HAL
archives-ouvertes.fr

Geometrical Method Using Simplicial Cones for Overdetermined Nonnegative Blind Source Separation: Application to Real PET Images

Wendyam Ouedraogo, Antoine Souloumiac, Mériem Jaidane, Christian Jutten

HAL Id: hal-00687436

https://hal.archives-ouvertes.fr/hal-00687436

Submitted on 13 Apr 2012

Geometrical Method Using Simplicial Cones for Overdetermined Nonnegative Blind Source Separation: Application to Real PET Images

Wendyam S.B. Ouedraogo[1,2,3], Antoine Souloumiac[1],
Meriem Jaidane[2], and Christian Jutten[3]

[1] CEA, LIST, Laboratoire d'Outils pour l'Analyse de Données,
Gif-sur-Yvette, F-91191, France
[2] Unité Signaux et Systèmes, École Nationale d'Ingénieurs de Tunis,
BP 37, 1002 Tunis, Tunisie
[3] GIPSA-lab, UMR 5216 CNRS, Université de Grenoble, 961 rue de la Houille
Blanche BP 46 F-38402 Grenoble Cedex, France
wendyam-serge-boris.ouedraogo@cea.fr

Abstract. This paper presents a geometrical method for solving the overdetermined Nonnegative Blind Source Separation (N-BSS) problem. Considering each column of the mixed data as a point in the data space, we develop a Simplicial Cone Shrinking Algorithm for Unmixing Nonnegative Sources (SCSA-UNS). The proposed method estimates the mixing matrix and the sources by fitting a simplicial cone to the scatter plot of the mixed data. It requires weak assumption on the sources distribution, in particular the independence of the different sources is not necessary. Simulations on synthetic data show that SCSA-UNS outperforms other existing geometrical methods in noiseless case. Experiment on real Dynamic Positon Emission Tomography (PET) images illustrates the efficiency of the proposed method.

Keywords: Blind Source Separation, Nonnegativity, Simplicial Cone, Minimum Volume, Facet.

1 Introduction

We deal with the problem of Nonnegative Blind Source Separation (N-BSS) in noiseless, linear intantaneous mixture case. The mixture model is given by:

$$\underset{m \times p}{X} = \underset{m \times n}{A} \ \underset{n \times p}{S} \tag{1}$$

where m is the number of observations, n is the number of sources and p is the number of samples. X, A and S are respectively the given observations matrix, the unknown mixing matrix and the hidden nonnegative sources matrix. N-BSS consists on retrieving S and A given only X. Possible directions for solving problem (1) are the geometrical approaches. These methods are very natural and intuitive and require weak assumption on the sources distribution. The first geometrical method

was introduced by Puntonet et al. [1] for separating two sources having bounded probability density functions. The mixing matrix is retrieved by finding the slopes of the parallelogram containing the scatter plot of mixed data. Babaie-Zadeh et al. [2] propose another geometrical method applicable to more than two sources based on clustering the observations points and fitting a line (hyper-plane) to each cluster to recover the mixing matrix. By assuming the (very strong) condition ***local dominance of the sources*** (i.e. for every source there is at least one instant where the underlined source is active and all the others are not), several authors propose estimating the mixing matrix by looking for the vertices of the convex hull of the scatter plot of mixed data [3][4]. These methods can unfortunately be very slow and demanding very large size samples, specially for large scale problem due to the convex hull computing. Noting that when the sources are nonnegative, the scatter plot of mixed data is contained in the simplicial cone generated by the mixing matrix, other geometrical methods were proposed for solving problem (1) by looking for the Minimum Volume (MV) simplicial cone containing the mixed data [5][6]. MV like methods do not require local dominance of sources. But the simplicial cone generated by the mixing matrix must be well recognizable from the scatter plot of mixed data. This weaker condition implies that there should be at least $n-1$ mixed data points on, or close to, each ***facet*** of this cone. It's also necessary to specify that beside the nonnegativity, some MV like methods developed for hyperspectral data processing [5][6] require ***full additivity of the sources*** (i.e. the sum on every column of the sources matrix is equal to one).

