Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 1996-93

From matrix to tensor: multilinear algebra and signal processing[1]

Lieven De Lathauwer and Bart De Moor[2]

December 1996

Keynote paper
4th Int. Conf. on Mathematics in Signal Processing
University of Warwick, UK, Dec. 17-19, 1996
Conference Digest, Part I, pp. 1-11.

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-96-93.ps.Z*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 18 05, Fax 32/16/32 19 70, WWW: *http://www.esat .kuleuven.ac.be/sista*. E-mail: *Lieven.DeLathauwer@esat.kuleuven.ac.be*, *Bart.DeMoor@esat.kuleuven.ac.be*.
This work was supported in part by the Flemish Government (Concerted Research Action MIPS), the Fund for Scientific Research - Flanders (FWO: G.0292.95), the Belgian Government (Interuniversity Attraction Poles IUAP-17 and IUAP-50), the European Commission (Human Capital and Mobility Network SIMONET) and the European Community Research program ESPRIT, Basic Research Working Group nr. 6620 (ATHOS) and the Flemish Institute for Support of Scientific-Technological Research in Industry (I.W.T). L. De Lathauwer is a Research Assistant supported by the I.W.T.; B. De Moor is a Research Associate of the F.W.0. and an associate professor at the K.U.Leuven.

Abstract

In recent years research on Higher-Order Statistics has raised some fundamentally new insights in the area of mathematical signal processing. From an algebraic point of view this is reflected by an emerging interest in the assets of multinear algebra, and the development of numerical "tensor" tools, complementary to the arsenal of existing matrix techniques. We give an overview of some important tensor algebraic results, and some of their implications in signal processing.

The algebraic track of our argument starts from the question how the EVD/SVD-concept could be generalized to the higher-order case. It appears that the rank properties of higher-order tensors, which are very different from their matrix counterparts, do not allow for an easy extension: instead, the generalization of different EVD/SVD-properties can lead to different multilinear decompositions. As far as orthogonal decompositions are concerned, we discuss the transformations to "all-orthogonality" resp. "maximal diagonality". In addition, we investigate the expansion of a higher-order tensor in a sum of non-orthogonal rank-1 components.

These multilinear techniques find a major application in the problem of Blind Source Separation, also known as Independent Component Analysis. Orthogonal tensor decompositions can be used to obtain the solution from the covariance matrix and a higher-order cumulant tensor of the data. As an alternative, non-orthogonal tensor decompositions allow to solve the problem by resorting only to the higher-order cumulant, in a way that is conceptually blind to additive Gaussian noise. A second basic type of problem that can be tackled with multilinear algebra, is the Blind Factor Analysis of multiway datasets. The signal processing part of the exposition is illustrated with real-life examples.

From matrix to tensor: multilinear algebra and signal processing*

Lieven De Lathauwer   Bart De Moor

K.U.Leuven - E.E. Dept.- ESAT - SISTA
Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
tel: 32/16/321805   fax: 32/16/321970
email: Lieven.DeLathauwer@esat.kuleuven.ac.be
Bart.DeMoor@esat.kuleuven.ac.be
URL: http://www.esat.kuleuven.ac.be/sista/sista.html

Abstract

In recent years research on Higher-Order Statistics has raised some fundamentally new insights in the area of mathematical signal processing. From an algebraic point of view this is reflected by an emerging interest in the assets of multinear algebra, and the development of numerical "tensor" tools, complementary to the arsenal of existing matrix techniques. We give an overview of some important tensor algebraic results, and some of their implications in signal processing.
The algebraic track of our argument starts from the question how the EVD/SVD-concept could be generalized to the higher-order case. It appears that the rank properties of higher-order tensors, which are very different from their matrix counterparts, do not allow for an easy extension: instead, the generalization of different EVD/SVD-properties can lead to different multilinear decompositions. As far as orthogonal decompositions are concerned, we discuss the transformations to "all-orthogonality" resp. "maximal diagonality". In addition, we investigate the expansion of a higher-order tensor in a sum of non-orthogonal rank-1 components.
These multilinear techniques find a major application in the problem of Blind Source Separation, also known as Independent Component Analysis. Orthogonal tensor decompositions can be used to obtain the solution from the covariance matrix and a higher-order cumulant tensor of the data. As an alternative, non-orthogonal tensor decompositions allow to solve the problem by resorting only to the higher-order cumulant, in a way that is conceptually blind to additive Gaussian noise. A second basic type of problem that can be tackled with multilinear algebra, is the Blind Factor Analysis of multiway datasets. The signal processing part of the exposition is illustrated with real-life examples.

