ECCTD'03 - European Conference on Circuit Theory and Design, September 1-4, 2003, Cracow, Poland

The Generalized Higher Order Singular Value Decomposition and the Oriented Signal-to-signal Ratios of Pairs of Signal Tensors and their Use in Signal Processing

Joos Vandewalle[1] , Lieven De Lathauwer[2], and Pierre Comon[3]

Abstract – **Two new generalizations of tensor concepts for signal processing are presented. These generalizations are typically relevant for applications where one tensor consists of valuable measured data or signals, that should be retained, while the second tensor contains data or information that should be rejected. First the higher order singular value decomposition for a single tensor is extended to pairs of tensors; this is the multilinear equivalent of the generalized or quotient SVD (GSVD, QSVD) for pairs of matrices. Next the notion of oriented signal-to-signal ratios that was derived for pairs of matrices is extended to pairs of tensors. These signal to signal ratios can be linked to the previously defined generalized higher order singular value decomposition.**

1 INTRODUCTION

In recent years more and more instances of applications occur, where the data have more than two indices and hence are not organized in a matrix but in a tensor, also called a multiway or multidimensional array. Let us mention here psychometrics [16], chemometrics [7, 8, 15, 17] and statistical signal and image processing [6, 9, 10, 18]. In typical image applications, the 4 different indices of a 4th order image tensor can be the x, y, t and color axes. A recent application is the websearch tensor. Here we will mainly work in a signal processing, but many concepts can be carried over to the other domains. Typically the methods of matrix theory and related numerical computations [1] are no longer adequate and valuable for tensors. Therefore a number of studies [10-14, 19-20] have been performed to extend some matrix concepts to tensors, like the higher order singular value decomposition or the canonical decomposition of a tensor. Also the use of tensors for finding independent components [9] in signals is a topic of current interest.

Many applications occur where the measured data lead to two tensors A and B rather than to a single tensor A. Typically one tensor (called A) contains the valuable information, whereas the other (called B) contains information that is irrelevant and hence should be rejected. This is the class of problems that are tackled in this paper. It turns out that there are several ways to extend the relevant matrix concepts, and that their extensions are not trivial.

The matrix counterpart of this paper, was already known in the 80ies with the work of Golub and Van Loan [1] on the Generalized Singular Value Decomposition of a matrix pencil and its numerically reliable computation. We studied [5] the oriented signal to signal ratios for the matrix case, their relevance for signal processing [2-5] and related computational issues. In [10, 11] we discussed a possible multilinear generalization of the Singular Value Decomposition (SVD), called the Higher Order SVD (HOSVD). The different n-rank values can easily be read from this decomposition. In [12-14, 19] some techniques to compute the least-squares approximation of a given tensor by a tensor with rank 1 or with a prespecified n-rank are discussed.

In section 2 the generalized higher order singular value decomposition of a pair of tensors is presented, which generalizes the HOSVD. The oriented signal to signal ratio of a tensor pair is described in section 3 and is brought in relation to the generalized HOSVD. The paper ends with conclusions and views on signal processing applications.

Let us conclude the introduction with some basic definitions and a comment on the unavoidably complicated notation that is used. The scalar product $<\mathcal{A},\mathcal{B}>$ of two tensors $\mathcal{A},\mathcal{B} \in \Re^{I_1 \times I_2 \times \ldots \times I_N}$ is defined in a straightforward way as

$$<\mathcal{A},\mathcal{B}> = \sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\ldots\sum_{i_N=1}^{I_N} a_{i_1 i_2 \ldots i_N}\, b_{i_1 i_2 \ldots i_N} \qquad (1)$$

The Frobenius-norm of a tensor is defined as $\|\mathcal{A}\| = \sqrt{<\mathcal{A},\mathcal{A}>}$ Two tensors are called orthogonal when their scalar product is zero.

