Optik 140 (2017) 145–150

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Non-orthogonal joint diagonalization algorithm preventable ill conditioned solutions for blind source separation

Tiaojun Zeng, Li Gou*, Junhang Wu

School of Information Science & Technology, Shihezi University, Shihezi 610031, Xinjiang Autonomous Region, China

ARTICLE INFO

Article history:
Received 5 January 2017
Accepted 4 April 2017

Keywords:
Blind signal separation
Dual-matrix method
Constrained terms
Joint diagonalization

ABSTRACT

The general joint diagonalization problem involves estimating the separating matrix only given mixing signals interrelated matrix set. For nonorthogonal joint diagonalization based on the weighted least-squares criterion, the algorithms may converge to trivial (zero) solution. Certainly, the trivial solution can simply be avoided by adopting some constraint on the diagonalizing matrix or penalty terms. However, free of zero solution is not enough, especially for the blind signal separation (BSS). Actually, ill-conditioned diagonalizer even though nonzero makes the objective function unstable or even divergence in the process optimization. Therefore, it is necessary to prevent the iterative solutions from degenerating ill-conditioned forms. To solve this problem, a novel nonleast-squares criterion for non-orthogonal joint diagonalization is proposed. It is imposed constrainted terms on diagonalizers, which are induced form the mathematic define of the ill condition matrix. Finally, Computer simulations indicate that the new algorithm yields diagonalizers which not only minimize the diagonalization error but also have as small condition numbers as possible, meanwhile, degenerate solutions are avoided strictly.

1. Introduction

Aiming to recover the sources from the observations without priori knowledge on the mixing procedure and the sources, blind source separation (BSS) is a powerful signal processing method. In recent years, many scholars have made deep research on this problem, and have made many achievements. In reference [1], Xu proposed a multi-stage algorithm (MSA) which is different with the classic second-order blind identification method using Givens rotations to derive the unitary matrix, the MSA seeks one column of the unitary through solving a symmetrical tri-quadric cost function. In ref. [2], Li proposed proposes some improved legible cost functions based on linear prediction for BSS. Their optimization is simpler and can be converted into a generalized eigendecomposition of a corresponding matrix pencil, etc. Owing to its broad application potential, BSS has been extensively employed in fields such as biomedicine, image processing, and speech signal processing [3–5], etc.

The issue of approximate joint diagonalization (JD) of a set of target matrices has been researched extensively in the past few decades and has used to implement its wide-ranging applications in a variety of signal processing fields, such as blind source separation (BSS) [6], blind identification [7], and blind wave beamforming [8].

* Corresponding author.
E-mail addresses: zengtiaojun@163.com (T. Zeng), guolishzu@163.com (L. Gou), 965743423@qq.com (J. Wu).

http://dx.doi.org/10.1016/j.ijleo.2017.04.011

Joint diagonalization algorithms play a crucial role in the success of this type of separation methods. By far, a variety of algorithms have been proposed for the joint diagonalization problem. Several algorithms based on joint diagonalization (JD) have been proposed to solve the BSS problem [9–16]. They are generally divided into two categories: the Orthogonal Joint Diagonalization (OJD) problem [9–11] and the Non-Orthogonal Joint Diagonalization (NOJD) problem depending on whether requiring that the solution matrix is an orthogonal matrix or not [12–16]. A number of state-of-the-art algorithms were proposed to solve this problem. OJD algorithms restrict the diagonalizer to be orthogonal, and are applicable in BSS when the observations are prewhitened. However, because of some disadvantages in prewhitening phase in BSS, NJD has received increasing attention in recent years. Among the existing JD algorithms, most are based on the following criterion of minimizing diagonalization error:

$$F(\boldsymbol{A}) = \sum_{i} \sum_{m \neq n} ((\boldsymbol{A}\boldsymbol{M}_i\boldsymbol{A}^T)_{mn})^2. \tag{1}$$