*Research supported by the Flemish Government (Concerted Research Action MIPS), the Fund for Scientific Research - Flanders (FWO: G.0292.95), the Belgian Government (Interuniversity Attraction Poles IUAP-17 and IUAP-50), the European Commission (Human Capital and Mobility Network SIMONET) and the European Community Research program ESPRIT, Basic Research Working Group nr. 6620 (ATHOS) and the Flemish Institute for Support of Scientific-Technological Research in Industry (I.W.T). L. De Lathauwer is a Research Assistant supported by the I.W.T.; B. De Moor is a Research Associate of the F.W.0. and an associate professor at the K.U.Leuven.

1 Introduction

Multilinear algebra is the algebra of higher-order tensors. In this paper an $(I_1 \times I_2 \times \ldots \times I_N)$-tensor is intuitively considered as an ordered $(I_1 \times I_2 \times \ldots \times I_N)$-table of numerical values, given with respect to the bases chosen in $\mathbb{R}^{I_1}$, $\mathbb{R}^{I_2}$, ..., $\mathbb{R}^{I_N}$.

In the past, famous scientists like Gauss, Kronecker, Cayley, Weyl and Hilbert have indirectly contributed to the development of tensor algebra through research on homogeneous polynomials: their results can be casted in a tensor framework via the isomorphism between homogeneous polynomials and super-symmetric tensors (which are invariant under arbitrary index permutations), in the same way as quadratic forms and symmetric matrices are related. An interesting discussion of the early literature can be found in [6]. Recently, (numerical) multilinear algebra has received a great deal of attention through research in Higher-Order Statistics (HOS), where the basic quantities (higher-order moments, cumulants and spectra) of a stochastic vector are super-symmetric higher-order tensors [17].

Literature in fundamental multilinear algebra is often based on the "index-notation" and "Einstein summation convention" (e.g. [15], Chapt. 1-3). It is a matter of practice to get acquainted with this notation, and to learn to benefit from its assets. For this introductory paper however, we have chosen to use a matrix-like notation.

In Section 2 we briefly introduce some basic material, such as the definition of scalar product and outer product, the concept of tensor rank, and the way in which a tensor changes by multiplication with a matrix. Sections 3 and 4 deal with orthogonal transformations of a tensor to make it "all-orthogonal" or "maximal diagonal" respectively. Section 5 discusses a non-orthogonal tensor transformation, used to obtain an expansion in a minimal number of rank-1 components. Section 6 explains how these tensor decompositions can be used in Independent Component Analysis (ICA), which is a refinement of the concept of Principal Component Analysis (PCA). In Section 7 we describe a second fundamentally new signal processing tool, namely the Blind Factor Analysis (BFA) of multiway datasets. The exposition is intended to give a very elementary introduction to the basic aspects of the topics that we have mentioned; for more details the reader is invited to consult the references.

All the concepts described in this paper, are formulated in terms of real-valued tensors. The generalization to the complex case is always possible, but from time to time it is more involved from a notational as well as computational point of view.

2 Basic Definitions

First, the definition of an outer product generalizes expressions of the type AB^T, in which A and B are vectors:

Definition 1 (Outer product) *The outer product $\mathcal{A} \circ \mathcal{B}$ of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_P}$ and a tensor $\mathcal{B} \in \mathbb{R}^{J_1 \times J_2 \times \ldots \times J_Q}$, is defined by*

$$(\mathcal{A} \circ \mathcal{B})_{i_1 i_2 \ldots i_P j_1 j_2 \ldots j_Q} \stackrel{\text{def}}{=} a_{i_1 i_2 \ldots i_P} b_{j_1 j_2 \ldots j_Q}$$

for all values of the indices.

Next, we give straightforward generalizations of the scalar product, orthogonality and Frobenius-norm:

Definition 2 (Scalar product) *The scalar product $\langle \mathcal{A}, \mathcal{B} \rangle$ of tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is defined as*

$$\langle \mathcal{A}, \mathcal{B} \rangle \stackrel{\text{def}}{=} \sum_{i_1} \sum_{i_2} \ldots \sum_{i_N} b_{i_1 i_2 \ldots i_N} a_{i_1 i_2 \ldots i_N}$$

Definition 3 (Orthogonality) *Tensors of which the scalar product equals 0, are mutually orthogonal.*

Definition 4 (Frobenius-norm) *The Frobenius-norm of a tensor $\mathcal{A}$ is given by*

$$\|\mathcal{A}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$$

Multiplications of a higher-order tensor with a matrix can be defined as follows:

Definition 5 *The mode-n product of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, denoted by $\mathcal{A} \times_n \mathbf{U}$, is an $(I_1 \times I_2 \times \ldots J_n \ldots \times I_N)$-tensor defined by*

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \ldots j_n \ldots i_N} = \sum_{i_n} a_{i_1 i_2 \ldots i_n \ldots i_N} u_{j_n i_n}$$

for all index values.

