Lecture Notes in Computer Science, vol. 3975, pp. 213-224, 2006.

[2] E. Acar, C. A. Bingol, H. Bingol, R. Bro, and B. Yener, "Multiway analysis of epilepsy tensors," *Bioinformatics*, vol. 23, no. 13, pp. i10–i18, 2007.

[3] E. Acar and B. Yener, "Unsupervised multiway data analysis: a literature survey," *IEEE Trans. Knowledge and Data Engineering*, to appear.

[4] A.H. Andersen and W.S. Rayens, "Structure-seeking multilinear methods for the analysis of fMRI data," *NeuroImage*, vol. 22, 2004, pp. 728–739.

[5] P. Baranyi, Y. Yam, D. Tikk, and R.J. Patton, "Trade-off between approximation accuracy and complexity: TS controller design via HOSVD based complexity minimization," *Studies in fuzziness and soft computing*, vol. 128, 2003, pp. 249–277.

[6] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[7] M. Berry, M. Browne, A. Langville, P. Pauca, R. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Comp. Stat. and Data Anal.*, vol. 52, 2007, pp. 155–173.

[8] R. Bro, H.A.L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemometrics*, vol. 17, no. 5, 2003, pp. 274–286.

[9] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proc.-F*, vol. 140, 1994, pp. 362–370.

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

[11] D. Chen and R. Plemmons, "Nonnegativity constraints in numerical analysis," in: A. Bultheel and R. Cools, Eds., *Proc. Symposium on the Birth of Numerical Analysis*, Leuven, Belgium, Oct. 2007, World Scientific Press, to appear.

[12] A. Cichocki, R. Zdunek and S. Amari, "Nonnegative matrix and tensor factorization," *IEEE Signal Processing Magazine*, Jan. 2008, pp. 142–145.

[13] A. Cichocki, A.H. Phan, and C. Caiafa, "Flexible HALS algorithms for sparse non-negative matrix/tensor factorization," *Proc. Conf. for Machine Learning for Signal Processing*, Oct. 16–19, 2008, Cancun, Mexico.

[14] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, April 1994.

[15] P. Comon and B. Mourrain, "Decomposition of quantics in sums of powers of linear forms," Signal Process., vol. 53, 1996, pp. 93–108.

[16] A.L.F. de Almeida, G. Favier and J.C.M. Mota, "PARAFAC-based unified tensor modeling for wireless communication systems with application to blind multiuser equalization," *Signal Proc.*, vol. 87, 2007, pp. 337–351.

[17] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, vol. 41, no. 6, Jan. 1999, pp. 391–407.

[18] B. De Moor, M. Moonen, Eds., *SVD and Signal Processing, III. Algorithms, Applications and Architectures*, Elsevier, Amsterdam, 1995.

[19] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, No. 4, April 2000, pp. 1253–1278.

[20] 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 Anal. Appl.*, vol. 21, No. 4, April 2000, pp. 1324–1342.

[21] L. De Lathauwer, B. De Moor, and J. Vandewalle, , "An introduction to independent component analysis," *J. Chemometrics*, vol. 14, 2000, pp. 123–149.

[22] L. De Lathauwer, B. De Moor, J. Vandewalle, "Independent Component Analysis and (Simultaneous) Third-Order Tensor Diagonalization", *IEEE Trans. Signal Processing*, vol. 49, no. 10, Oct. 2001, pp. 2262–2271.

[23] L. De Lathauwer and J. Vandewalle, "Dimensionality reduction in higher-order signal processing and rank-$(R_1, R_2, \ldots, R_N)$ reduction in multilinear algebra", *Lin. Alg. Appl.*, vol. 391, Nov. 2004, pp. 31–55.

[24] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 26, pp. 295–327, 2004.

[25] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.

[26] L. De Lathauwer and A. de Baynast, "Blind deconvolution of DS-CDMA signals by means of decomposition in rank-$(1, L, L)$ terms," *IEEE Trans. Signal Processing*, vol. 56, no. 4, April 2008, pp. 1562–1571.

[27] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part I: Lemmas for partitioned matrices," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1022–1032.

[28] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1033–1066.

[29] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms — Part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1067–1083.

[30] B. De Moor and M. Moonen, Eds., *SVD and Signal Processing, III. Algorithms, Applications and Architectures*, Elsevier, Amsterdam, 1995.

[31] E.F. Deprettere, Ed., *SVD and Signal Processing. Algorithms, Applications and Architectures*, North-Holland, Amsterdam, 1988.

