Signal Processing 145 (2018) 78–90

Contents lists available at ScienceDirect

Signal Processing

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

A new class of block coordinate algorithms for the joint eigenvalue decomposition of complex matrices

Rémi André[a,b,1,*], Xavier Luciani[a,b,1], Eric Moreau[a,b,1]

[a] Laboratoire des Sciences de l'Information et des Systèmes, UMR CNRS 7296, France

[b] Aix Marseille Université, CNRS, ENSAM, 13397 Marseille et Université de Toulon, La Garde, 83957 France

ARTICLE INFO

Article history:
Received 22 June 2017
Revised 20 November 2017
Accepted 22 November 2017

Keywords:
Joint eigenvalue decomposition
Block coordinate procedure
Canonical polyadic decomposition
Source separation
Joint diagonalization
DS-CDMA

ABSTRACT

Several signal processing problems can be written as the joint eigenvalue decomposition (JEVD) of a set of noisy matrices. JEVD notably occurs in source separation problems and for the canonical polyadic decomposition of tensors. Most of the existing JEVD algorithms are based on a block coordinate procedure and require significant modifications to deal with complex-valued matrices. These modifications decrease algorithms performances either in terms of estimation accuracy of the eigenvectors or in terms of computational cost. Therefore, we propose a class of algorithms working equally with real- or complex-valued matrices. These algorithms are still based on a block coordinate procedure and multiplicative updates. The originality of the proposed approach lies in the structure of the updating matrix and in the way the optimization problem is solved in $\mathbb{C}^{N\times N}$. That structure is parametrized and allows to define up to five different JEVD algorithms. Thanks to numerical simulations, we show that, with respect to the more accurate algorithms of the literature, this approach improves the estimation of the eigenvectors and has a computational cost significantly lower. Finally, as an application example, one of the proposed algorithm is successfully applied to the blind source separation of Direct-Sequence Code Division Multiple Access signals.

1. Introduction

Joint EigenValue Decomposition (JEVD), also called joint diagonalization by similarity, is an important issue for a number of signal processing applications such as directions of arrival estimation [1], joint angle-delay estimation [2], multi-dimensional harmonic retrieval [3], Independent Component Analysis (ICA) [4–8], Canonical Polyadic Decomposition (CPD) of tensors [9–11] and statistics [12].

JEVD problems occur when a set of K non-defective matrices $\mathbf{M}^{(k)}$ shares the same basis of eigenvectors:

$$\boldsymbol{M}^{(k)} = \boldsymbol{A}\boldsymbol{D}^{(k)}\boldsymbol{A}^{-1}, \quad \forall k = 1, \ldots, K \tag{1}$$

where the invertible matrix $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ is the common matrix of eigenvectors and the K matrices $\boldsymbol{D}^{(k)} \in \mathbb{C}^{N\times N}$ are all diagonal and contain the eigenvalues of corresponding $\mathbf{M}^{(k)}$ matrices. The goal is then to estimate $\boldsymbol{A}$ or $\boldsymbol{A}^{-1}$ from matrices $\mathbf{M}^{(k)}$.

Naive approaches consist in considering the K eigenvalue decompositions separately or a linear combination of those. The main problem is then that if some eigenvalues are degenerated or very close, the corresponding eigenvectors cannot be correctly identified. Moreover, in practice this problem is accentuated by the presence of noise because small perturbations of a matrix can strongly affect its eigenvectors [9,13]. As a consequence, a recommended solution is to decompose the whole matrix set jointly. One usual way of doing is to make matrices $\mathbf{M}^{(k)}$ as diagonal as possible within the same change of basis. That's why JEVD can be seen as a joint diagonalization problem. More precisely we speak of joint diagonalization by similarity in contrast with the Joint Diagonalization by Congruence (JDC) problem for which the inverse of the matrix $\boldsymbol{A}$ in (1) is replaced by the conjugate transpose of $\boldsymbol{A}$. Of course, JDC and JEVD are equivalent if $\boldsymbol{A}$ is a unitary matrix but this is not necessary the case here. Tensor CPD is a common application to JDC and JEVD [14,15]. However JEVD and JDC are two different mathematical problems that appear in distinct applications. Moreover the JEVD can be seen has a more general problem since the JDC can be very easily rewritten into a JEVD [16] while it seems that the opposite is not true. Hence JEVD deserves a particular attention and specific algorithms although these are often inspired by JDC algorithms.