In this notation, the matrix product $\mathbf{A} = \mathbf{U}^{(1)} \cdot \mathbf{B} \cdot \mathbf{U}^{(2)^T}$ takes the form of the "symmetric" expression $\mathbf{A} = \mathbf{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)}$. The mode-$n$ product allows to express the effect of a basis transformation in $\mathbb{R}^{I_n}$ on the tensor $\mathcal{A}$.

In tensor terminology, column vectors, row vectors, ... will be denoted as mode-1 vectors, mode-2 vectors, etc. The definitions of column rank, row rank and rank of a matrix can then be generalized as follows:

Definition 6 (Mode-n rank) *The mode-n vectors of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ are the I_n-dimensional vectors obtained from $\mathcal{A}$ by varying the index i_n and keeping the other indices fixed. The mode-n rank of $\mathcal{A}$, denoted by $R_n = \mathtt{rank_n}(\mathcal{A})$, is the dimension of the vector space generated by the mode-n vectors.*

Definition 7 (Rank-1 tensor) *An Nth-order tensor $\mathcal{A}$ has rank 1 when it equals the outer product of N vectors $U^{(1)}$, $U^{(2)}$, ..., $U^{(N)}$:*

$$\mathcal{A} = U^{(1)} \circ U^{(2)} \circ \ldots \circ U^{(N)}.$$

Definition 8 (Rank) *The rank of an arbitrary Nth-order tensor $\mathcal{A}$, denoted by $R = \mathtt{rank}(\mathcal{A})$, is the minimal number of rank-1 tensors that yield $\mathcal{A}$ in a linear combination.*

Higher-order rank-related properties are thoroughly different from their matrix counterparts. The various mode-n ranks of a given tensor are not necessarily the same, and can also be different from the rank. From the definitions we have that always $R_n \leq R$. Examples are given in [13].

3 The Higher-Order SVD

A first way to generalize the SVD of matrices, is as follows:

Definition 9 (HOSVD) *The Higher-Order Singular Value Decomposition (HOSVD) of a real $(I_1 \times I_2 \times \ldots \times I_N)$-tensor $\mathcal{A}$ is given by:*

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \ldots \times_N \mathbf{U}^{(N)}, \tag{1}$$

in which:

- $\mathbf{U}^{(n)} = \left[U_1^{(n)} U_2^{(n)} \ldots U_{I_n}^{(n)}\right]$ *is an orthogonal $(I_n \times I_n)$-matrix $(1 \leq n \leq N)$.*
- $\mathcal{S}$ *is a real $(I_1 \times I_2 \times \ldots \times I_N)$-tensor of which the subtensors $\mathcal{S}_{i_n=\alpha}$, obtained by fixing the nth index to α, have the properties of:*
  - *all-orthogonality: two subtensors $\mathcal{S}_{i_n=\alpha}$ and $\mathcal{S}_{i_n=\beta}$ are orthogonal for all possible values of n, α and β subject to $\alpha \neq \beta$:*

    $$\langle \mathcal{S}_{i_n=\alpha}, \mathcal{S}_{i_n=\beta} \rangle = 0 \quad \textit{when} \quad \alpha \neq \beta$$

  - *ordering:*

    $$\|\mathcal{S}_{i_n=1}\| \geq \|\mathcal{S}_{i_n=2}\| \geq \ldots \geq \|\mathcal{S}_{i_n=R_n}\| > \|\mathcal{S}_{i_n=R_n+1}\| = \ldots = \|\mathcal{S}_{i_n=I_N}\| = 0$$

    for all possible values of n.

The Frobenius-norms $\|\mathcal{S}_{i_n=i}\|$, symbolized by $\sigma_i^{(n)}$, are called mode-n singular values of $\mathcal{A}$ and the vector $U_i^{(n)}$ is an ith mode-n singular vector. The decomposition is visualized for third-order tensors in Fig. 1.

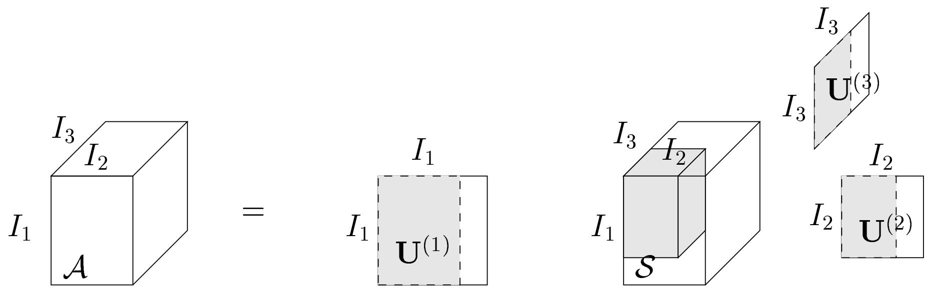

Figure 1: Visualization of the HOSVD for a third-order tensor.

