THEORETICAL FOUNDATIONS OF SECOND-ORDER-STATISTICS-BASED BLIND SOURCE SEPARATION FOR NON-STATIONARY SOURCES

Akira TANAKA, Hideyuki IMAI and Masaaki MIYAKOSHI

Division of Computer Science,
Graduate School of Information Science and Technology, Hokkaido University
Kita-14, Nishi-9, Kita-ku, Sapporo, 060-0814, Japan

ABSTRACT

The aim of "Blind Source Separation" is to recover mutually independent unknown source signals from observations obtained through an unknown linear mixture system. Simultaneous diagonalization of correlation matrices (second-order statistics) of observations is one of the resolutions, when the unknown source signals are non-stationary. Although it is trivial that the true separation matrix simultaneously diagonalizes all the correlation matrices, it is not well investigated whether a simultaneous diagonalizer of the correlation matrices is always a separation matrix. In this paper, we give explicit solutions of simultaneous diagonalizers of the correlation matrices and we also clarify the condition that the solutions always achieve the blind source separation.

1. INTRODUCTION

The aim of "Blind Source Separation" is to recover mutually independent unknown source signals only from observations obtained through an unknown linear mixture system. As resolutions of the problem, many methods have been proposed [1]. It was reported that simultaneous diagonalization of correlation matrices (second-order statistics) of observations achieves the blind source separation, when the unknown source signals are non-stationary (see [2] for instance). Although it is trivial that the true (unknown) separation matrix simultaneously diagonalizes all the correlation matrices, it is not well investigated (except limited cases [3, 4]) whether a simultaneous diagonalizer of the correlation matrices is always a separation matrix.

In this paper, we give explicit solutions of simultaneous diagonalizers of the correlation matrices and we also clarify a necessary and sufficient condition that the solutions always achieve the blind source separation for general cases.

This work was partially supported by Grant-in-Aid No.16700001 for Young Scientist (B), Ministry of Education, Culture, Sports, Science and Technology, Japan.

2. OVERVIEW OF SECOND-ORDER-STATISTICS-BASED BLIND SOURCE SEPARATION

In this section, we formulate the second-order-statistics-based blind source separation. We adopt unitary spaces as signal spaces so that we can deal with the problem written in the short time Fourier domain.

Let n, m $(\leq n)$, and t be the number of observations, the number of sources, and the time index (or the frame index in the short time Fourier domain), respectively. Let $\boldsymbol{s}(t) \in \mathbf{C}^m$ be the mutually independent zero-mean unknown source signal vector and let $A \in \mathbf{C}^{n\times m}$ be an unknown mixture matrix with $\mathrm{rank}(A) = m$. We assume that the observation vector $\boldsymbol{x}(t) \in \mathbf{C}^n$ is given by the following model:

$$\boldsymbol{x}(t) = A\boldsymbol{s}(t). \tag{1}$$

The aim of the blind source separation in this formulation is to obtain a matrix $W \in \mathbf{C}^{m\times n}$ that recovers independency between any two elements of $\boldsymbol{y}(t)$ given by

$$\boldsymbol{y}(t) = W\boldsymbol{x}(t). \tag{2}$$

Let R_k, $(k \in \{1,\ldots,K\})$ be the correlation matrices of $\boldsymbol{x}(t)$ defined by

$$R_k = E_{T_k}[\boldsymbol{x}(t)\boldsymbol{x}^*(t)], \tag{3}$$

where X^* and E_{T_k} denote the adjoint matrix (or vector) of X and the expectation operator over intervals $t \in T_k$, respectively. We assume that $T_i \neq T_j$ for any $i, j \in \{1,\ldots,K\}$ with $i \neq j$. On the basis of Eq.(1) and the independency of the unknown source signals, R_k can be written by

$$R_k = A\Lambda_k A^*, \tag{4}$$

where Λ_k is a diagonal matrix defined by

$$\Lambda_k = E_{T_k}[\boldsymbol{s}(t)\boldsymbol{s}^*(t)]. \tag{5}$$

The key idea of the second-order-statistics-based blind source separation is that if a certain matrix W makes WR_kW^* diagonal matrices for all $k \in \{1,\ldots,K\}$, the matrix W is deeply related to A_L^{-1} (a left inverse matrix of A), when the unknown source signals are non-stationary. In many previous works, simultaneous diagonalization of R_k is formulated as the problem to find a minimizer of a criterion similar to

$$J(W) = \sum_k ||WR_kW^* - \mathrm{diag}(WR_kW^*)||^2, \qquad (6)$$

(see [2] for instance), where $||X||^2 = \mathrm{tr}[XX^*]$ and $\mathrm{diag}(X)$ denotes the diagonal matrix whose diagonal elements and their order are equal to those of X, and the minimization is generally achieved by the gradient descent method. Here, we introduce the following useful notation.

