An Efficient Algorithm to Estimate Mixture Matrix in Blind Source Separation using Tensor Decomposition

Yan-liang Zhang*, Geng Li

School of Computer Science & Technology, Henan Polytechnic University, Jiaozuo 454001,China

Abstract

The estimation of mixing matrix is a key step to solve the problem of blind source separation. The existing algorithm can only estimate the matrix of well-determined, over-determined and under-determined in condition of sparse source. Scaling and permutation ambiguities lie in both factor matrix of tensor Canonical Decomposition and mixing matrix in blind source separation. With this property, the estimation of mixing matrix can be transformed into tensor Canonical Decomposition of observed signal's statistic. The decomposition can be realized by the method of alternating least squares. The theoretical analysis and simulations show that the method proposed in this paper is an efficient algorithm to estimate well-determined, over-determined and under-determined mixing matrix.

***Keywords**: Blind Source Separation(BSS); Tensor; Canonical Decomposition; Alternating Least Squares(ALS)*

1 Introduction

Blind source separation is an area of great interest due to its ability to separate multiple independent sources from array observations without requiring a priori knowledge of the location of sources or the geometry of sensor array[1]. Such flexibility has made BSS a potential technique in a variety of applications, such as multitalker speech separation from multimicrophone audio recordings, elimination of interuser interference in wireless communications and biomedical signals processing like ECG and EEG.

There are two steps to estimate mixture matrix. First, estimate the mixture matrix. Second, recover source signals. The accuracy of estimated mixture matrix dramatically impact the result of BSS. There are many algorithms, such as Fast Approximate Joint Diagonalization(FAJD)[2] , Alternating Colums Diagonal Cetres(ACDC)[3],etc. that can solve well-determined, over-determined BSS problem. But, these algorithms are not suitable for estimating under-determined mixture matrix. Under-determined means the numbers of unknown variable exceed the equations and the matrix to be estimated is "short and fat" .To solve this problem, it is regularly assumed that the source signals are sparse in time domain or transform domain[4]. In this situation, for the clustering characteristic of sparse mixing signals, scatter diagram of the observed data assembled in direction vector of mixing matrix, and under-determined mixing matrix could be obtained by local maximization algorithm. This method is effective for seperating time domain sparse signal,such as audio signals. But it is not applicable to most Sub-Gaussian signals that are not satisfy the hypothesis of sparseness. Delathauwer propose the method of tensor decomposition to estimate the under-determined mixture matrix, that need not the hypothesis of sparseness of source signals[5]. It is a new idea for under-determined BSS.

In this paper, the problem of estimation mixture matrix of well-determined, super-determined, under-determined is addressed by canonical decomposition of tensor, and the canonical decomposition is obtained by the method of alternating least squares.

2 Relation between BSS and Tensor's Canonical Decomposition

2.1 Model of BSS

The mixing model of BSS is below

$$\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t)+\mathbf{n}(t) \tag{1}$$

where $\mathbf{x}(t)\in\mathbb{R}^J$ represent mixture signal vector of J dimension, $\mathbf{s}(t)\in\mathbb{R}^R$ is source signal vector of R dimension whose component is independent each other, $\mathbf{n}(t)\in\mathbb{R}^J$ indicate additive noise which is independent with source signal, $\mathbf{A}\in\mathbb{R}^{J\times R}$ express a full row rank matrix which is to be identificated.

The identification of $\mathbf{A}$ is first step to recover source signals. There are inherent scaling and permutation uncertainty of column in estimation matrix $\mathbf{A}$. The estimation of $\mathbf{A}$ writes $\hat{\mathbf{A}}$,then these two uncertainty expressed as follow:

$$\mathbf{A} = \hat{\mathbf{A}}\mathbf{PD} \tag{2}$$

where, $\mathbf{P}$ represent $R \times R$ permutation matrix, $\mathbf{D}$ indicate $R \times R$ non-singular diagonal matrix.

2.2 Tensor's Canonical Decomposition and Its Essence Uniqueness Condition

A tensor is a multidimendional array,and it is a higher order extensions of the matrix.There are two forms in tensor decompositions:canonical decomposition and tucker decomposition. Tensor decomposition is widely used in psychometrics and chemometrics. For its powerful data analysis functions, during the last decades, researcher has begun to apply it in information area,such as signal processing, machine vision, data mining,etc.
The identification of mixing matrix can be transformed into the tensor's canconical decomposition. For elaborating this method, introduce the conception of canconical decomposition.