This decomposition can be obtained by computing any $\mathbf{U}^{(n)}$ as the left singular matrix of a matrix $\mathbf{A}_{(n)}$, in which all the mode-n vectors are stacked as columns; the tensor $\mathcal{S}$ then follows from Eq. (1). A consequence of this strong link with the matrix decomposition, is that many properties of the tensor decomposition show a strong analogy with the matrix case as well (uniqueness, link with EVD, perturbation properties, ...) For this reason, the decomposition has been denoted as *the* HOSVD in[8].

Imposing the condition of all-orthogonality, instead of diagonality, on tensor $\mathcal{S}$, implies that the HOSVD is always defined. As a matter of fact, $\mathcal{S}$ cannot be diagonal in general, which means that the HOSVD does not necessarily reveal the rank of $\mathcal{A}$. On the other hand, the number of non-zero (significant) mode-n singular values corresponds to the mode-n rank (in a numerical sense) of $\mathcal{A}$.

4 "Maximal Diagonality"

Another way to deal with the fact that a generic tensor cannot be diagonalized by orthogonal transformations, when looking for a tensorial kind of SVD, is to claim that the transformed tensor is "as diagonal as possible" (in least-squares sense):

Definition 10 *For every real $(I_1 \times I_2 \times \ldots \times I_N)$-tensor $\mathcal{A}$, the decomposition*

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \ldots \times_N \mathbf{U}^{(N)}, \tag{2}$$

in which:

- $\mathbf{U}^{(n)} = \left[U_1^{(n)} U_2^{(n)} \ldots U_{I_n}^{(n)}\right]$ *is an orthogonal $(I_n \times I_n)$-matrix $(1 \leq n \leq N)$,*
- $\mathcal{S}$ *is a real $(I_1 \times I_2 \times \ldots \times I_N)$-tensor with the following properties:*
  - *maximal diagonality (in least-squares sense): for all $\mathcal{S}'$, obtained by an orthogonal decomposition of $\mathcal{A}$, we have that $\sum_i |s_{ii\ldots i}|^2 \geq |s'_{ii\ldots i}|^2$,*
  - *ordering:*
  $$|s_{11\ldots 1}| \geq |s_{22\ldots 2}| \geq \ldots \geq |s_{II\ldots I}|,$$
  in which $I = \mathtt{min}\{I_1, I_2, \ldots, I_N\}$,

defines a generalized SVD. The diagonal entries $s_{ii\ldots i}$ and the vectors $U_i^{(n)}$ can be considered as higher-order singular values resp. singular vectors of $\mathcal{A}$. The decomposition is visualized for third-order tensors in Fig. 2.

The computation of the decomposition, for super-symmetric higher-order tensors, was first studied by Comon, who proposed a Jacobi-type iteration algorithms [5]. Related computation schemes have been explored in [10] and [12].

5 Canonical Decomposition

An expansion of a tensor $\mathcal{A}$ as a sum of $\mathtt{rank}(\mathcal{A})$ rank-1 terms, generically involves non-orthogonal components:

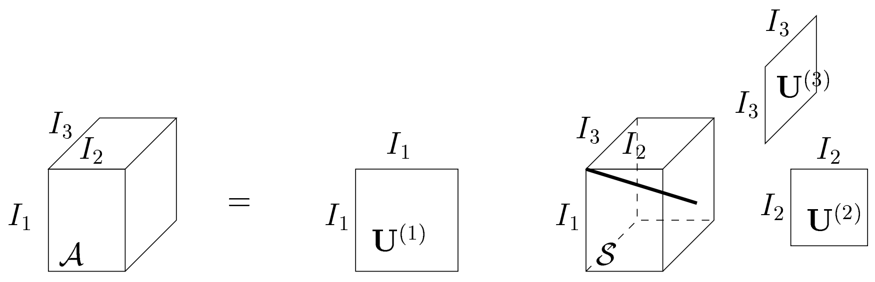

Figure 2: Visualization of the optimal diagonalization, in least-squares sense, of a third-order tensor by orthogonal transformations.

Definition 11 (CANDECOMP) *A Canonical Decomposition (CANDECOMP) or Parallel Factors (PARAFAC) Decomposition of a real $(I_1 \times I_2 \times \ldots \times I_N)$-tensor $\mathcal{A}$ is a decomposition of $\mathcal{A}$ as a linear combination of a minimal number of rank-1 terms:*

$$\mathcal{A} = \sum_r^R \lambda_r \, U_r^{(1)} \circ U_r^{(2)} \circ \ldots \circ U_r^{(N)}. \tag{3}$$

The decomposition is visualized for third-order tensors in Fig. 3.