Definition 1 *[4] If there exists a non-singular matrix $P \in \mathbf{C}^{n\times n}$ that has exactly one non-zero element in each row and column, satisfying*

$$A = PB, \ \ A, \ B \in \mathbf{C}^{n\times m}, \qquad (7)$$

the matrices A and B are said to be essentially equal and denoted by $A \doteq B$.

It is guaranteed that $J(W)$ is minimized and the solution W surely achieves the separation, when the solution obtained by the gradient descent method is incidentally equal to

$$W \doteq A_L^{-1}. \qquad (8)$$

However, it is not so trivial that all minimizers of $J(W)$ achieve the separation, even if the unknown source signals are mutually independent and non-stationary. In [3, 4], conditions that achieve the separation in the case of $K = 2$ are discussed. However, consideration for general cases such as $K \geq 3$ is needed, since two *n.n.d* Hermitian matrices can be always simultaneously diagonalized [5], even if they do not have the structure written by Eq.(4).

3. EXPLICIT SOLUTIONS OF SIMULTANEOUS DIAGONALIZER

In this section, we give explicit solutions of simultaneous diagonalizers of R_k as a preparation for analyzing the relation between simultaneous diagonalization of correlation matrices and the blind source separation. The following theorem plays an essential role.

Theorem 1 [5] *Let $A_k \in \mathbf{C}^{n\times n}$ $(k \in \{1, \ldots, K\})$ be n.n.d. matrices and let $B = \sum_k A_k$. There exists a non-singular matrix M that makes M^*A_kM diagonal matrices for all $k \in \{1, \ldots, K\}$, if and only if*

$$A_iB^-A_j = A_jB^-A_i \qquad (9)$$

holds for any $i, j \in \{1, \ldots, K\}$, where B^- denotes an arbitrary generalized inverse matrix of B[5].

Hereafter, we give explicit solutions of simultaneous diagonalizers of R_k based on Theorem 1. Note that derivation of the explicit solutions is basically along with the proof of Theorem 1 with some supplements that make the solutions as general as possible.

We can make a non-singular matrix $\hat{A} = [A \ \ \tilde{A}] \in \mathbf{C}^{n\times n}$, where $\tilde{A} \in \mathbf{R}^{n\times(n-m)}$ is a matrix consisting of the orthogonal basis of $\mathcal{N}(A^*)$ (the null space of A^*), since $\mathrm{rank}(A) = m$. Thus, we can rewrite R_k as

$$R_k = \hat{A}\hat{\Lambda}_k\hat{A}^*, \qquad (10)$$

with

$$\hat{\Lambda}_k = \begin{bmatrix} \Lambda_k & O \\ O & O \end{bmatrix} \in \mathbf{C}^{n\times n}.$$

Although $\hat{A}^{-1}$ is unknown, it surely exists and simultaneously diagonalizes R_k for all $k \in \{1, \ldots, K\}$. Thus, on the basis of Theorem 1,

$$R_iB^-R_j = R_jB^-R_i \qquad (11)$$

holds for any $i, j \in \{1, \ldots, K\}$ with

$$B = \sum_k R_k = \hat{A}\hat{\Sigma}\hat{A}^*, \ \ \hat{\Sigma} = \left(\sum_k \hat{\Lambda}_k\right). \qquad (12)$$

Note that $\mathrm{rank}(B) = m$ holds.