This paper presents a MV like method for solving overdetermined N-BSS problem called Simplicial Cone Shrinking Algorithm for Unmixing Nonnegative Sources (SCSA-UNS). This work establishes an extension of [9] in which we only consider the determined case with nonnegative mixing matrix. Section 2 reviews the geometrical view of the N-BSS problem and derives the main idea of MV like methods. In section 3, we describe the proposed SCSA-UNS method. Section 4 presents simulation results on synthetic data and real Dynamic Positon Emission Tomography (PET) data and comparisons with other MV like methods. Finally section 5 presents the conclusions and future works.

2   Geometrical View of N-BSS Problem

Let's review the geometrical view of the N-BSS problem we described in [9]. For a given matrix $W = [w_1, w_2, \cdots, w_n]$, we define the ***Simplicial Cone*** $Span^+(W)$ generated by W by :

$$Span^+(W) = \left\{ z \, \big| \, z = Wy \text{ with } y \in \mathbb{R}^n_+ \right\} \tag{2}$$

We also define the ***positive orthant*** $\mathbb{R}^n_+$ as being the simplicial cone generated by the identity matrix I_n: $\mathbb{R}^n_+ = Span^+(I_n)$

By considering each column x_i of X ($1 \leq i \leq p$) as a point in the n dimension data space, it comes that when the sources are nonnegative, the mixed data form a cloud of points contained in the simplicial cone generated by the mixing matrix.

$$\{x_i \, | \, x_i \in X, 1 \leq i \leq p\} \subseteq Span^+(A) \tag{3}$$

One can thus imagine to estimate A (up to the classical (positive here) scaling and permutation BSS indetermination's) by finding a simplicial cone containing all the mixed data. But without any additional condition there is an infinite number of such cones. So for recovering the mixing matrix, we require the scatter plot of the mixed data to **fill enough** $Span^+(A)$ and we look for the Minimum Volume (i.e. the smallest) simplicial cone containing the mixed data [7]. Filling enough means that $Span^+(A)$ must be well recognizable from $\{x_i, x_i \in X, 1 \leq i \leq p\}$ (i.e. there should be at least $n-1$ mixed data points on, or close to, each **facet** of $Span^+(A)$ [8]). This intuitive and natural condition will be defined more formally in a future work.

3 Geometrical Method Using Simplicial Cones for Overdetermined Nonnegative Blind Source Separation

3.1 Determined Case : Simplicial Cone Shrinking Algorithm for Unmixing Nonnegative Sources (SCSA-UNS)

We first restrict to determined case with full column rank nonnegative mixing matrix (i.e. $m = n$ and $A \geq 0$). This case is considered in [9]. SCSA-UNS aims at estimating A by finding the Minimum Volume simplicial cone containing all the mixed data. In this objective, we propose a criterion for measuring the volume of a given simplicial cone and an algorithm to minimize this criterion.

Proposed Criterion : We define $V(W)$, the volume of a simplicial cone $Span^+(W)$ generated by a given square matrix $W = [w_1, w_2, \cdots, w_n]$, where w_i is the i-th column of W, by :

$$V(W) = \frac{|\det(W)|}{\|w_1\| \times \|w_2\| \times \cdots \times \|w_n\|} \tag{4}$$

$V(W)$ strictly represents the "aperture" of the simplicial cone $Span^+(W)$, it is positive and upper bounded by 1 (Hadamard's Inequality).

The task of estimating the mixing matrix can then be reduced to solving the following optimization problem:

$$W^* = \underset{W \geq 0, \ W^{-1}X \geq 0}{\arg\min} V(W) \tag{5}$$

Proposed Algorithm : We define the R_k^l like matrices by :

$$R_k^l = \begin{pmatrix} 1 & 0 & \cdots & 0 & r_{1k}^l & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & r_{2k}^l & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & r_{k-1k}^l & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & r_{k+1k}^l & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{nk}^l & 0 & \cdots & 1 \end{pmatrix} \text{ where } r_{ik}^l \geq 0, \ \forall \ 1 \leq i \leq n, i \neq k \tag{6}$$