[32] V. de Silva and L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1084–1127.

[33] M. De Vos, A. Vergult, L. De Lathauwer, W. De Clercq, S. Van Huffel, P. Dupont, A. Palmini, and W. Van Paesschen, "Canonical decomposition of ictal EEG reliably detects the seizure onset zone", *Neuroimage*, vol. 37, no. 3, Sep. 2007, pp. 844–854.

[34] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?," *Proc. 17th Annual Conf. on Neural Information Processing Systems (NIPS 2003)*, Vancouver, Whistler, Canada.

[35] D.M. Dunlavy, T.G. Kolda, and and W.P. Kegelmeyer, "Multilinear algebra for analyzing data with multiple linkages,". Tech. Rep. SAND2006-2079, Sandia National Laboratories, Albuquerque, NM and Livermore, CA, April 2006.

[36] L. Eldén and B. Savas, A Newton–Grassmann method for computing the best multilinear rank-(r_1, r_2, r_3) approximation of a tensor, *SIAM J. Matrix Anal. Appl*, to appear.

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

[38] F.L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, no. 1, 1927, pp. 164–189.

[39] F.L. Hitchcock, "Multiple invariants and generalized rank of a p-way matrix or tensor," *J. Math. Phys.*, vol. 7, no. 1, 1927, pp. 39–79.

[40] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[41] M. Ishteva, L. De Lathauwer, P.-A. Absil, and S. Van Huffel, "Dimensionality reduction for higher-order tensors: algorithms and applications," *Int. Journal of Pure and Applied Mathematics*, vol. 42, No. 3, 2008, pp. 337–343.

[42] M. Ishteva, L. De Lathauwer, P.-A. Absil, S. Van Huffel, "Differential-geometric Newton algorithm for the best rank-(R_1, R_2, R_3) approximation of tensors," *Numerical Algorithms*, to appear.

[43] I. Joliffe, *Principal Component Analysis*, John Wiley & Sons, 2005.

[44] H. Kiers and I. Van Mechelen, "Three-way component analysis: principles and illustrative application," *Psychological Methods*, vol. 6, 2001, pp. 84–110.

[45] D. Kim, S. Sra, and I.S. Dhillon, "Fast projection-based methods for the least squares nonnegative matrix approximation problem," *Statistical Analysis and Data Mining*, vol. 1, no. 1, Feb. 2008, pp. 38–51.

[46] E. Kofidis and P.A. Regalia, "On the best rank-1 approximation of higher-order supersymmetric tensors," *SIAM J. Matrix Anal. Appl.*, vol. 23, 2002, pp. 863–884.

[47] T.G. Kolda, "Orthogonal tensor decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 23, 2001, pp. 243–255.

[48] T.G. Kolda, B.W. Bader, and J.P. Kenny, "Higher-order web link analysis using multilinear algebra," ICDM 2005: *Proc. 5th IEEE Int. Conf. on Data Mining (ICDM 2005)*, Nov. 2005, pp. 242–249.

[49] T.G. Kolda and B.W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, to appear.

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

[51] J.B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Lin. Alg. Appl.*, vol. 18, 1977, pp. 95–138.

[52] D.D. Lee and H. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, 1999, pp. 788–791.

[53] D.D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," *Advances in Neural Information Processing Systems*, vol. 13, 2001, pp. 556–562.

[54] C.D. Moravitz Martin and C.F. Van Loan, "A Jacobi-type method for computing orthogonal tensor decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1219–1232.

[55] M. Mørup, L.K. Hansen, S.M. Arnfred, "Algorithms for sparse non-negative Tucker," *Neural Computation*, vol. 20, no. 8, 2008, pp. 2112–2131.

[56] C. Navasca, L. De Lathauwer, and S. Kindermann, "Swamp reducing technique for tensor decomposition," *Proc. 16th European Signal Processing Conference (EUSIPCO 2008)*, Aug. 25–29, 2008, Lausanne, Switzerland.

[57] D. Nion and L. De Lathauwer, "A tensor-based blind DS-CDMA receiver using simultaneous matrix diagonalization," *Proc. VIII IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2007)*, June 17–20, 2007, Helsinki, Finland.

[58] D. Nion and L. De Lathauwer, "An enhanced line search scheme for complex-valued tensor decompositions. Application in DS-CDMA." *Signal Processing*, vol. 88, No. 3, March 2008, pp. 749–755.