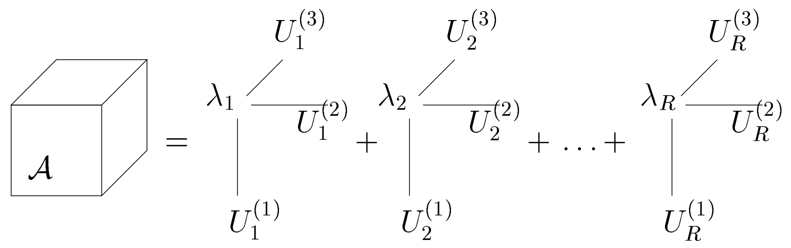

Figure 3: Visualization of the CANDECOMP for a third-order tensor.

Even if non-orthogonal rank-1 terms are allowed, the minimal number of terms is in general not bounded by $\mathtt{min}\{I_1, I_2, \ldots, I_N\}$. For generic rank values and rank upperbounds we refer to [6] (the paper addresses the case of super-symmetric tensors, but many theorems can easily be generalized for unsymmetric tensors). For computational aspects we refer to [11, 1, 2].

The uniqueness properties of the CANDECOMP are very different from (and also much more complicated than) their matrix equivalents. Where a decomposition of a matrix $\mathbf{A}$ as a sum of $\mathtt{rank}(\mathbf{A})$ rank-1 terms is usually made unique - up to some trivial indeterminacies - by imposing orthonormality constraints, the uniqueness of the CANDECOMP depends on conditions of linear independence. As an example, we assume the special but important case of an $(I \times I \times I)$-tensor $\mathcal{A}$ with rank $R = I$. It can be verified that the CANDECOMP of $\mathcal{A}$ is essentially unique if e.g. (a) the set $\{U_r^{(1)}\}_{(1 \leq r \leq R)}$ is linearly independent, (b) the set $\{U_r^{(2)}\}_{(1 \leq r \leq R)}$ is linearly independent, and (c) the set $\{U_r^{(3)}\}_{(1 \leq r \leq R)}$ does not contain collinear vectors [16].

6 Independent Component Analysis

In a first subsection we introduce some basic elements of HOS. The second subsection deals with the topic of ICA itself.

6.1 Moments and cumulants of a stochastic vector

Moment tensors of a real stochastic vector are defined as follows:

Definition 12 (Moment) *The Nth order moment tensor $\mathcal{M}_X^{(N)}$ of a real stochastic vector X is defined by the element-wise equation:*

$$(\mathcal{M}_X^{(N)})_{i_1 i_2 \ldots i_N} = \mathtt{Mom}(X_{i_1}, X_{i_2}, \ldots, X_{i_N}) \stackrel{\text{def}}{=} \mathtt{E}\{X_{i_1} X_{i_2} \ldots X_{i_N}\} \tag{4}$$

$\mathtt{E}$ denotes the expectation. On the other hand, cumulants, up to order four, of a real stochastic vector are defined as follows:

Definition 13 (Cumulant) *For a real zero-mean stochastic vector X the cumulants, up to order four, are given by:*

$$\begin{aligned}
(C_X)_i = \mathtt{Cum}(X_i) &\stackrel{\text{def}}{=} \mathtt{E}\{X_i\} && (5)\\
(\mathbf{C}_X)_{i_1 i_2} = \mathtt{Cum}(X_{i_1}, X_{i_2}) &\stackrel{\text{def}}{=} \mathtt{E}\{X_{i_1} X_{i_2}\} && (6)\\
(\mathcal{C}_X^{(3)})_{i_1 i_2 i_3} = \mathtt{Cum}(X_{i_1}, X_{i_2}, X_{i_3}) &\stackrel{\text{def}}{=} \mathtt{E}\{X_{i_1} X_{i_2} X_{i_3}\} && (7)\\
(\mathcal{C}_X^{(4)})_{i_1 i_2 i_3 i_4} = \mathtt{Cum}(X_{i_1}, X_{i_2}, X_{i_3}, X_{i_4}) &\stackrel{\text{def}}{=} \mathtt{E}\{X_{i_1} X_{i_2} X_{i_3} X_{i_4}\} - \mathtt{E}\{X_{i_1} X_{i_2}\}\mathtt{E}\{X_{i_3} X_{i_4}\} \\
&\quad - \mathtt{E}\{X_{i_1} X_{i_3}\}\mathtt{E}\{X_{i_2} X_{i_4}\} \\
&\quad - \mathtt{E}\{X_{i_1} X_{i_4}\}\mathtt{E}\{X_{i_2} X_{i_3}\} && (8)
\end{aligned}$$

For every component X_i of X that has a non-zero mean, X_i has to be replaced in these formulas, except the first one, by $X_i - \mathtt{E}\{X_i\}$.