Proposition 1. *For a given nonnegative matrix W, the volume $V(W)$ of $Span^+(W)$ decreases when W is multiplied to the right by R_k^l, $\forall\, l$ and $\forall\, 1 \leq k \leq n$.*

Proof. For fixed l, fixed k $(1 \leq k \leq n)$ and a given nonnegative square matrix $W = [w_1, w_2, \cdots, w_n]$, let $U = WR_k^l$ $(U = [u_1, u_2, \cdots, u_n])$. The task is to demonstrate that $V(U) \leq V(W)$:

1. $|\det(U)| = \left|\det(WR_k^l)\right| = |\det(W)| \left|\det(R_k^l)\right| = |\det(W)|$ because $\det(R_k^l) = 1$
2. $\forall\, j \neq k, u_j = w_j \implies \|u_j\| = \|w_j\|$
3. For $j = k$, $\forall\, 1 \leq i \leq n$, $u_{ik} = w_{ik} + \sum_{q=1, q \neq k}^{n} w_{iq} r_{qk}^l \geq w_{ik}$ because w_{iq} and r_{qk}^l are all nonnegative. Therefore $(u_{ik})^2 \geq (w_{ik})^2$ which leads to $\|u_k\| \geq \|w_k\|$

According to the definition of the volume of a simplicial cone given by (4), 1., 2. and 3. allow us to conclude that $V(U) \leq V(W)$

For solving problem (5), SCSA-UNS starts from an initial simplicial cone containing all the mixed data (typically, the positive orthant $Span^+(I_n)$) and iteratively decreases its volume by performing several sweeps of n right-multiplications of the matrix which generated this initial cone ($W_0 = I_n$) by the R_k^l matrices $(1 \leq k \leq n)$. At each iteration, the matrix R_k^l is computed so to keep the mixed data inside the new simplicial cone. Details for computing the R_k^l are given in Appendix 1. Let $W \geq 0$ be the matrix which generated the current simplicial cone (W is the current estimation of the mixing matrix and $Y = W^{-1}X \geq 0$ is the current estimation of the sources). The algorithm stops when one cannot decrease anymore $V(W)$ without creating negative values in Y. This often corresponds to the convergence of W to the true mixing matrix A. However, it may happen that $V(W)$ does not decrease anymore during the iterations while W has not converged yet to A. This freezing[1] situation arises when there is at least one zero value on each row of Y. To avoid this problem, we suggest applying, after each sweep l, an orthogonal linear transformation Q_l to Y (and Q_l^T to W) to delete the zeros values of Y without increasing $V(W)$. The details of computing the unfreezing matrices Q_l are given in Appendix 2.

The whole estimated mixing matrix and sources are given by:

$$A = W_0 \prod_l \left[\left(\prod_{k=1}^{n} R_k^l \right) Q_l^T \right] \quad \text{and} \quad S = A^{-1}X \tag{7}$$

3.2   Proposed Method for Overdetermined Case

When the number of observations is greater than the number of sources ($m > n$), we propose to first perform a dimension reduction of the mixed data and afterward run the SCSA-UNS algorithm on the reduced mixed data.

Dimension Reduction : The dimension reduction is performed by the classical Principle Component Analysis (PCA). We compute the Singular Value Decomposition (SVD) of the mixed data and we keep only the n largest singular values and the corresponding singular vectors.

---

[1] See Appendix 2 for more details.

$$X \approx E \ F \ G^T \atop m \times p \quad m \times nn \times nn \times p \tag{8}$$

The SCSA-UNS algorithm should be executed on G^T but G^T is not necessarily nonnegative, so its scatter plot is not necessarily contained in the positive orthant. One must then find an initial simplicial cone containing all the dimension reduced mixed data G^T, since $Span^+(I_n)$ is not necessarily suited any more.