[59] D. Nion and L. De Lathauwer, "Block Component Model Based Blind DS-CDMA Receivers," *IEEE Trans. Signal Processing*, to appear.

[60] P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, 1994, pp. 111–126.

[61] P. Paatero, "The multilinear engine — A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model," *Journal of Computational and Graphical Statistics*, vol. 8, pp. 854–888, 1999.

[62] J.-M. Papy, L. De Lathauwer, and S. Van Huffel, "Exponential data fitting using multilinear algebra. The single-channel and multichannel case." *Num. Lin. Alg. Appl.*, vol. 12, no. 8, Oct. 2005, pp. 809–826.

[63] J.-M. Papy, L. De Lathauwer, and S. Van Huffel, "Exponential data fitting using multilinear algebra: the decimative case." *J. Chemometrics*, to appear.

[64] A.H. Phan and A. Cichocki, "Fast and efficient algorithms for nonnegative Tucker decomposition," ISNN-2008, Sept. 21–26, 2008, Beijing, China, Springer LNCS 2008.

[65] M. Rajih, P. Comon, R.A. Harshman, "Enhanced line search: a novel method to accelerate parafac," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, 2008, pp. 1128–1147.

[66] B. Savas and L. Eldén, "Handwritten digit classification using higher order singular value decomposition," *Pattern Recognition*, vol. 40, no. 3, March 2007, pp. 993–1003.

[67] B. Savas and L.-H. Lim, Best multilinear rank approximation of tensors with quasi-Newton methods on Grassmannians, Tech. Rep. LITH-MAT-R-2008-01-SE, Dept. Mathematics, Linköpings Universitet, 2008.

[68] N.D. Sidiropoulos, G.B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. on Signal Processing*, vol. 48, no. 3, pp. 810–823, March 2000.

[69] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Processing*, vol. 48, 2000, pp. 2377–2388.

[70] N. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *Journal of Chemometrics*, vol. 14, 2000, pp. 229–239.

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

[72] A. Stegeman, J.M.F. ten Berge, and L. De Lathauwer, "Sufficient conditions for uniqueness in candecomp/parafac and indscal with random component matrices," *Psychometrika*, vol. 71, no. 2, June 2006, pp. 219–229.

[73] A. Stegeman and N.D. Sidiropoulos, "On Kruskal's uniqueness condition for the candecomp/parafac decomposition," *Lin. Alg. Appl.*, vol. 420, 2007, pp. 540–552.

[74] A. Stegeman, "Low-rank approximation of generic $p \times q \times 2$ arrays and diverging components in the Candecomp/Parafac model," *SIAM J. Matris Anal. Appl.*, vol. 30, no. 3, 2008, pp. 988–1007.

[75] A. Stegeman and L. De Lathauwer, "A method to avoid diverging components in the Candecomp/Parafac model for generic $I \times J \times 2$ arrays," *SIAM J. Matrix Anal. Appl.*, to appear.

[76] G. Tomasi and R. Bro, "A comparison of algorithms for fitting the PARAFAC model," *Comp. Stat. & Data Anal.*, vol. 50, 2006, pp. 1700–1734.

[77] L.R. Tucker, "The extension of factor analysis to three-dimensional matrices", in: H. Gulliksen and N. Frederiksen, Eds., *Contributions to mathematical psychology*, Holt, Rinehart & Winston, NY, 1964, pp. 109–127.

[78] L.R. Tucker, "Some mathematical notes on three-mode factor analysis", *Psychometrika*, vol. 31, 1966, pp. 279–311.

[79] R. Vaccaro, Ed., *SVD and Signal Processing, II. Algorithms, Applications and Architectures*, Elsevier, Amsterdam, 1991.

[80] M.A.O. Vasilescu, D. Terzopoulos, "Multilinear subspace analysis for image ensembles," *Proc. Computer Vision and Pattern Recognition Conf. (CVPR '03)*, vol.2, Madison, WI, June 2003, pp. 93–99.

[81] M.A.O. Vasilescu, D. Terzopoulos, "TensorTextures: multilinear image-based rendering," *Proc. ACM SIGGRAPH 2004*, Los Angeles, CA, August 2004, pp. 336–342.

[82] S.A. Vorobyov, Y. Rong, N.D. Sidiropoulos, A.B. Gershman, "Robust iterative fitting of multilinear models," *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 2678–2689, Aug. 2005.

[83] M. Welling and M. Weber, "Positive tensor factorization," *Pattern Recogn. Lett.*, vol. 22, no. 12, pp. 1255–1261, 2001.