Some crucial properties are [17, 18]:

- *Multilinearity:* if a real stochastic vector X is transformed into a stochastic vector $\tilde{X}$ by a matrix multiplication $\tilde{X} = \mathbf{A} \cdot X$, with $\mathbf{A} \in \mathbb{R}^{J \times I}$, then we have:

$$\begin{aligned}
\mathcal{M}_{\tilde{X}}^{(N)} &= \mathcal{M}_X^{(N)} \times_1 \mathbf{A} \times_2 \mathbf{A} \ldots \times_N \mathbf{A} && (9)\\
\mathcal{C}_{\tilde{X}}^{(N)} &= \mathcal{C}_X^{(N)} \times_1 \mathbf{A} \times_2 \mathbf{A} \ldots \times_N \mathbf{A} && (10)
\end{aligned}$$

- *Partitioning of independent variables:* if a subset of I stochastic variables $X_1, \ldots, X_I$ is independent of the other variables, then we have:

$$\mathtt{Cum}(X_1, X_2, \ldots, X_I) = 0 \tag{11}$$

  This property does not hold in general for moments. A consequence of this property is that a higher-order cumulant of a stochastic vector with mutually independent components, is a diagonal tensor.

- *Non-Gaussianity:* if Y is a Gaussian variable with the same mean and variance as a given stochastic variable X, then holds for $N \geq 3$:

$$\mathcal{C}_X^{(N)} = \mathcal{M}_X^{(N)} - \mathcal{M}_Y^{(N)} \tag{12}$$

  Higher-order cumulants of a Gaussian variable are 0.

6.2 Independent Component Analysis

Assume the following basic linear statistical model:

$$Y = \mathbf{M}X + N \tag{13}$$

in which $Y \in \mathbb{R}^I$ is denoted as the *observation vector*, $X \in \mathbb{R}^J$ is called the *source vector* and $N \in \mathbb{R}^I$ represents additive *noise*. $\mathbf{M} \in \mathbb{R}^{I \times J}$ is the *mixing matrix*.

The concept of Independent Component Analysis can be formulated as follows:

The goal of Independent Component Analysis (ICA), or Blind Source Separation (BSS), consists of the estimation of the transfer matrix $\mathbf{M}$ *and/or the corresponding realizations of the source vector* X*, given only realizations of the output vector* Y*, under the following assumptions:*

- *The columns of* $\mathbf{M}$ *are linearly independent.*
- *The components of* X *are mutually statistically independent, as well as statistically independent from the noise components.*

The second assumption is the key ingredient for ICA. It is a very strong hypothesis, but also quite natural in lots of applications. From an algebraic point of view it implies that:

$$\mathbf{C}_Y = \mathbf{C}_X \times_1 \mathbf{M} \times_2 \mathbf{M} + \mathbf{C}_N \tag{14}$$
$$\mathcal{C}_Y^{(n)} = \mathcal{C}_X^{(n)} \times_1 \mathbf{M} \times_2 \mathbf{M} \ldots \times_n \mathbf{M} + \mathcal{C}_N^{(n)} \tag{15}$$

in which $\mathbf{C}_X$ and $\mathcal{C}_X^{(n)}$ are diagonal, and in which $\mathcal{C}_N^{(n)}$ vanishes if the noise is Gaussian.

The ICA-assumptions do not allow to distinguish between the signal and the noise term in Eq. (13). Hence the source signals will be estimated as $\hat{X}$, by a simple matrix multiplication:

$$\hat{X} = \mathbf{W}^H Y \tag{16}$$

Eq. (14) leads to a classical *Principal Component Analysis* (PCA), which only allows to estimate the sources as well as the mixing matrix up to an orthogonal transformation. Up to the noise term, Eq. (15) is a CANDECOMP of the cumulant of Y, which leads to higher-order-only solutions to the ICA-problem, conceptually blind for additive Gaussian noise [11]. Substitution of the PCA-subresults in Eq. (15) leads to an orthogonal transformation of $\mathcal{C}_Y^{(n)}$ into a diagonal tensor. Theoretically the unknown orthogonal factor can thus be estimated via an HOSVD (diagonality is a special case of all-orthogonality) [7].

However, when dealing with sample statistics, and taking the noise terms into account, the "maximal diagonality" approach is more robust [5].