Widening Procedure for Initialization : If all the observations are not nonnegative, we have developed a procedure for finding a convenient initial simplicial cone containing all the mixed for the SCSA-UNS algorithm. This procedure is based on widening $Span^+(I_n)$ by multiplying I_n by D_k^l matrices, with a structure similar to R_k^l, but with negative entries in order to increase the volume of $Span^+(I_n)$ up to enclose all the mixed data. The details of computation of the D_k^l are not given here due to lack of space. This widening procedure is used on G^T to compute W_1 and Y_1 so that:

$$G^T = W_1 \ Y_1 \quad \text{with} \ Y_1 \geq 0 \atop n \times p \quad n \times nn \times p \tag{9}$$

Estimating the Mixing Matrix and the Sources : The SCSA-UNS algorithm is finally executed on Y_1 to give $Y_1 = W_2 \ Y_2 \atop n \times p \quad n \times nn \times p$.
The whole mixing matrix and sources are estimated by:

$$A \approx E \ F \ W_1 \ W_2 \text{ and } S \approx Y_2 \atop m \times n \quad m \times nn \times n_n \times nn \times n \quad n \times p \quad n \times p \tag{10}$$

4   Simulations and Discussions

Case 1: Synthetic Data
The proposed method is first evaluated on synthetic data and compared with two other MV methods: MVSA [5] and MVES [6]. To make sure that the mixed data *fill enough* the simplicial cone generated by the mixing matrix, the nonnegative sources have been generated using "random sparse uniform distribution generator" with 64% of non-zero elements in the sources matrix. We consider two cases: full additive sources and non full additive sources. The mixing matrix has Gaussian random entries. We set $m = 20$, $n = 5$ and $p = 10000$. Comparison criteria are the CPU time to converge T and the separation error E_{sep} defined by (11). The smaller E_{sep}, the better the separation.

$$E_{sep} = \frac{1}{n(n-1)} [\sum_i (\sum_j \frac{\left|(W^{-1}A)_{ij}\right|}{\max_k \left|(W^{-1}A)_{ik}\right|} - 1) + \sum_j (\sum_i \frac{\left|(W^{-1}A)_{ij}\right|}{\max_k \left|(W^{-1}A)_{kj}\right|} - 1)] \tag{11}$$

Table 1 records the average performance indices for 50 independent Monte-Carlo runs. One may note that when the sources are full additive, the three methods perform a good separation but SCSA-UNS is faster than MVSA and MVES. However when the sources are not full additive, SCSA-UNS still performs a perfect separation while MVSA and MVES do not.

Table 1. Average performance indices

	Index	SCSA-UNS	MVSA	MVES
Full additive sources	E_{sep}	$3,56.10^{-10}$	$1,23.10^{-12}$	$2,67.10^{-8}$
	$T(s)$	0.69	$2,21$	$8,37$
Non full additive sources	E_{sep}	$1,54.10^{-9}$	$7,89$	$6,01$
	$T(s)$	$0,83$	$4,76$	$5,59$

Case 2: Real Dynamic Positon Emission Tomography (PET) Images
Simulations have also been performed on real Dynamic Positon Emission To-
mography (PET) data to study the pharmacokinetics of the [18F]-FDG (Fluo-
roDeoxyGlucose) tracer on human brain. The main objective is to estimate the
arterial pharmacokinetic also called Arterial Input Function (AIF) using only
the dynamic TEP images with no arterial blood sampling (rAIF) which is too
invasive for routine clinic use. The rAIF is only considered here as the reference
AIF to assess the proposed estimator accuracy. As a matter of fact, an accurate
estimation of the AIF allows a quantitative measurement which is indispensable
for an efficient treatment evaluation in oncology. We have 19 human brain PET
images recorded during $33mn$. Each 3D PET image is reshaped to form one row
of the observations matrix X. The number of observations is $m = 19$ and the
number of samples is $p = 266742$. We set the number of sources to $n = 4$.