Definition Ⅰ [6]**.** Outer product.The outer product of three vectors $\mathbf{u} \in \mathbb{R}^M$ 、 $\mathbf{v} \in \mathbb{R}^N$ 、 $\mathbf{w} \in \mathbb{R}^P$ is $\underline{\mathbf{Z}} \in \mathbb{R}^{M \times N \times P}$ which is a three order tensor. It writes $\underline{\mathbf{Z}} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}$, where each element of $\underline{\mathbf{Z}}$ can obtained by

$$z_{mnp} = u_m v_n w_p \quad m = 1,...,M; n = 1,...,N; p = 1,...,P \tag{3}$$

Definition Ⅱ [6]**.** Rank One Tensor. A three order tensor $\underline{\mathbf{Z}} \in \mathbb{R}^{M \times N \times P}$ is rank one if it can be written as the outer product of three vectors, i.e.,

$$\underline{\mathbf{Z}} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w} \tag{4}$$

Definition Ⅲ [6]**.**Tensor's rank. Every tensor can be expressed as a sum of rank one tensors. The rank of a general tensor $\underline{\mathbf{Y}}$ is defined to be the minimum number of rank one tensors with which it is possible to express $\underline{\mathbf{Y}}$ as a sum.

Definition Ⅳ [6]**.**Canonical Decomposition. The definition of $\underline{\mathbf{Y}} \in \mathbb{R}^{M \times N \times P}$ canonical decomposition is

$$\underline{\mathbf{Y}} = \sum_{i=1}^{rank(\underline{\mathbf{Y}})} \mathbf{u}_i \circ \mathbf{v}_i \circ \mathbf{w}_i = \mathbf{U}, \mathbf{V}, \mathbf{W} \tag{5}$$

where, $\mathbf{u}_i \in \mathbb{R}^M$, $\mathbf{v}_i \in \mathbb{R}^N$, $\mathbf{w}_i \in \mathbb{R}^P$, $rank(\underline{\mathbf{Y}})$ is the rank of tensor $\underline{\mathbf{Y}}$,writes $I = rank(\underline{\mathbf{Y}})$; $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_I]$, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_I]$, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_I]$ are factor matrix of canonical decomposition.

Below we discuss the essence uniqueness condition of $\underline{\mathbf{Y}}$ canonical decomposition. Apparently, there are inherent scaling and permutation uncertainty in factor matrix. If $\alpha_i \beta_i \gamma_i = 1 (1 \le i \le I)$ stands, then $\underline{\mathbf{Y}} = \sum_{i=1}^{I} (\alpha_i \mathbf{u}_i) \circ (\beta_i \mathbf{v}_i) \circ (\gamma_i \mathbf{w}_i)$ work. This is called scaling uncertainty. Supposing $\mathbf{P}$ is a $I \times I$ permutation matrix, then $\underline{\mathbf{Y}} = \mathbf{U}, \mathbf{V}, \mathbf{W} = \mathbf{UP}, \mathbf{VP}, \mathbf{WP}$ work, and this is called permutation uncertainty. If there are only scaling and permutation indeterminacy in factor matrix of the tensor, then canonical decomposition is called uniqueness.

In conclusion, there are only scaling and permutation indeterminacy both in the factor matrix of canonical decomposition and the estimated mixture matrix of BSS. This common is theoretical basis for transforming the problem of estimating the mixture matrix into tensor's canonical decomposition.

3 Tensor Method to Estimate Mixture Matrix

3.1 Model Transformation

The second-order correlation matrix of mixture signals in BSS model (1) is

$$\mathbf{C}_k = E\left[\mathbf{x}(t)\mathbf{x}^{\mathrm{T}}(t+\tau_k)\right] = \mathbf{A}\mathbf{D}_k\mathbf{A}^{\mathrm{T}} \quad k = 1,...,K \tag{6}$$

where $\mathbf{D}_k = E\left[\mathbf{s}(t)\mathbf{s}^{\mathrm{T}}(t+\tau_k)\right]$ is a diagonal matrix. To confirm robustness of the estimation, it is supposed that $K \gg R$ .Assemble a three order tensor with matrix $\mathbf{C}_k \in \mathbb{R}^{J \times J}$ as follows

$$\underline{\mathbf{C}} \in \mathbb{R}^{J \times J \times K} : \ (\underline{\mathbf{C}})_{ijk} = (\mathbf{C}_k)_{ij} ; i, j = 1,...,J ; \quad k = 1,...,K . \tag{7}$$