It is clear that ICA is a very basic problem; its application range is very wide. The most well-known application is perhaps blind spatial multiplexing in telecommunications (e.g. [3], in which an efficient algorithm based on simultaneous matrix diagonalization is presented). Fig. 4 gives a biomedical example [9]. The left figure displays 8 channels of cutaneous potential signals of a pregnant woman. The large pulses correspond to the mother electrocardiogram (ECG); channels 1 to 3 clearly contain an additional weaker signal, which corresponds to the fetal ECG. The right figure displays the sources estimated with the ICA-algorithm of [5]. Channels 1 to 3, and to some extent also channel 4, contain the components of the mother ECG. Channels 6 and 8 show the fetal ECG.

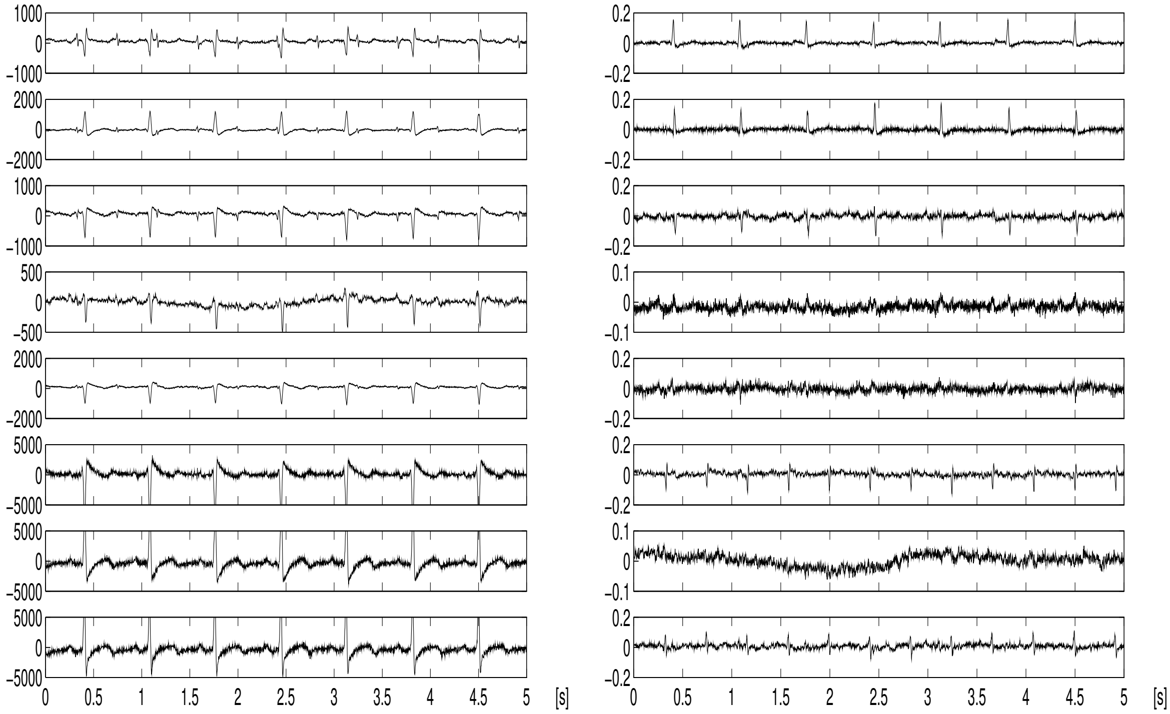

Figure 4: Left: 8-channel set of cutaneous potential recordings. Right: source estimates obtained via ICA.

7 Blind Factor Analysis

Blind Factor Analysis (BFA) is a fundamentally new signal processing tool, based on the CANDECOMP. Rather than discussing the concept from a theoretical point of view, we illustrate the technique by means of an application in Chemometrics [19, 20].

About 60 % of the equipment used for chemical analysis is bilinear (tandem massa spectrometry, 2D-NMR, . . .): the output of the device is a matrix, formed by rank-1 contributions of the various chemical components. E.g. for emission-excitation fluorescence data, each row is an emission spectrum and each column an excitation scan. Theoretically these row- and column-vectors are linear combinations of the emission and excitation spectra of the components, which have to be determined. At a matrix-level however, an infinite number of matrix decompositions is possible, and the chemical analysis cannot be realized without a priori knowledge. On the other hand, it is often necessary to examine several samples. If its identifiability conditions are satisfied, then the CANDECOMP of the corresponding third-order dataset allows to determine the components and their concentrations in a blind way.

8 Conclusion

In 1995, there was an LA-NET discussion on the question whether numerical linear algebra was dead (it wasn't). We hope to have demonstrated that numerical multinear algebra at least, is a new challenging and important field of research. In this summary paper we have briefly proposed three tensor generalizations of the SVD of matrices, each focusing on a different aspect of the matrix decomposition. Independent Component Analysis and Blind Factor Analysis are two examples of new signal processing methodologies based on tensor techniques.