Fig 1 shows the pharmacokinetics compartments estimated by the SCSA-UNS
algorithm. Every subfigure represents the normalized kinetics (estimated mixing
matrix) over the first four minutes (lower left) and the corresponding spatial

(a) Arterial

(b) Veinous

(c) Tissue
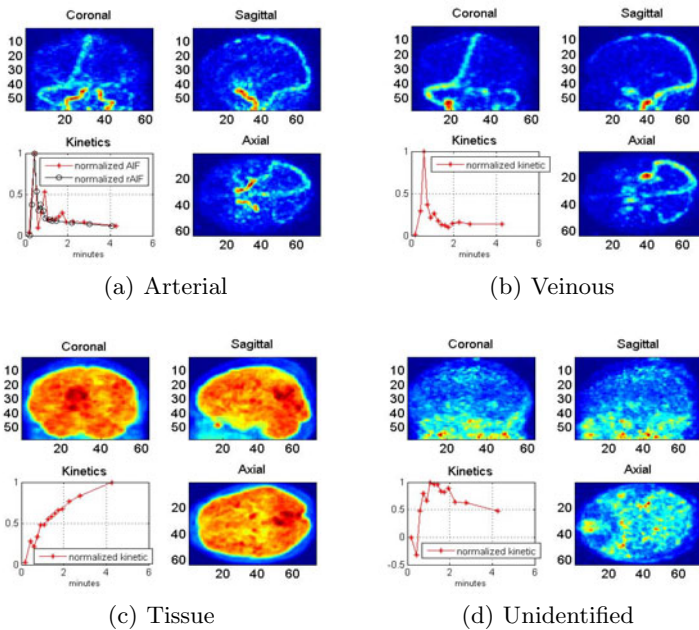
(d) Unidentified

Fig. 1. Estimated pharmacokinetics compartments

distributions (estimated sources) according to the three views coronal (upper left), sagittal (upper right) and axial (lower right). Three of the estimated compartments are identified to be the Arterial compartment (Fig 1.a), the Veinous one (Fig 1.b) and the Tissue one (Fig 1.c). Fig 1.a (lower left) shows that the normalized estimated AIF correctly approximates the normalized rAIF obtained by blood sampling, which was the main objective. However, one may note negative values on the kinetic of the unidentified compartment (lower left of Fig 1.d) which we attribute to measure noise.

5   Conclusions an Future Works

In this paper, we present a geometrical method for separating nonnegative sources. In overdetermined case, the proposed method first reduces the dimension of mixed data by performing the classical PCA and afterward runs the SCSA-UNS algorithm on the reduced data. The SCSA-UNS algorithm, estimates the mixing matrix by fitting a minimum volume simplicial cone to the scatter plot of observations. Unlike other geometrical methods that require local dominance or full additivity of the sources, SCSA-UNS only requires the mixed data to fill enough the simplicial cone generated by the mixing matrix. Simulations on synthetic data have showned that the proposed method always performs a good separation in noiseless case and runs faster than other MV methods. The proposed method also gave very promising results on real Dynamic PET images. As future works, we will investigate how to avoid the freezing situations without having to compute unfreezing matrices and we will consider the noisy case.

References

1. Puntonet, C., Mansour, A., Jutten, C.: A Geometrical Algorithm for Blind Source Separation. In: Gretsi 1995, Juan-les-Pins, pp. 273–276 (September 1995)
2. Babaie-Zadeh, M., Mansour, A., Jutten, C., Marvasti, F.: A Geometric Approach for Separating Several Speech Signals. In: Puntonet, C.G., Prieto, A.G. (eds.) ICA 2004. LNCS, vol. 3195, pp. 798–806. Springer, Heidelberg (2004)
3. Lazar, C., Nuzillard, D., Nowé, A.: A New Geometrical BSS Approach for Non Negative Sources. In: Vigneron, V., Zarzoso, V., Moreau, E., Gribonval, R., Vincent, E. (eds.) LVA/ICA 2010. LNCS, vol. 6365, pp. 530–537. Springer, Heidelberg (2010)
4. Chan, T.H., Ma, W.K., Chi, C.Y., Wang, Y.: A Convex Analysis Framework for Blind Separation of Non-Negative Sources. IEEE Transactions on Signal Processing 56, 5120–5134 (2008)
5. Li, J., Bioucas-Dias, J.M.: Minimum Volume Simplex Analysis: A Fast Algorithm to Unmix Hypersectral Data. In: IEEE International Symposium on Geoscience and Remote Sensing, vol. 3, pp. 250–253 (2008)
6. Chan, T.H., Chi, C.Y., Huang, Y.M., Ma, W.K.: A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing. IEEE Transactions on Signal Processing 57, 4418–4432 (2009)
7. Craig, M.D.: Minimum-Volume Transforms for Remotely Sensed Data. IEEE Transactions on Geoscience and Remote Sensing 32, 542–552 (1994)