Define matrix $\mathbf{D} \in \mathbb{R}^{K \times R}$ as follows

$$(\mathbf{D})_{kr} = (\mathbf{D}_k)_{rr} , \quad k = 1,...,K , \quad r = 1,...,R . \tag{8}$$

Then tensor $\underline{\mathbf{C}}$ can be decomposed as

$$\underline{\mathbf{C}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{d}_r = \mathbf{A}, \mathbf{A}, \mathbf{D} \tag{9}$$

where $\mathbf{a}_r$ and $\mathbf{d}_r$ are r-th column of $\mathbf{A}$ and $\mathbf{D}$,respectivly. Formula (9) is the canonical decomposition of tensor $\underline{\mathbf{C}}$.

If the canonical decomposition of $\underline{\mathbf{C}}$ is unique, then mixture matrix of BSS can be estimated with it. There provided a sufficient condition that canonical decomposition exists solely in literature[5]. From the sufficient condition, when the number of receiving sensor is constant, the maximum of source signal that can guarantee uniqueness of canonical decomposition, can be deduced.

Table 1: Relation between sensor number J and maximum of source number R_{max} allowed

J	2	3	5	8	10	15	18	23
R_{max}	2	4	10	26	41	92	132	216

From table 1 we can see that, fixing the number of receiving sensor J , so long as the number of source signal satisfies $R \le R_{max}$, the method of canonical decomposition can be used to estimate mixture matrix. So, tensor method can estimate the mixture matrix of well-determined, super-determined and under-determined.

4Alternating Least Squares(ALS) Algorithm to Realize Canonical Decomposition

In this section, we realize canonical decomposition by minimizing the cost function. Choosed function follows as

$$f(\mathbf{A},\mathbf{D}) = \left\| \underline{\mathbf{C}} - \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{a}_r \circ \mathbf{d}_r \right\|_F^2 \qquad (10)$$

For this multi-variable optimization problem, cyclic minimizing is a common method. The thought of this method is, partitioning variable set into several subsets, utilize optimization algorithm to calculate one of subsets in every step, and regard other subsets as constant, repeat this process until cost function is convergence.

In accordance with the formula (10), we adopt ALS algorithm to realize cyclic minimizing.

The first, second and third slice matrix of tensor $\underline{\mathbf{C}}$ are respectively

$$\mathbf{E}_i \in \mathbb{R}^{J\times K}, i = 1,...,J; \qquad (11)$$

$$\mathbf{F}_j \in \mathbb{R}^{K\times J},\ j = 1,...,J; \qquad (12)$$

$$\mathbf{G}_k \in \mathbb{R}^{J\times J}, k = 1,\ ...,K\ ; \qquad (13)$$

Combining formula(9) and symmetry of tensor $\underline{\mathbf{C}}$, we can deduce

$$\mathbf{G}_k = \mathbf{A}diag_k(\mathbf{D})\mathbf{A}^{\mathrm{T}}, \quad k = 1,...,K \qquad (14)$$

$$\mathbf{E}_i = \mathbf{A}diag_i(\mathbf{A})\mathbf{D}^{\mathrm{T}}, \quad i = 1,...,J \qquad (15)$$

$$\mathbf{F}_j = \mathbf{D}diag_j(\mathbf{A})\mathbf{A}^{\mathrm{T}}, \quad j = 1,...,J \qquad (16)$$

Formula(14) can be rewritten as

$$\begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \vdots \\ \mathbf{G}_K \end{bmatrix} = \begin{bmatrix} \mathbf{A}diag_1(\mathbf{D}) \\ \mathbf{A}diag_2(\mathbf{D}) \\ \vdots \\ \mathbf{A}diag_K(\mathbf{D}) \end{bmatrix} \mathbf{A}^{\mathrm{T}} \qquad (17)$$