In order to facilitate the distinction between scalars, vectors, matrices and higher-order tensors, the type of a quantity will be reflected by its representation: scalars are denoted by lower-case letters (a, b...), vectors are written as lower-case italic letters (*a,b,..*), matrices correspond to bold-face capitals (**A, B**...) and tensors are written as calligraphic letters ($\mathcal{A}$, $\mathcal{B}$...). This notation is consistently used for entries of a given structure. The entry with row index i and

[1] Department of Electrical Engineering,(ESAT), Kasteelpark Arenberg 10, Katholieke Universiteit Leuven B3001, Belgium. e-mail: [joos.vandewalle,lieven.delathauwer]@esat.kuleuven.ac.be, tel.: +32 16 321052, fax: +32 16321970].
[2] ETIS, CNRS, 6 av du Ponceau, 95014 Cergy-Pontoise cedex, France [lieven.delathauwer@ensea.fr]
[3] I3S, CNRS, 2000 route des Lucioles, Sophia-Antipolis cedex, France [comon@i3s.unice.fr].

column index j in a matrix $\mathbf{A}$ i.e. $(\mathbf{A})_{ij}$ is symbolized by a_{ij}. Also $(\mathbf{A})_i = a_i$ and $(\mathcal{A})_{i1\ i2\dots\ iN} = a_{i1\ i2\dots\ iN}$. Furthermore, the i-th column vector of a matrix $\mathbf{A}$ is denoted as A_i, i.e., $\mathbf{A}= [A_1\ A_2\ \dots]$. To enhance the overall readability, we have made one exception to this rule: as we frequently use the characters i, j, r and n in the meaning of indices (counters), I, J, R and N will be reserved to denote the index upper bounds, unless stated otherwise. We define *the n-mode product* of a tensor $\mathcal{A} \in \Re^{I_1 \times I_2 \times \dots \times I_N}$ with a matrix $\mathbf{U} \in \Re^{J_n x I_n}$ as the $(I_1 \times I_2 \times .. I_{n-1} \times J_n \times I_{n+1} \times .. I_N)$ tensor with entries

$$\left(\mathcal{A} \times_n \mathbf{U}\right)_{i_1 i_2 .. i_{n-1} j_n i_{n+1} .. i_N} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 .. i_{n-1} i_n i_{n+1} .. i_N} u_{j_n i_n}.$$

Although the results can be easily generalized for complex tensors and matrices, we describe these here only for real, in order not to complicate the notation.

2 GENERALIZED HIGHER ORDER SINGULAR VALUE DECOMPOSITION OF A MATRIX PAIR

2.1 The multilinear SVD or HOSVD of a tensor

A multilinear SVD (HOSVD) of a tensor has recently been discussed in [10,11]. Through an unfolding (see figure 1) of an N-th order tensor as a matrix we can apply the regular matrix SVD and obtain the orthogonal or unitary transformation matrices. This matrix unfolding cuts the tensor resp. along vertical, frontal and horizontal planes into patches, and pastes these patches resp. into the matrices $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$ and $\mathbf{A}_{(3)}$

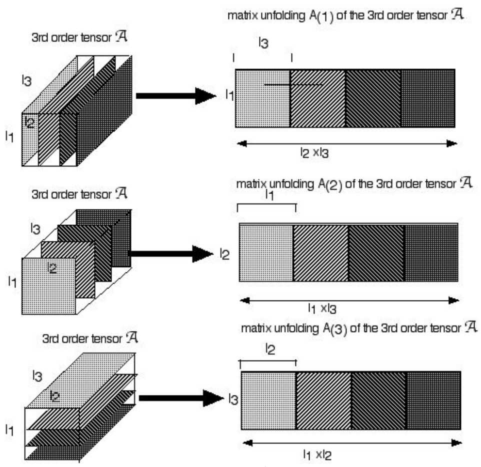

Figure 1 Matrix unfolding of a 3rd order tensor $\mathcal{A}$ produces 3 matrices $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$ and $\mathbf{A}_{(3)}$ to which the matrix SVD can be applied.

Theorem 1 [HOSVD] Every $(I_1 x\ I_2 x \dots I_N)$ tensor $\mathcal{A}$ can be decomposed as the product of N or- thogonal matrices and an all-orthogonal tensor $\mathcal{S}$, i.e.,

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)} \qquad (2)$$

in which all $\mathbf{U}^{(n)}$ are orthogonal $(I_n x I_n)$ matrices for n=1,..N, and in which $\mathcal{S}$ is an $(I_1 x\ I_2 x \dots I_N)$ tensor whose subtensors $\mathcal{S}_{in=\alpha}$ and $\mathcal{S}_{in=\beta}$ are orthogonal for all possible values of n, α and β subject to $\alpha\neq\beta$, i.e.,

$$\left\langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \right\rangle = 0 \text{ when } \alpha\neq\beta \qquad (3)$$

The subtensors of S are ordered as follows:

$$\left\|\mathcal{S}_{i_n=1}\right\| \geq \left\|\mathcal{S}_{i_n=2}\right\| \geq \dots \left\|\mathcal{S}_{i_n=I_N}\right\| \geq 0 \qquad (4)$$

for all possible values of n.