* Corresponding author.
E-mail addresses: randre@univ-tln.fr (R. André), luciani@univ-tln.fr (X. Luciani), moreau@univ-tln.fr (E. Moreau).

[1] EURASIP member.

https://doi.org/10.1016/j.sigpro.2017.11.016

Indeed most of them resort to an iterative block coordinate procedure adapted from the original Jacobi method [17]. This means that matrix $\boldsymbol{A}$ (or $\boldsymbol{A}^{-1}$) is built by successive multiplicative updates. Each update involves a small set of parameters (with respect to N^2) that allows to build the updating matrix. JEVD algorithms differ one from another in the way these parameters are defined and computed. Several families of algorithms can then be identified. Algorithms of the first family look for the updating matrix in the form of a QR factorization [3,18]. Actually these methods only aim to estimate the eigenvalues by triangulazing the matrix set, thereby these are not considered here. More recently, three algorithms based on the polar decomposition were proposed: SHear RoTation algorithm (SH-RT) [19], Joint Unitary Shear Transformation (JUST) [20] and Joint Diagonalization algorithm based on Targeting hyperbolic Matrices (JDTM) [11]. When dealing with real-valued matrix sets, these algorithms build the updating matrix as a product of an orthogonal rotation matrix and a symmetric matrix. These two matrices are estimated separately one after an other. Following the idea of Souloumiac in [21], each matrix is parametrized by a rotation or an hyperbolic rotation angle. Thanks to simple trigonometric and hyperbolic trigonometric properties the optimization problems can then be rewritten as a simple EVD of a 2×2 matrix. The extension of these methods to the complex field involve 2 reals angles and 2 real hyperbolic angles to define the unitary and the hermitian matrix. The estimation of the hyperbolic angles is not an easy task. In JUST, a suboptimal scheme is used since the angles are estimated separately. In addition the first one is approximated. In JDTM both angles are estimated conjointly in some optimal way but the procedure requires to find the zeros of an eleventh degree polynomial. Finally a third family of algorithms based on the LU factorization and called Joint Eigenvalue decomposition algorithms based on Triangular matrices (JET) was introduced in [8]. Here the updating matrix is an elementary lower triangular matrix. For complex valued matrix sets, real and imaginary parts of the updating matrix are not estimated conjointly but one after the other. Hence, at each (real or imaginary) update, the set of available solutions is limited to the set of real matrices or to the set of imaginary matrices alternatively.

As it has just been seen the extension of the previous methods to the complex case implies to map $\mathbb{C}$ to $\mathbb{R}^2$ and involves some suboptimal optimization schemes. Few algorithm comparisons have been made in the complex case so far. More precisely, it has been shown in [8] that in this case JDTM requires many iterations to converge hence a very high computational cost with respect to JET algorithms. In addition, JET performs better than JDTM when the noise level is low. However JET performances decrease significantly with the noise level. As a consequence we propose here a new class of algorithms that work equally for real- or complex-valued matrices.

The proposed approach has common points with the previous methods: it is also based on a block coordinate procedure and the updating matrix is still computed as a product of factorization matrices. However its originality is twofold. First, these matrices are estimated conjointly from a simple eigenvalue decomposition of a 2×2 matrix. This strategy is inspired from [22] for the JDC problem. Second, we propose a parametrized expression of the updating matrix that cover different matrix factorizations. We then show that we can switch from one matrix factorization to another by changing only one parameter. This allows to define a class of five algorithms sharing the same structure. One of these algorithms was briefly presented in [23]. Moreover, we have the possibility to pass from one version to another at each new iteration of the optimization process hence defining a sixth algorithm. By estimating conjointly all the parameters of the updating matrix, we expect a better trade-off between convergence speed and robustness to the noise level than the above mentioned approaches. In addition, our approach allows to reduce the numerical complexity of the block coordinate step with respect to classical polar decomposition based algorithms. Eventually, this approach has the advantage to work in $\mathbb{C}^{N \times N}$ throughout the process. In other words, no modifications are required to deal with real- or complex-valued matrices.

The paper is organized as follow: in Section 2, we recall the principle of block coordinate JEVD algorithms. In Section 3, we describe the proposed method and the algorithms. In Section 4, we compare the numerical complexity of these algorithms to the ones of the existing algorithms. Section 5 is dedicated to numerical simulations. We have evaluated the performances of the proposed algorithms to compute the JEVD of complex matrices according to several scenarios. Comparisons are made with all the other JEVD algorithms. Finally in Section 6, we show how the proposed approach can be used to achieve the blind source separation of telecommunication signals.