From formula (17), least square estimation of mixing matrix $\mathbf{A}$ can be inferred as

$$\hat{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} \mathbf{A}diag_1(\mathbf{D}) \\ \mathbf{A}diag_2(\mathbf{D}) \\ \vdots \\ \mathbf{A}diag_K(\mathbf{D}) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \vdots \\ \mathbf{G}_K \end{bmatrix} \qquad (18)$$

where symbol“ † ”denotes moore-penrose inverse matrix. Similarly, least square estimation of $\mathbf{D}$ and $\mathbf{A}$ can be obtained respectively as

$$\hat{\mathbf{D}}^{\mathrm{T}} = \begin{bmatrix} \mathbf{A}diag_1(\mathbf{A}) \\ \mathbf{A}diag_2(\mathbf{A}) \\ \vdots \\ \mathbf{A}diag_J(\mathbf{A}) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \\ \vdots \\ \mathbf{E}_J \end{bmatrix} \qquad (19)$$

$$\hat{\mathbf{A}}^{\mathrm{T}} = \begin{bmatrix} \mathbf{D}diag_1(\mathbf{A}) \\ \mathbf{D}diag_2(\mathbf{A}) \\ \vdots \\ \mathbf{D}diag_J(\mathbf{A}) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \vdots \\ \mathbf{F}_J \end{bmatrix} \qquad (20)$$

Random initialize matrix $\mathbf{A}$ and $\mathbf{D}$, then update $\hat{\mathbf{A}}$ and $\hat{\mathbf{D}}$ according to formula (18),(19)and (20).In process of updating, $\mathbf{A}$ and $\mathbf{D}$ on the right side of equation are replaced by $\hat{\mathbf{A}}$ and $\hat{\mathbf{D}}$ that are obtained by last step of estimation.

4 Simulation Results

In this section, we will compare the performance of our proposed method with the other algorithms presented in paper [2, 3]. We mark algorithm in this paper, in paper[2],[3]respectively as GTCD, FAJD and ACDC.

IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 1, No 2, January 2014
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
www.IJCSI.org

To measure the difference between the mixing matrix and the estimated matrix, the performance index, called the relative error (REER), is adopted. The REER has been frequently used in the evaluation in area of BSS[7]. The REER is given by

$$REER = \mathrm{E}\left\{ \frac{\left\| \mathbf{A} - \hat{\mathbf{A}} \right\|_{\mathrm{F}}}{\left\| \mathbf{A} \right\|_{\mathrm{F}}} \right\} \tag{21}$$

where, $\hat{\mathbf{A}}$ is the estimation of mixing matrix $\mathbf{A}$,under the condition that their columns are unitized and are eliminated permutation indeterminacy. For the convenience of comparing, transform the REER into decibel(dB) using formula $10\log_{10}(REER)$. It is evident that, lower the decibel of relative error is , preciser the estimation of mixing matrix is.

The second-order correlation matrices are generated by formula(6), where the mixing matrix $\mathbf{A}$ is $M \times N$,diagonal matrix $\mathbf{D}_k$ is $N \times N$,and their elements are random generated with standard normal distribution in each experiment. These matrices are assembled into third order tensor $\underline{\mathbf{C}} \in \mathbb{R}^{M \times M \times K}$ following formula(7).

Experiments 1. In this experiment, we compare the performance of eatimation well-determined matrix between the GTCD algorithm and FAJD algorithm based on non-unitary joint diagonalization. Assume the number of correlation function $K = 100$. The iteration stop condition is set as

$$\left| \left(f_n\left(\hat{\mathbf{A}}', \hat{\mathbf{D}}'\right) - f_{n-1}\left(\hat{\mathbf{A}}', \hat{\mathbf{D}}'\right) \right) / f_n\left(\hat{\mathbf{A}}', \hat{\mathbf{D}}'\right) \right| < 10^{-6} \tag{22}$$

where the definition of f function follows formula (10).

Note:1) The unit of Runtime is seconds.2)REER is measured with dB. The same below.

Estimate the mixing matrix of different M using algorithm GTCD and FAJD respectively, and each scenario is repeated 100 times independently. The mean of REER and run time in each case are presented in Tab.2.The table illustrates that the accuracy of GTCD is better than that of FAJD. Except the low dimension case, such as $M = N = 4$, $M = N = 6$, the run time of FAJD outnumber GTCD. And the run time of FAJD rapid increase with the dimension augment of mixing matrix.
Experiments 2. In this experiment, we compare the performance of estimation super-determined matrix between the GTCD algorithm and ACDC algorithm based on non-unitary joint diagonalization. Set different value of M and N satisfying $M > N$ and independently repeat 100 times. The mean of REER and run time in each case are presented in Tab.3. As can be seen from the table that, in each case, the precision of GTCD is better than that of ACDC and the run time of ACDC outnumber TCD. Similar to experiment 1, the run time of FAJD rapid increase with the dimension augment of mixing matrix.