Observe that the remaining tensor $\mathcal{S}$ is not diagonal unlike the matrix case, but the information in the original tensor $\mathcal{A}$ is compressed in an all-orthogonal tensor, in which, generally speaking, the strongest contributions occur for the smallest values of the indices.

2.2 The generalized multilinear SVD or HOSVD of a tensor pair

When two tensors $\mathcal{A}$ and $\mathcal{B}$ have the same range for one of their indices (here we assume that this is the first index I_1), then we can define the *new concept of the generalized HOSVD of this pair.*

Theorem 2 [*Generalized HOSVD*] Given an arbitrary pair of tensors with the same first index I_1, say, $(I_1 x\ I_2 x \dots I_N)$ tensor $\mathcal{A}$ and $(I_1 x\ J_2 x \dots J_M)$ tensor $\mathcal{B}$, then the generalized HOSVD of this tensor pair exists and is basically unique, where $\mathcal{A}$ and $\mathcal{B}$ are decomposed each as product of orthogonal matrices and one nonsingular $(I_1 x\ I_1)$-matrix $\mathbf{W}$ and all-orthogonal tensors $\mathcal{S}$ and $\mathcal{R}$, i.e.,

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{W} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)} \qquad (5)$$

$$\mathcal{B} = \mathcal{R} \times_1 \mathbf{W} \times_2 \mathbf{V}^{(2)} \dots \times_M \mathbf{V}^{(M)} \qquad (6)$$

where all $\mathbf{U}^{(n)}$ for n=2,..N (resp. $\mathbf{V}^{(n)}$ for n=2,..M) are orthogonal $(I_n x I_n)$ (resp. $(J_n x J_n)$) matrices, and in which $\mathcal{S}$ (resp. $\mathcal{R}$) is a real $(I_1 x\ I_2 x \dots I_N)$ tensor of which the subtensors $\mathcal{S}_{in=\alpha}$ and $\mathcal{S}_{in=\beta}$ are orthogonal for all possible values of n, α and β when $\alpha\neq\beta$, i.e.,

$$\left.\begin{array}{l} \left\langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \right\rangle = 0 \\ \left\langle \mathcal{R}_{i_n=\alpha}, \mathcal{R}_{i_n=\beta} \right\rangle = 0 \end{array}\right. \text{when } \alpha\neq\beta \qquad (7)$$

The subtensors of S and R are ordered as follows:

$$\left\|\mathcal{S}_{i_1=1}\right\| / \left\|\mathcal{R}_{i_1=1}\right\| \geq \left\|\mathcal{S}_{i_1=2}\right\| / \left\|\mathcal{R}_{i_1=2}\right\| \geq \dots \geq \left\|\mathcal{S}_{i_1=I_1}\right\| / \left\|\mathcal{R}_{i_1=I_1}\right\| \geq 0 \qquad (8).$$

The proof stems from the fact that we can apply the generalized SVD [1, 4] on the matrix unfolding along the first index of the tensors $\mathcal{A}$ and $\mathcal{B}$ and the regular SVD on the matrix unfoldings along the other indices of $\mathcal{A}$ and $\mathcal{B}$. Observe that the matrix **W** is not orthogonal, like in the matrix case. So its columns or rows are not mutually orthogonal. The conditions under which a tensor pair with more than one common dimension have a similar decomposition, with more than one nonsingular matrix **W**, are still under investigation.

3 ORIENTED SIGNAL TO SIGNAL RATIOS OF TENSOR PAIRS

3.1 The oriented signal to signal ratio of a matrix pair

In [5] the notion of oriented energy of a vector signal was defined as follows. We consider an (mxn) matrix A**,** whose rows typically correspond to sensors and whose columns to time instants when the signals of the sensors are measured. The oriented energy $E_e(A)$ is then the energy that is sensed in the direction of a unit vector e, i. e.,

$$E_e(\mathbf{A}) = \sum_{i=1}^{n} (e^T a_i)^2 = \left\| e^T \mathbf{A} \right\|^2 \qquad (9)$$