where matrix $A \in \mathbb{R}^{N \times M}$ (in this paper, we discuss the problem in real number domain) is called a diagonalizer conveniently and $\boldsymbol{M}_i = R_{\boldsymbol{x}}(\tau_i)(i = 1, 2, \ldots, N)$ is the correlation matrix of observed signal vector $\boldsymbol{x}(t)$ with time lag τ_i. It is clear that the solution under this model sometimes tends to degenerate zero or near zero because of the mathematical properties. Of course, these solutions can be excluded to the optimization process by adding constraint terms in the criterion. Many articles have discussed this issue. But there is a problem rarely discussed: how to avoid the the diagonalizer tends to the ill conditioned form. When this problem occurs, it will undermine the stability of the algorithm that lead the algorithm not to convergence, although it is implied, does not necessarily occur. Zhou proposed a NJD algorithm by imposing variety of constraints on diagonalizers, which yields small condition number diagonalizers [15]. But the algorithm has two shortcomings: (1) the imposed constrained terms indirectly limit the upper bound of diagonalizers, but they do not clear their own ranges. That may lead the conditioned numbers of diagonalizers to converge large values even ill-conditioned solutions. (2) In order to solve NJD problem, they transform it into a constrained optimization problem. Compared to the unconstrained optimization problem, the constraint form is more difficult to solve. Moreover, the adopted off-diag cost function is the special case of the general form, it is proven that the solution under the off model rarely converge to the separable matrix even though its cost function converges to zero [14]. In this paper, a new method is proposed to solve these problems. The expression of the constrain term imposed on the cost function is derived from the mathematical definition of the ill conditioned matrix. Under the action of the constraint term, the condition number of the diagonalizer is limited to the range of the specified value. In other words, the condition number can be adaptively controlled, that will accelerate the convergence rate of iterative algorithm. Finally, the NJD problem is transformed into a unconstrained optimization which is easier programming.

The rest of the paper is organized as follows. The issue of the condition number of a diagonalizer is raised in Section 2. Section 3 shows the experimental results and related discussions. Conclusions are drawn in Section 4. Notation: boldfaced capital and lowercase letters denote matrices and the column vectors, respectively. $(\cdot)^H$ is the Hermitian transpose, $(\cdot)^T$ is the transpose, $(\cdot)^*$ is the complex conjugate. $\|\cdot\|$ represents the Frobenius norm and $E(\cdot)$ denotes the expectation operator.

2. Proposed approach

In the BSS context, denote the observed signals $\boldsymbol{x}(t) = \{x_i(t), \boldsymbol{i} = 1, 2, \ldots, M\}^T$ by matrix equation $\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{s}(t) + \boldsymbol{n}(t)$, where $\boldsymbol{W} = (w_{ij})_{M \times N}$ is the mixing matrix, $\boldsymbol{W} = (w_{ij})_{M \times N}(M \geq N)$ is the source signal vector, and $\boldsymbol{n}(t)$ is the additive noise vector. Under some assumptions for solving the BSS problem following: (1) The unknown mixing matrix $\boldsymbol{W}$ is of full-column rank. (2) The source signals are zero-mean, statistically mutually independent stationary signals. (3) The additive noise, which could be colored in the space domain and white in the time domain, is independent from the source signals.

The covariance matrices of the array output form the following structure [9]:

$$R_{\boldsymbol{x}}(0) = E\{\boldsymbol{x}(t)\boldsymbol{x}^T(t)\} = \boldsymbol{W}\boldsymbol{D}_0\boldsymbol{W}^T + R_{\boldsymbol{n}} \tag{2}$$

$$R_{\boldsymbol{x}}(\tau_i) = E\{\boldsymbol{x}(t)\boldsymbol{x}^T(t+\tau_i)\} = \boldsymbol{W}\boldsymbol{D}_i\boldsymbol{W}^T, \tau_i \neq 0 \tag{3}$$

where $R_{\boldsymbol{s}}(\tau_i) = E\{\boldsymbol{s}(t)\boldsymbol{s}^T(t+\tau_i)\} = Diag[\rho_1^i(\tau_i), \rho_2^i(\tau_i), \ldots, \rho_N^i(\tau_i),] = \boldsymbol{D}_i$. BSS aims to identify the mixture matrix so that the source signals can be recovered. Eqs. (2) and (3) imply that when $\tau_i \neq 0$, the noise effect can be eliminated during mixture matrix identification. When $\boldsymbol{M}_i = R_{\boldsymbol{x}}(\tau_i)$, $i = 1, 2, \ldots, N$, the following equation is obtained:

$$\boldsymbol{M}_i = \boldsymbol{W}\boldsymbol{D}_i\boldsymbol{W}^T \tag{4}$$

Premultiplying $\boldsymbol{M}_i$ in Eq.(4) by $\boldsymbol{W}$'s inverse matrix $\boldsymbol{A}$ and postmultiplying by $\boldsymbol{A}^T$ yields:

$$\boldsymbol{A}\boldsymbol{M}_i\boldsymbol{A}^T = \boldsymbol{D}_i \tag{5}$$

Based on the analysis, we construct a cost function thus:

$$F(\boldsymbol{A}, \{\boldsymbol{D}_i\}) = \sum_{i}^{N} \|\boldsymbol{D}_i - \boldsymbol{A}\boldsymbol{M}_i\boldsymbol{A}^T\|^2. \tag{6}$$