Theorem 2 *$R_iB^-R_j$ is invariant for any B^-.*

Proof Let B_1^- and B_2^- be different generalized inverse matrices of B, then

$$B_2^- = B_1^- + Y - B_1^-BYBB_1^- \qquad (13)$$

holds with a certain matrix $Y \in C^{n\times n}$ [5]. Thus,

$$\begin{aligned} & R_iB_1^-R_j - R_iB_2^-R_j \\ = \ & R_iB_1^-R_j - R_i(B_1^- + Y - B_1^-BYBB_1^-)R_j \\ = \ & R_iYR_j - R_iB_1^-BYBB_1^-R_j \\ = \ & R_iYR_j - R_iYR_j = O \end{aligned}$$

holds, since $\mathcal{R}(R_k) \subset \mathcal{R}(B)$, $(k \in \{1, \ldots, K\})$, where $\mathcal{R}(X)$ denotes the range of the matrix X. □

On the basis of Theorem 2, we adopt the Moore-Penrose generalized inverse matrix[5] of B written by B^+ as B^- in the following contents. Let $B^+ = LL^*, (L \in \mathbf{C}^{n\times m})$ be a certain full-rank decomposition of B^+. From Eq.(11),

$$\begin{aligned} & (L^*R_iL)(L^*R_jL) \\ = \ & L^*(R_iB^+R_j)L = L^*(R_jB^+R_i)L \\ = \ & (L^*R_jL)(L^*R_iL) \end{aligned}$$

holds for any $i, j \in \{1, \ldots, K\}$, which means that $L^* R_i L$ and $L^* R_j L$ are commutable Hermitian matrices and they share all eigenvectors. Thus, it is concluded that $L^* R_k L$, $(k \in \{1, \ldots, K\})$ also share all eigenvectors. Let T be the matrix consisting of the eigenvectors of $L^* R_k L$ satisfying

$$L^* R_k L = T D_k T^*, \ k \in \{1, \ldots, K\}, \tag{14}$$

where D_k denotes *p.d.* diagonal matrices consisting the eigenvalues of $L^* R_k L$, then $T^* L^* R_k L T$, $(k \in \{1, \ldots, K\})$ are diagonal matrices, which means that

$$W = (LT)^* \tag{15}$$

simultaneously diagonalizes R_k. Note that procedures for calculating T are revealed in [6].

Lemma 1 *[7] Let $Z = LL^* = L_1 L_1^*$ be two arbitrary full-rank decompositions of n.n.d. Hermitian matrix Z, then there exists a unique unitary matrix C that makes $L_1 C = L$.*

Theorem 3 *A simultaneous diagonalizer Eq.(15) is invariant for any full-rank decomposition of B^+.*

Proof Let $B^+ = LL^* = L_1 L_1^*$ be different full-rank decompositions of B^+ with $L_1 \neq L$. According to Lemma 1, there exists a unique unitary matrix C that makes $L_1 C = L$. Thus, on the basis of Eq.(14),

$$L_1^* R_k L_1 = C L^* R_k L C^* = C T D_k T^* C^* \tag{16}$$

holds, which means that column vectors of CT are eigenvectors of $L_1^* R_k L_1$. Thus, a simultaneous diagonalizer based on L_1 is given by

$$W = (L_1 C T)^* = (L C^* C T)^* = (LT)^*, \tag{17}$$

which concludes the proof. □

In terms of diagonalization, any W satisfying $W \doteq (LT)^*$ is also a simultaneous diagonalizer of R_k

4. SIMULTANEOUS DIAGONALIZATION AND BLIND SOURCE SEPARATION

In this section, we discuss the relation between the blind source separation and simultaneous diagonalizers of R_k obtained in the previous section.

If a certain simultaneous diagonalizer W of R_k given by Eq.(15) achieves the blind source separation, then

$$X = WA \doteq I_m \tag{18}$$

must hold. Thus, we investigate properties of the matrix X in Eq.(18). Let us consider the Hermitian matrix written by

$$B_S = (\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{A}^{-1}. \tag{19}$$