8. Henry, R.C.: History and fundamentals of multivariate air quality receptor models. Chemometrics and Intelligent Laboratory Systems 37, 37–42 (1997)
9. Ouedraogo, W.S.B., Souloumiac, A., Jaidane, M., Jutten, C.: Simplicial Cone Shrinking Algorithm for Unmixing Nonnegative Sources. In: Accepted to ICASSP 2012 (2012)

Appendix 1: Computing R_k^l

$W \geq 0$ and $Y = W^{-1}X \geq 0$ are the current estimations of A and S at iteration $k - 1$ of sweep l, respectively. At iteration k, we look for a matrix R_k^l, so that $U = W R_k^l$ verify the following three conditions:

i) $U \geq 0$

ii) $V(U) \leq V(W)$

iii) $Z = U^{-1}X \geq 0$ $\left(Z = U^{-1}X = (R_k^l)^{-1}W^{-1}X = (R_k^l)^{-1}Y \right)$

The first two conditions are automatically satisfied because W and R_k^l are nonnegative and due to Proposition 1. From the definition of R_k^l given by (6), one may demonstrate that $[Z]_{ij} = [Y]_{ij} - r_{ik}^l[Y]_{kj}, \forall \, 1 \leq i \leq n, \forall \, 1 \leq j \leq p$. For fixed i, $[Z]_{ij} \geq 0 \Leftrightarrow r_{ik}^l \leq \frac{[Y]_{ij}}{[Y]_{kj}}, \forall \, 1 \leq j \leq p$.

A convenient R_k^l matrix can then be computed by taking:

$$r_{kk}^l = 1 \text{ and } r_{ik}^l = \min_{1 \leq j \leq p} \frac{[Y]_{ij}}{[Y]_{kj}}, [Y]_{kj} \neq 0, \text{ for } i \neq k \tag{12}$$

Appendix 2: Computing the Unfreezing Matrix Q_l

Before giving details of computation of the unfreezing matrix Q_l, lets explain how arises the freezing situation. Given $W \geq 0$ and $Y = W^{-1}X \geq 0$ the current estimated mixing matrix sources respectively and $U = W R_k^l$:

$V(U) = V(W) \Leftrightarrow R_k^l = I_n \Longleftrightarrow r_{ik}^l = \delta_{ik}$. For $i \neq k$ and according to (12), $r_{ik}^l = 0 \Longleftrightarrow \exists \, 1 \leq j \leq p$ so $[Y]_{ij} = 0$ and $[Y]_{kj} \neq 0$.

The freezing arises if, at sweep l, the algorithm finds $R_k^l = I_n, \forall \, 1 \leq k \leq n$. This situation happens when there are at least one zero value on each row of the current estimated sources matrix (i.e. when $\forall \, 1 \leq i \leq n, \exists \, 1 \leq j \leq p, [Y]_{ij} = 0$). To avoid this problem, we suggest applying an orthogonal linear transformation Q_l to Y (and Q_l^T to W) to delete the zeros values of Y without increasing $V(W)$. We then introduce the unfreezing matrix Q_l so that $X = W Q_l^T Q_l Y = HT$. The current estimated mixing matrix and sources become $H = W Q_l^T$ and $T = Q_l Y$. We look for a matrix Q_l so that $Q_l^T Q_l = I_n$ and $T > 0$.

For findind such a Q_l matrix, we define the criterion J by:

$$J(Q) = \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{T_{ij}} \text{ where } 1_{T_{ij}} = \begin{cases} 1 \text{ if } T_{ij} = 0 \\ 0 \text{ elsewhere} \end{cases} \tag{13}$$

The unfreezing matrix Q_l can be computed by solving the following equation:

$$Q_l = \underset{Q^T Q = I_n}{\arg\min} \, J(Q) \tag{14}$$

We developed a regularized gradient algorithm for solving the optimization problem (14) [9] but this method is not described here due to lack of space.