References

[1] J.-F. Cardoso, "Iterative Techniques for Blind Source Separation using only Fourth-Order Cumulants", *Signal Processing VI: Theories and Applications, Proc. EUSIPCO-92*, Brussels, Aug. 24-27, 1992, Vol. 2, pp. 739-742.

[2] J.-F. Cardoso, "Super-Symmetric Decomposition of the Fourth-Order Cumulant Tensor. Blind Identification of More Sources than Sensors." *Proc. ICASSP-91*, Vol. 5, pp. 3109-3112.

[3] J.-F. Cardoso, A. Souloumiac, "Blind beamforming for non-Gaussian signals", *IEE Proc.-F*, Vol. 140, No. 6, 1994, pp. 362-370.

[4] J.D. Carroll, S. Pruzansky, "The CANDECOMP-CANDELINC Family of Models and Methods for Multidimensional Data Analysis", in: H.G. Law, C.W. Snyder, J.A. Hattie, R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis*, Praeger, N.Y., 1984, pp. 372-402.

[5] P. Comon, "Independent Component Analysis, A New Concept?" *Signal Processing*, Special Issue *Higher Order Statistics*, Vol. 36, No. 3, April 1994, pp. 287-314.

[6] P. Comon, B. Mourrain, "Decomposition of Quantics in Sums of Powers of Linear Forms", *Signal Processing*, Special Issue *Higher Order Statistics*, Vol. 53, Nos. 2-3, Sept. 1996, pp. 93-108.

[7] L. De Lathauwer, B. De Moor, J. Vandewalle, "Blind Source Separation by Higher-Order Singular Value Decomposition", *Proc. EUSIPCO-94*, Edinburgh, Scotland, U.K., Sept. 13-16, 1994, Vol. 1, pp. 175-178.

[8] L. De Lathauwer, B. De Moor, J. Vandewalle, "A Singular Value Decomposition for Higher-Order Tensors", ESAT/SISTA report 94-31, K.U.Leuven, submitted to: *SIAM Journal on Matrix Anal. and Appl.*.

[9] L. De Lathauwer, B. De Moor, J. Vandewalle, "Fetal Electrocardiogram Extraction by Source Subspace Separation", *Proc. IEEE SP / ATHOS Workshop on HOS*, Girona, Spain, June 12-14, 1995, pp. 134-138.

[10] L. De Lathauwer, P. Comon, B. De Moor, J. Vandewalle, "Higher-Order Power Method - Application in Independent Component Analysis", *Proc. 1995 Int. Symp. on Nonlinear Theory and its Applications (NOLTA'95)*, Las Vegas, USA, Dec. 10-14, 1995, Vol. 1, pp. 91-96.

[11] L. De Lathauwer, B. De Moor and J. Vandewalle, "Independent Component Analysis based on Higher-Order Statistics Only", *Proc. 8th IEEE SP Workshop on Stat. Signal and Array Process. (SSAP-96)*, Corfu, Greece, June 24-26, 1996, pp. 356-359.

[12] L. De Lathauwer, B. De Moor and J. Vandewalle, "Blind Source Separation by Simultaneous Third-Order Tensor Diagonalization", *Signal Processing VIII: Theories and Applications, Proc. EUSIPCO-96*, Trieste, Italy, Sept. 10-13, 1996, Vol. 3, pp. 2089-2092.

[13] L. De Lathauwer, *Signal Processing by Higher-Order Tensors*, Ph.D. thesis, K.U.Leuven, E.E.Dept.-ESAT, Belgium (in preparation).

[14] R.A. Harshman, M.E. Lundy, "The PARAFAC Model for Three-Way Factor Analysis and Multidimensional Scaling," in: H.G. Law, C.W. Snyder, J.A. Hattie, R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis*, Praeger, N.Y., 1984, pp. 122-215.

[15] D.C. Kay, *Theory and Problems of Tensor Calculus*, McGraw-Hill, 1988.

[16] S.E. Leurgans, R.T. Ross, R.B. Abel, "A Decomposition for Three-Way Arrays", *SIAM Journal on Matrix Anal. and Appl.*, Vol. 14, No. 4, Oct. 1993, pp. 1064-1083.

[17] P. McCullagh, *Tensor Methods in Statistics*, Chapman and Hall, Londen, 1987.

[18] C.L. Nikias, J.M. Mendel, "Signal Processing with Higher-Order Spectra", *IEEE Signal Proc. Mag.*, July 1993, pp. 10-37.

[19] E. Sanchez, B.R. Kowalski, "Tensorial Calibration: II. Second-Order Calibration", *Journal of Chemometrics*, Vol. 2, pp. 265-280, 1988.

[20] E. Sanchez, B.R. Kowalski, "Tensorial Resolution: a Direct Trilinear Decomposition", *Journal of Chemometrics*, Vol. 4, pp. 29-45, 1990.