LABORATOIRE

INFORMATIQUE, SIGNAUX ET SYSTÈMES
DE SOPHIA ANTIPOLIS
UMR 6070

JOINT APPROXIMATE DIAGONALIZATIONS AND APPLICATION TO VIRTUAL ARRAYS

Laurent Albera, Pierre Comon

Projet ASTRE

Rapport de recherche
ISRN I3S/RR–2003-27–FR

Août 2003

LABORATOIRE I3S: Les Algorithmes / Euclide B – 2000 route des Lucioles – B.P. 121 –
06903 Sophia-Antipolis Cedex, France – Tél. (33) 492 942 701 – Télécopie : (33) 492 942 898
http://www.i3s.unice.fr/I3S/FR/

RÉSUMÉ :

Le probleme de l'identification aveugle de melanges de sources independantes est souvent relie a la diagonalisation de certains tenseurs. Ce probleme est pose ici en termes de diagonalisation non conventionnelle simultanee approximative de plusieurs matrices. En effet, une transformation congruente identique est appliquee a chacune des matrices; celle de gauche etant rectangulaire de rang plein, et celle de droite etant unitaire. Nous decrivons une application en traitement d'antenne, et proposons quelques algorithmes sous-optimaux.

MOTS CLÉS :

Tenseurs, Separation Aveugle de Sources, Systemes sous-determines

ABSTRACT:

The problem of Blind Identification of linear mixtures of independent random processes is often related to the diagonalization of some tensors. This problem is posed here in terms of a non conventional joint approximate diagonalization of several matrices. In fact, a single congruent transform is applied to each of these matrix; the left transform being rectangular full rank, and the right one being unitary. The application in antenna signal processing is described, and suboptimal numerical algorithms are proposed.

KEY WORDS :

Tensors, Blind Source Separation, Under-determined Systems

Joint Approximate Diagonalizations and Application to Virtual Arrays

L. Albera and P. Comon*

Lab. I3S, 2000 route des Lucioles, B.P.121, F-06903 Sophia-Antipolis cedex

I3S Research Report
ISRN I3S/RR-2003-27-FR

August 19, 2003

Abstract

The problem of Blind Identification of linear mixtures of independent random processes is often related to the diagonalization of some tensors. This problem is posed here in terms of a non conventional joint approximate diagonalization of several matrices. In fact, a single congruent transform is applied to each of these matrices, the left transform being rectangular full rank, and the right one being unitary. The application in antenna signal processing is described, and suboptimal numerical algorithms are proposed.

KEY WORDS : Tensors, Blind Source Separation, Under-determined systems

*this paper has been submitted to *Linear Algebra Appl. on May 5, 2003, with two other co-authors of Thales Communications: Pascal Chevalier and Anne Ferreol*

1 Introduction

xx def of underdetermined mixtures

In order to face the underdetermined mixtures case, several methods have been developed. Some papers focus on blind source extraction [7] [3], which is a difficult problem since underdetermined mixtures are not linearly invertible, while others, as herein, favour Blind Source Identification (BSI) of the mixture matrix [1] [2] [4] [6] [5] [7] [8]. The methods proposed in [1] [2] [4] only exploit the information contained in the data FO statistics whereas the one proposed in [8] exploits the information contained in the second characteristic function of the observations. As for Lee et al. [7], they maximize the probability of the data conditionally to the mixture matrix. However, all these methods have drawbacks in operational contexts. Indeed, the method [1] is still very difficult to implement and does not ensure the BSI of the source steering vectors when the sources have the same kurtosis. The BSI methods [2] [4] assume FO non-circularity and thus fail in separating FO circular sources. Besides, the theory developed in [2] only confines itself to the three sources and two sensors case. Although the method [7] succeeds in identifying the steering vectors of up to four speech signals with only two sensors, the authors need sparsity conditions, and do not address the general case when all sources are always present. Eventually, the method [8] has been developed only for real mixtures of real-valued sources, and the issue of robustness with respect to an over estimation of the number of sources remains open.

From the linear algebra viewpoint, it is shown in section 2 that the BSI problem can be expressed in the form of the problem below, even in the underdetermined case.

Problem 1 *Given N matrices, $\mathbf{\Gamma}(n)$, $1 \le n \le N$, each of size $M \times P$, $M \ge P$, find a full rank $M \times P$ matrix $\boldsymbol{\mathcal{A}}$, N diagonal matrices $\mathbf{\Lambda}(n)$, and a unitary $P \times P$ matrix, $\boldsymbol{V}$, such that*

$$\mathbf{\Gamma}(n) \approx \boldsymbol{\mathcal{A}}\,\mathbf{\Lambda}(n)\,\boldsymbol{V}^{\mathsf{H}}$$

Note that this problem differs from the GSVD setting in two respects. First, the unitary matrix $\boldsymbol{V}$ must be the same for every $\mathbf{\Gamma}(n)$; second, there may be more than two matrices ($N \ge 2$). As a counterpart, the diagonalizations cannot be exact, in general. For this reason, an appropriate optimization will be defined.