Of course the oriented energy varies for varying orientations of the unit vector e. In figure 2 the left picture shows how the oriented energy for two matrices **A** and **B** varies for a case where m=2. It turns out [5] that the directions of maximal oriented energy of a matrix **A** are mutually orthogonal and correspond to the directions of the left singular vectors of the SVD of **A**. This allows to find vectors and subspaces where a maximal signal contribution is present in the vector signal. Next, one can define the oriented signal to signal ratio of a pair of matrices **A** and **B** as follows

$$E_e(\mathbf{A})/E_e(\mathbf{B}) = \sum_{i=1}^{n} (e^T a_i)^2 / \sum_{i=1}^{n} (e^T b_i)^2 \qquad (10)$$

Again the oriented signal to signal ratio varies for varying sensing direction e (see right picture of figure 2). The direction of maximal signal to signal ratio shows a direction of a linear combination of rows where the signal **A** is relatively more pronounced than signal **B**. When the generalized SVD (GSVD or also called QSVD) [1-5] of this matrix pair (**A**, **B**) is computed, the rows of the nonsingular matrix $\mathbf{W}^{-1}$ correspond with the directions of maximal oriented signal to signal ratio. Since the matrix is nonsingular but not necessarily orthogonal, these directions are not mutually orthogonal.

3.2 The oriented signal to signal ratio of a tensor pair

The oriented energy $E_e(\mathcal{A})$ of a tensor signal $\mathcal{A}$ is the energy that is sensed in the direction of a unit vector e i. e.

$$E_e(\mathcal{A}) = \sum_{i_2=1}^{I_2} \cdots \sum_{i_n=1}^{I_N} \left(\sum_{i_1=1}^{I_1} (e_{i_1} a_{i_1 i_2 \ldots i_n}) \right)^2 = \left\| \mathcal{A} \times_1 e^T \right\|_F^2 \qquad (11)$$

Of course the oriented energy varies for varying orientations of the unit vector e. The directions of maximal oriented energy of a tensor $\mathcal{A}$ are mutually orthogonal and correspond with the directions of the left singular vectors of the SVD of 1-unfolding $\mathbf{A}_{(1)}$ of the tensor $\mathcal{A}$. These directions correspond also with the columns of $\mathbf{U}^{(1)}$ in the HOSVD of the tensor $\mathcal{A}$. Next, one can define the oriented signal to signal ratio of a pair of tensors, $\mathcal{A}$ and $\mathcal{B}$, that have the same first index I_1, as follows

$$\frac{E_e(\mathcal{A})}{E_e(\mathcal{B})} = \frac{\sum_{i_2=1}^{I_2} \cdots \sum_{i_n=1}^{I_N} \left(\sum_{i_1=1}^{I_1} (e_{i_1} a_{i_1 i_2 \ldots i_n}) \right)^2}{\sum_{i_2=1}^{I_2} \cdots \sum_{i_n=1}^{I_N} \left(\sum_{i_1=1}^{I_1} (e_{i_1} b_{i_1 i_2 \ldots i_n}) \right)^2} = \frac{\left\| \mathcal{A} \times_1 e^T \right\|_F^2}{\left\| \mathcal{B} \times_1 e^T \right\|_F^2} \qquad (12)$$

t index I_1 , as follows

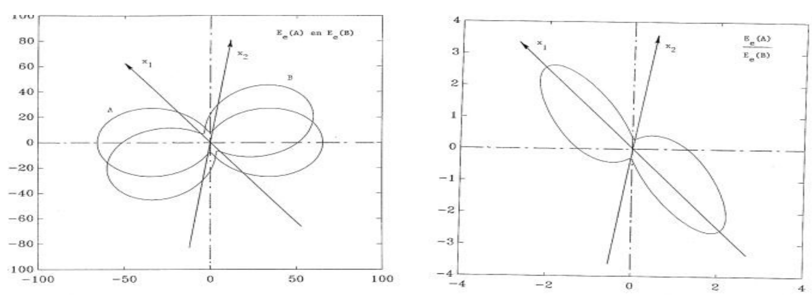

Figure 2 (left) oriented energy plots, (right) oriented signal-to-signal plot

Again the oriented signal to signal ratio varies for varying sensing direction *e*. The direction of maximal signal to signal ratio shows a direction of a linear combination of rows where the signal $\mathcal{A}$ is relatively more pronounced than signal $\mathcal{B}$. When the generalized HOSVD of the tensor pair ($\mathcal{A}$, $\mathcal{B}$) is computed according to theorem 2, it turns out that the rows of $\mathbf{W}^{-1}$ correspond to these directions. Again these are not mutually orthogonal.

4 CONCLUSIONS

Several notions that are relevant for tensor signal pairs have been defined. It would be nice to extend these even for the cases where more than one index among the pair of tensors is in common.