Table 2: Accuracy and runtime of two algorithms in well-determined case

	M=6,N=6		M=10,N=10		M=16,N=16		M=20,N=20		M=40,N=40	
	REER	Runtime	REER	Runtime	REER	Runtime	REER	Runtime	REER	Runtime
GTCD	-155.5	1.2034	-156.3	1.2655	-161.8	1.4023	-156.5	1.8284	-155.1	7.4014
FAJD	-95.4	0.1764	-92.6	0.7216	-96.2	2.7487	-82.2	10.1810	-11.0	423.3555

Table 3: Accuracy and runtime of two algorithms in super-determined case

	M=6,N=4		M=10,N=6		M=16,N=10		M=20,N=12		M=40,N=30	
	REER	Runtime	REER	Runtime	REER	Runtime	REER	Runtime	REER	Runtime
GTCD	-345.4	0.3042	-339.4	0.3770	-321.6	0.5089	-330.9	0.6537	-306.7	3.0930
ACDC	-84.4	1.4134	-95.9	2.4805	-100.9	7.8012	-107.1	13.9125	-121.91	158.4669

Table 4: Accuracy and runtime of GTCD in under-determined case

	M=4,N=5	M=5,N=6	M=6,N=8	M=8,N=10	M=10,N=16	M=15,N=20	M=18,N=30	M=23,N=40
REER	-155.1	-149.4	-146.5	-144.2	-140.0	-134.6	-111.3	-72.6
Runtime	1.4748	1.5918	2.5185	3.0151	4.2947	5.4100	9.1283	15.6118

Experiments 3. The performance of GTCD algorithm for estimating under-determined mixing matrix is analysed in this experiment. Assume different value of M and N and independently repeat 100 times in each case. The mean of REER and run time are illustrate in Tab.4. It is indicated that the GTCD can address the problem of under-determined BSS, but the ACDC and FAJD can not resolve this problem.

5 Conclusion

In this paper, the estimation of mixing matrix is transformed into the problem of tensor canonical decomposition. Tensor decomposition is achieved by ALS algorithm. The proposed method not only can estimate the mixing matrix of well-determined and super-determined, but also can address under-determined matrix. The expriments show that the accuracy and the run time of our algorithm is improver than that of the existing algorithm based on non-unitary joint diagonalization.

Acknowledgement

I would like to express my gratitude to the support of the National Natural Science Funds of China [61201299，61074102]and the Doctoral Scientific Funds of Henan polytechnic University [648753].

References

[1] Jumin ZHAO, Dengao LI, Haiyan ZHANG, Qinglun MA. Second Order Blind Source Separation Algorithm. Journal of Computational Inforamtional Systems. 2011, 7 (3) : 979- 983

[2]LI X L, ZHANG X D. Non-orthogonal approximate joint diagonalization free of degenerate solution. IEEE Trans. Signal Process., 2007, 55(51): 1803–1814.

[3]YEREDOR A. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. IEEE Trans. Signal Process., 2002, 50(7): 1545–1553.

[4]AISSA-EL-BEYA,LINH-TRUNG ,ABED-MERAIM K,et al.Underdetermind blind separation of nondisjoint sources in the time-frequency domain. IEEE Transactions on Signal Processing, 2007,55(3): 897–907.

[5]DE LATHAUWER L, JOSÉPHINE CASTAING, Cardoso, Jean-Franois. Fourth-Order Cumulant-Based Blind Identification of Underdetermined Mixtures. IEEE Transactions on Signal Processing, 2007,55(6): 2965-2973.

[6]KOLDA T G,BADER B W. Tensor Decompositions and Applications. SIAM Review, 2009,51(3): 455–500.

[7]DE LATHAUWER L, JOSÉPHINE CASTAING. Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization. IEEE Transactions on Signal Processing, 2008,56(3): 1096–1105.

Yan-liang Zhang received Doctor degree at Xidian University in 2011. She has been employed at Henan Polytechnic University since the summer of 2001. She has been supported by the National Natural Science Foundation of China. Her research fields are blind source separation and independent component analysis.

Geng Li born in 1982, Master of Engineering. . His research interests include wireless communication and Internet of things.