Notations. In the following scalars are denoted by a lower case (a), vectors by a boldface lower case ($\mathbf{a}$) and matrices by a boldface upper case ($\boldsymbol{A}$). a_i is the i-th element of vector $\mathbf{a}$ and $A_{i,j}$ is the (i, j)-th element of matrix $\boldsymbol{A}$. Operator $||\bullet||$ is the Frobenius norm of the argument matrix. Real and complex fields are denoted by $\mathbb{R}$ and $\mathbb{C}$ respectively. Operator $\text{ZDiag}\{\bullet\}$ sets to zero the diagonal of the argument matrix. $\boldsymbol{I}$ is the identity matrix. Modulus and conjugate of any complex number z are denoted by $|z|$ and $\bar{z}$ respectively. $k \in [1, K]_{\mathbb{N}}$ is the sequence of natural integers from 1 to K.

2. A block coordinate procedure

The JEVD problem consists in finding a matrix, $\boldsymbol{B}$, which jointly diagonalizes the given set of matrices $\mathbf{M}^{(k)}$ in equation (1). $\boldsymbol{B}$ is called the diagonalizing matrix and can be considered as an estimate of $\boldsymbol{A}^{-1}$ up to a permutation and a scaling indeterminacy. This indeterminacy is inherent to the JEVD problem. An important uniqueness result has been shown in [9]. Let us define matrix $\boldsymbol{\Omega}$ as

$$\boldsymbol{\Omega} = \begin{pmatrix} D_{11}^{(1)} & \cdots & D_{11}^{(K)} \\ \vdots & \cdots & \vdots \\ D_{NN}^{(1)} & \cdots & D_{NN}^{(K)} \end{pmatrix}. \tag{2}$$

The JEVD is unique up to a permutation and a scaling indeterminacy if and only if the rows of $\boldsymbol{\Omega}$ are two by two distinct (*i.e.* $\forall m$, n with $m \neq n$, $\boldsymbol{\omega}_{m.} - \boldsymbol{\omega}_{n.} \neq \mathbf{0}$ where $\boldsymbol{\omega}_{m.}$ and $\boldsymbol{\omega}_{n.}$ are the m^{th} and the n^{th} rows of $\boldsymbol{\Omega}$ respectively). In the following, we will always assume that this uniqueness condition is satisfied.

We want to build $\boldsymbol{B}$ such that matrices $\widehat{\boldsymbol{D}}^{(k)}$ defined by:

$$\widehat{\boldsymbol{D}}^{(k)} = \boldsymbol{B}\boldsymbol{M}^{(k)}\boldsymbol{B}^{-1}, \quad \forall k \in [1, K]_{\mathbb{N}} \tag{3}$$

are as diagonal as possible. The method is iterative. At each iteration, $\boldsymbol{B}$ is multiplicatively updated with a matrix $\boldsymbol{X}$ in the following way:

$$\boldsymbol{B} \leftarrow \boldsymbol{X}\boldsymbol{B} \tag{4}$$

and the set of matrices $\widehat{\boldsymbol{D}}^{(k)}$ is consequently updated as:

$$\widehat{\boldsymbol{D}}^{(k)} \leftarrow \boldsymbol{X}\widehat{\boldsymbol{D}}^{(k)}\boldsymbol{X}^{-1}, \quad \forall k \in [1, K]_{\mathbb{N}}. \tag{5}$$

We expect that at the end of the iterative process, matrices $\widehat{\boldsymbol{D}}^{(k)}$ are diagonalized and $\boldsymbol{B}\boldsymbol{A}$ is close to the product between an invertible diagonal matrix and a permutation matrix (This closeness can be measured thanks to the criterion proposed in [6]). $\boldsymbol{B}$ is initialized from an appropriate initial guess $\boldsymbol{B}_0$ and before the first iteration we set $\widehat{\boldsymbol{D}}^{(k)} \leftarrow \boldsymbol{B}_0\boldsymbol{M}^{(k)}\boldsymbol{B}_0{}^{-1}, \quad \forall k \in [1, K]_{\mathbb{N}}$. The choice of $\boldsymbol{B}_0$ will be discussed in Section 5.

At each iteration, $\boldsymbol{X}$ is computed thanks to a block coordinate procedure. Here, it means that $\boldsymbol{X}$ is built from a set of $N(N-1)/2$