Acknowledgments

This work is supported by several institutions:
(1) Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology), IDO, (2) the Fund for Scientific Research-Flanders (F.W.O.) projects G.0240.99 (Multilinear Generalisations of the Singular Value Decomposition and Applications in Signal Processing and System Identification) and G.0407.02 support vector machines, and G.0262.97; the F.W.O. Research Communities ICCoS (Identification and Control of Complex Systems) and ANMMM (Advanced Numerical Methods for Mathematical Modelling), (3) the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs - Interuniversity Poles of Attraction Programme (IUAP V-22 (2002-2006):. (4) Lieven De Lathauwer and Pierre Comon hold a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); L. De Lathauwer has also an honorary post-doctoral research position with the K.U.Leuven. Joos Vandewalle is a Full Professor with the K.U.Leuven, currently on a sabbatical leave at the I3S laboratory of the CNRS at Sophia Antipolis, France sponsored by F.W.O.

References

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

[2] Doclo S., Moonen M., ``GSVD-Based Optimal Filtering for Single and Multi-Microphone Speech Enhancement'', *IEEE Trans on Signal Processing*, vol. 50, no. 9, Sep. 2002, pp. 2230-2244.

[3] Moonen M., Vandewalle J., ``A QSVD approach to on- and off-line state space identification'', *Int. Journal of Control*, vol.51, no.5, 1990, pp. 1133-1146.

[4] De Moor B., ``On the structure of generalized singular value and QR decompositions'', *SIAM J. Matrix Analysis & Applications*, vol.15-1, Jan. 1994, pp. 347-358.

[5] De Moor B., Staar J., Vandewalle J., ``Oriented energy and oriented signal-to-signal ratio concepts in the analysis of vector sequences and time series'', in *`SVD and Signal Processing: Algorithms, Applications and Architectures'*, E. Deprettere (ed.), Elsevier Science Publishers B.V. (North-Holland), 1988, pp. 209-232.

[6] C.J. Appellof and E.R. Davidson, "Strategies for analyzing data from video fluoromatric monitoring of liquid chromatographic effluents", *Analytical Chemistry*, 53, 1981, pp.~2053--2056.

[7] R. Bro, "PARAFAC. Tutorial & applications" Chemom. Intell. Lab. Syst., Special Issue 2nd Internet Conf. in Chemometrics (INCINC'96), 38, 1997, pp.~149--171.

[8] D. Burdick, X. Tu, L. McGown, and D. Millican, "Resolution of multicomponent fluorescent mixtures by analysis of the excitation-emission-frequency arrays", *Journal of Chemometrics*, 4 ,1990, pp.15-28.

[9] P. Comon, "Independent component analysis, a new concept?" *Signal Process., Special Issue on Higher Order Statistics,* 36, 1994, pp. 287-314.

[10] L. De Lathauwer, *Signal Processing Based on Multilinear Algebra*, Ph.D. thesis, K.U.Leuven, E.E.Dept. (ESAT), Belgium, 1997.

[11] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition", *SIAM J. Matrix Anal. Appl.,* 21, 2000, pp. 1253--1278.

[12] L. De Lathauwer, B. De Moor, and J. Vande- walle, "On the best rank-1 and rank-R1,R2, approximation of higher-order tensors", *SIAM J. Matrix Anal. Appl.,* 21, 2000, pp. 1324--1342.

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

[14] T. Kolda, "A counterexample to the possibility of an extension of the Eckart-Young low-rank approximation theorem for the orthogonal rank tensor decomposition", accepted for publication in: *SIAM J. Matrix Anal. Appl.*

[15]E. Sanchez and B.R. Kowalski, "Tensorial resolutiona direct trilinear decomposition", *Journal of Chemometrics,* 4, 1990, pp. 29-45.

[16] R. Sands and F. Young, "Component models for three-way data : an alternating least squares algorithm with optimal scaling features", *Psychometrika,* 45, 1980, pp. 39-67.

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

[18] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation", *IEEE Trans. Signal Processing*, 50, 2002, pp.1545-1553.

[19] T. Zhang and G.H. Golub, "Rank-one approximation to high order tensors", *SIAM J. Matrix Anal. Appl.,* 23, 2001, pp. 534--550.

[20] P. Comon, ``Tensor Decompositions,'' In J. G. McWhirter and I. K. Proudler, eds, Mathematics in Signal Processing V, pp. 1-24. Clarendon Press, Oxford, UK, 2002.