Throughout the paper, vectors (one-way arrays) are denoted with bold lowercase symbols, and matrices (2-way arrays) or tensors (higher order arrays) in bold uppercase. Transposition, conjugate transposition, and complex conjugation are denoted respectively with superscripts $(^{\mathsf{T}})$, $(^{\mathsf{H}})$, and $(^{*})$.

Matrix notation

First, define the following compact notation associated with the usual Kronecker product:

$$\boldsymbol{B}^{\otimes q} = \underbrace{\boldsymbol{B} \otimes \boldsymbol{B} \otimes \ldots \otimes \boldsymbol{B}}_{q \text{ times}} \qquad \text{with } \boldsymbol{B}^{\otimes 0} = 1 \tag{1}$$

where $\boldsymbol{B}$ is any $N \times P$ rectangular matrix; $\boldsymbol{B}^{\otimes q}$ is then $N^q \times P^q$. Next, define a columnwise Kronecker product, denoted $\oslash$. For any matrix $\boldsymbol{B}$, the columns of matrix $\boldsymbol{B}^{\oslash q}$ are defined as $\boldsymbol{b}_j^{\otimes q}$, if $\boldsymbol{b}_j$ denote the columns of $\boldsymbol{B}$. As a consequence, $\boldsymbol{B}^{\oslash q}$ is of size $N^q \times P$.

2 The BIOME method

The actual problem we face (cf. section 4) is a structured version of problem 1, and better described by

Problem 2 *Given N matrices, $\mathbf{\Gamma}(n)$, $1 \le n \le N$, each of size $N^q \times P$, $N^q \ge P$ but possibly $N < P$, find a full rank $N \times P$ matrix $\boldsymbol{A}$, a $P \times P$ diagonal matrix $\mathbf{\Lambda}$, and a unitary $P \times P$ matrix, $\boldsymbol{V}$, such that*

$$\mathbf{\Gamma}(n) \approx \boldsymbol{\mathcal{A}}\,\mathbf{\Lambda}(n)\,\boldsymbol{V}^{\mathsf{H}}$$

where $\mathbf{\Lambda}(n) = \mathrm{Diag}[A_{n1}, \ldots, A_{nP}]\,\mathbf{\Lambda}$, and $\boldsymbol{\mathcal{A}} = \boldsymbol{A}^{\oslash q-\ell} \oslash \boldsymbol{A}^{*^{\oslash \ell}}$.

By forming products of the form $\mathbf{\Gamma}(m)^{\#}\mathbf{\Gamma}(n)$, where $\mathbf{\Gamma}(m)^{\#}$ denotes the pseudo inverse of $\mathbf{\Gamma}(m)$, one can estimate $\boldsymbol{V}$ by computing the joint approximate Eigen Value Decomposition (EVD) of matrices:

$$\mathbf{\Gamma}(m)^{\#}\mathbf{\Gamma}(n) \approx \boldsymbol{V}\mathbf{\Lambda}(m,n)\boldsymbol{V}^{\mathsf{H}}$$

xx detail how...

Once $\boldsymbol{V}$ has been computed, Problem 2 reduces to find matrices $\boldsymbol{A}$ and $\mathbf{\Lambda}(n)$ such that $\mathbf{\Gamma}(n)\boldsymbol{V} \approx \boldsymbol{A}\mathbf{\Lambda}(n)$, $\forall n$, which is equivalent to problem 3 below, if $\mathbf{\Lambda}(n)$ are invertible.

Problem 3 *Given N matrices, $\mathbf{\Gamma}(n)$, $1 \le n \le N$, each of size $M \times P$, find a full rank $M \times P$ matrix $\boldsymbol{A}$, and $P \times P$ diagonal matrices $\boldsymbol{D}_n$ such that*

$$\mathbf{\Gamma}(n)\boldsymbol{D}(n) \approx \boldsymbol{\mathcal{A}}$$

Now, matrix $\boldsymbol{\mathcal{A}}$ is structured. Once it has been estimated, one can reuse Problem 3 another time to estimate $\boldsymbol{A}$ from $\boldsymbol{\mathcal{A}}$ (actually, one decreases q by one every time). xx explain better with the example of $q = 3$. xx

3 Solution to Problem 3

Matrices $\boldsymbol{A}$ and $\boldsymbol{D}(n)$ are obtained as stationary values of the Least Squares (LS) criterion below:

$$\varepsilon = \sum_{m=1}^{N} ||\mathbf{\Gamma}(m)\boldsymbol{D}(m) - \boldsymbol{A}||^2 \tag{2}$$