Lemma 2

$$B B_S = B_S B. \tag{20}$$

Proof Note that

$$B B_S = \hat{A} \hat{\Sigma} \hat{\Sigma}^+ \hat{A}^{-1}, \ \ B_S B = (\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{\Sigma} \hat{A}^*,$$

$$\hat{\Sigma} \hat{\Sigma}^+ = \hat{\Sigma}^+ \hat{\Sigma} = \begin{bmatrix} I_m & O \\ O & O \end{bmatrix},$$

hold, where I_m denotes the identity matrix of degree m. Also note that

$$\hat{A}^* \hat{A} = \begin{bmatrix} A^* A & O \\ O & I_{n-m} \end{bmatrix}$$

holds. Therefore,

$$\begin{aligned} & B B_S - B_S B \\ &= \hat{A} \hat{\Sigma} \hat{\Sigma}^+ \hat{A}^{-1} - (\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{\Sigma} \hat{A}^* \\ &= (\hat{A}^*)^{-1} (\hat{A}^* \hat{A} \hat{\Sigma} \hat{\Sigma}^+ - \hat{\Sigma}^+ \hat{\Sigma} \hat{A}^* \hat{A}) \hat{A}^{-1} \\ &= (\hat{A}^*)^{-1} \left(\begin{bmatrix} A^* A & O \\ O & I_{n-m} \end{bmatrix} \begin{bmatrix} I_m & O \\ O & O \end{bmatrix} \right. \\ & \qquad \left. - \begin{bmatrix} I_m & O \\ O & O \end{bmatrix} \begin{bmatrix} A^* A & O \\ O & I_{n-m} \end{bmatrix} \right) \hat{A}^{-1} = O \end{aligned}$$

is concluded. □

Theorem 4 *B_S is the Moore-Penrose generalized inverse matrix of B.*

Proof

$$\begin{aligned} & B B_S B \\ &= (\hat{A} \hat{\Sigma} \hat{A}^*)(\hat{A}^{*-1} \hat{\Sigma}^+ \hat{A}^{-1})(\hat{A} \hat{\Sigma} \hat{A}^*) = \hat{A} \hat{\Sigma} \hat{\Sigma}^+ \hat{\Sigma} \hat{A}^* \\ &= \hat{A} \hat{\Sigma} \hat{A}^* = B, \\ & B_S B B_S \\ &= ((\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{A}^{-1})(\hat{A} \hat{\Sigma} \hat{A}^*)((\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{A}^{-1}) \\ &= (\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{\Sigma} \hat{\Sigma}^+ \hat{A}^{-1} = (\hat{A}^*)^{-1} \hat{\Sigma}^+ \hat{A}^{-1} = B_S, \end{aligned}$$

hold, and on the basis of Lemma 2 and the fact that B and B_S are Hermitian matrices,

$$\begin{aligned} (B B_S)^* &= B_S^* B^* = B_S B = B B_S, \\ (B_S B)^* &= B^* B_S^* = B B_S = B_S B, \end{aligned}$$

hold, which conclude the proof. □

According to Theorem 4, it is easy to show that

$$L_S = (\hat{A}^*)^{-1} \begin{bmatrix} I_m \\ O \end{bmatrix} \Sigma^{-1/2}, \tag{21}$$

satisfies $B^+ = L_S L_S^*$, where $\Sigma = (\sum_k \Lambda_k)$. Thus, adopting $L_S L_S^*$ as a full-rank decomposition of B^+ yields

$$\begin{aligned} & L_S^* R_k L_S \\ &= \Sigma^{-1/2} \left[\begin{array}{cc} I_m & O \end{array} \right] \hat{A}^{-1} \hat{A} \hat{\Lambda}_k \hat{A}^* (\hat{A}^*)^{-1} \\ & \quad \times \left[\begin{array}{c} I_m \\ O \end{array} \right] \Sigma^{1/2} \\ &= \Sigma^{-1/2} \Lambda_k \Sigma^{-1/2} = \Sigma^{-1} \Lambda_k, \end{aligned}$$

which means that $L_S^* R_k L_S$, $(k \in \{1, \ldots, K\})$ are already diagonalized. Let T_S be the unitary matrix consisting of the eigenvectors of $L_S^* R_k L_S$ satisfying

$$\Sigma^{-1} \Lambda_k = T_S \Sigma^{-1} \Lambda_k T_S^*, \quad k \in \{1, \ldots, K\}, \tag{22}$$

then a simultaneous diagonalizer of R_k is written by $W = (L_S T_S)^*$. We can not directly calculate L_S and T_S, since $\hat{A}$ is unknown. However, it is guaranteed that the solution $(L_S T_S)^*$ is essentially identical to the solution Eq.(15) on the basis of Theorem 3 and Theorem 4. Thus,