As a consequence, they satisfy the following system of equations:

$$\sum_i [\mathbf{\Gamma}(m)\boldsymbol{D}(m) - \boldsymbol{A}]_{ij}\Gamma^*_{ji}(m) = 0, \forall j$$
$$\sum_m [\mathbf{\Gamma}(m)\boldsymbol{D}(m) - \boldsymbol{A}]_{kj} = 0, \forall (k,j)$$

It is then not hard to obtain the closed form expression for $\boldsymbol{A}$:

$$\boldsymbol{A} = \frac{1}{M}\sum_m \mathbf{\Gamma}(m)\boldsymbol{D}(m) \tag{3}$$

By plugging back this solution in the system, one gets after some manipulations:

$$\boldsymbol{G}(j)\begin{pmatrix} D_{jj}(1) \\ D_{jj}(2) \\ \vdots \\ D_{jj}(M) \end{pmatrix} = 0,\ \forall j, \tag{4}$$

where

$$\boldsymbol{G}(j) \stackrel{\text{def}}{=} \begin{bmatrix} (M-1)[\mathbf{\Gamma}^{\mathsf{H}}(1)\mathbf{\Gamma}(1)]_{jj} & [\mathbf{\Gamma}^{\mathsf{H}}(1)\mathbf{\Gamma}(2)]_{jj} & \cdots & [\mathbf{\Gamma}^{\mathsf{H}}(1)\mathbf{\Gamma}(M)]_{jj} \\ [\mathbf{\Gamma}^{\mathsf{H}}(2)\mathbf{\Gamma}(1)]_{jj} & (M-1)[\mathbf{\Gamma}^{\mathsf{H}}(2)\mathbf{\Gamma}(2)]_{jj} & \cdots & [\mathbf{\Gamma}^{\mathsf{H}}(2)\mathbf{\Gamma}(M)]_{jj} \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{\Gamma}^{\mathsf{H}}(M)\mathbf{\Gamma}(1)]_{jj} & [\mathbf{\Gamma}^{\mathsf{H}}(M)\mathbf{\Gamma}(2)]_{jj} & \cdots & (M-1)[\mathbf{\Gamma}^{\mathsf{H}}(M)\mathbf{\Gamma}(M)]_{jj} \end{bmatrix}.$$

In other words, the solution to the LS problem under the constraint that $\sum_m |D_{jj}(m)|^2 = 1$ is obtained when the vector $[D_{jj}(1),\ D_{jj}(2),\ \ldots D_{jj}(M)]^{\mathsf{T}}$ is the right singular vector of matrix $\boldsymbol{G}(j)$ associated with the minimal singular value. Once every entry $D_{jj}(m)$ is obtained, matrix $\boldsymbol{A}$ can be calculated thanks to (3).

4 Application to Blind identification of Under-determined Mixtures

xx

References

[1] J.-F. CARDOSO, "Super-symetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors," in *ICASSP 91*, Toronto, Canada, May 1991, pp. 3109–3112.

[2] P. COMON, "Blind channel identification and extraction of more sources than sensors," in *SPIE Conference*, San Diego, US, July 19-24 1998, pp. 2–13.

[3] P. COMON and O. GRELLIER, "Non-linear inversion of underdetermined mixtures," in *ICA 99, IEEE Workshop on Indep. Comp. Anal. and Signal Separation*, Aussois, France, January 11-15 1999, pp. 461–465.

[4] L. de LATHAUWER, P. COMON, and B. de MOOR, "ICA algorithms for 3 sources and 2 sensors," in *Sixth Sig. Proc. Workshop on Higher Order Statistics*, Caesarea, Israel, June 14-16 1999, pp. 116–120.

[5] L. de LATHAUWER and B. de MOOR, "From matrix to tensor : Multilinear algebra and signal processing," in *Mathematics in Signal Processing IV*. Oxford, UK: Oxford University Press, 1998, pp. 1–15.

[6] L. de LATHAUWER, B. de MOOR, and J. VANDEWALLE, "On the best rank-1 and rank-(R1,R2,...RN) approximation of high-order tensors," *SIAM Jour. Matrix Ana.*, vol. 21, no. 4, pp. 1324–1342, 2000.

[7] T. W. LEE, M. S. LEWICKI, M. GIROLAMI, and T. J. SEJNOWSKI, "Blind source separation of more sources than mixtures using overcomplete representations," *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 87–90, 1999.

[8] A. TALEB, "An algorithm for the blind identification of N independent signal with 2 sensors," in *ISSPA 01, sixteenth symposium on signal processing and its applications*, Kuala-Lumpur, Malaysia, August 13-16 2001.