$$\begin{aligned} X &= WA = (LT)^* A = (L_S T_S)^* A \\ &= T_S^* \Sigma^{-1/2} \left[\begin{array}{cc} I_m & O \end{array} \right] \hat{A}^{-1} \hat{A} \left[\begin{array}{c} I_m \\ O \end{array} \right] = T_S^* \Sigma^{-1/2} \end{aligned}$$

holds with the columns of T being appropriately permuted in Eq.(14). Accordingly, the unitary matrix T_S plays an essential role for whether the simultaneous diagonalizer W achieves the separation or not. Accordingly, we obtain the following main theorem.

Theorem 5 *The second-order-statistics-based blind source separation is always achieved, if and only if the unitary matrix T_S satisfies*

$$T_S \doteq I_m. \tag{23}$$

Proof If $T_S \doteq I_m$ is satisfied,

$$X = WA = T_S^* \Sigma^{-1/2} \doteq \Sigma^{-1/2} \doteq I_m$$

holds, which means that the separation is trivially achieved.

Contrarily, if T_S is a unitary matrix that does not satisfy Eq.(23), then one or more rows of X must have two or more non-zero elements. Thus, the corresponding elements of $\boldsymbol{y}(t) = W\boldsymbol{x}(t) = X\boldsymbol{s}(t)$ must include mixed source signals, which means that the separation fails. □

Here, we consider the case that Eq.(23) is not satisfied in Theorem 5, that is, the case that simultaneous diagonalizers may fail the separation. When Eq.(23) is not satisfied, $\Sigma^{-1}\Lambda_k$, $(k \in \{1, \ldots, K\})$ necessarily share at least one eigenspace, whose dimension is larger than 1, derived from eigenvalues of multiple root in the same position. Let $\lambda_i^{(k)}$ and σ_i, $(i \in \{1, \ldots, m\})$ be i-th diagonal elements of Λ_k, $(k \in \{1, \ldots, K\})$ and Σ, respectively; and let

$$\boldsymbol{v}_i = [\lambda_i^{(1)} \ \lambda_i^{(2)} \ \cdots \ \lambda_i^{(K)}], \ \ \boldsymbol{u}_i = \boldsymbol{v}_i / \sigma_i.$$

If all $\Sigma^{-1}\Lambda_k$ have eigenvalues of multiple root in the same position such as $i_1, i_2, \cdots$, and i_ℓ-th elements with $\ell \le m$, $i_1 < i_2 < \cdots < i_\ell \in \{1, \ldots, m\}$, then

$$\boldsymbol{u}_{i_1} = \boldsymbol{u}_{i_2} = \cdots = \boldsymbol{u}_{i_\ell} \tag{24}$$

holds, which means that the transitions of $\lambda_i^{(k)}$ and $\lambda_j^{(k)}$ for any $i, j \in \{i_1, i_2, \ldots, i_\ell\}$ with respect to k are proportional, since $\boldsymbol{u}_i$ is normalized version of $\boldsymbol{v}_i$ by its L_1-norm. Thus, it is concluded that a simultaneous diagonalizer may fail the separation when Eq.(24) holds, even if the source signals are mutually independent and non-stationary. This result is a generalized version of the results obtained in [3, 4].

5. CONCLUSION

In this paper, we gave explicit solutions of simultaneous diagonalizers of the correlation matrices for the second-order-statistics-based blind source separation and also clarified a necessary and sufficient condition that the solutions always achieve the blind source separation.

6. REFERENCES

[1] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[2] L. C. Parra and C. V. Alvino, "Convolutive Blind Separation of Non-Stationary Sources," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 320–327, 2000.

[3] L. Tong, R. Liu, V. C. Soon, and Y. Huang, "Indeterminacy and Identifiability of Blind Identification," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 5, pp. 499–509, 1991.

[4] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A Blind Source Separation Technique Using Second-Order Statistics," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 434–444, 1997.

[5] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications*, John Wiley & Sons, 1971.

[6] M. L. Mehta, *Matrix Theory - Selected Topics and Useful Results*, les editions de phisique, 1989.

[7] C. R. Rao, *Linear Statistical Inference and its Applications (Second Edition)*, John Wiley & Sons, 2002.