A geometrical blind separation method for unconstrained-sum locally dominant sources

Axel Boulais, Yannick Deville and Olivier Berné

Institut de Recherche en Astrophysique et Planétologie (IRAP)
Toulouse University, UPS-OMP, CNRS
Toulouse France

SFPT-GH - May 11-13, 2016
Summary

1. Blind Source Separation
2. Geometrical framework
3. MASS
4. Hyperspectral unmixing in astrophysics
5. Experimental results
Blind Source Separation

- A generic signal processing problem

![Blind Source Separation Diagram](image)

Figure: Blind source separation problem (L sources and M observations (M ≥ L)).

- Different mixing operators: linear instantaneous, anechoic, convolutive, non-linear
- Different signals:
  - 1D (audio, communication, spectroscopy...)
  - 2D (images)
Linear Blind Source Separation

- **Instantaneous linear mixing:** \( x(m, n) = \sum_{\ell=1}^{L} a(m, \ell) \times s(\ell, n) \)

\[
\begin{pmatrix}
  x_1 \\
  \vdots \\
  x_M
\end{pmatrix}
\begin{pmatrix}
  a_{1,1} & \cdots & a_{1,L} \\
  \vdots & \ddots & \vdots \\
  a_{M,1} & \cdots & a_{M,L}
\end{pmatrix}
\begin{pmatrix}
  s_1 \\
  \vdots \\
  s_L
\end{pmatrix}
\]

- **3 classes of methods to solve the linear problem:**
  - Independent Component Analysis (ICA)
  - Non-Negative Matrix Factorization (NMF)
  - Sparse Component Analysis (SCA)

- **Subclass of methods:** NMF + SCA \( \rightarrow \) geometric methods
Linear Blind Source Separation

- **Single-source observations:**

\[
X \ (M \times N) = A \ (M \times L) \times S \ (L \times N)
\]

- **Sparsity assumption:** for each source, there exist at least one sample index \( n \) of observations for which these source is non-zero.
Sum-to-one constraint:

In many geometric methods (hyperspectral image unmixing in Earth observation):

$$\sum_{\ell=1}^{L} a(m, \ell) = 1 \quad \forall m \in \{1, \ldots, M\}$$  \hspace{1cm} (1)
Sum-to-one constraint:

- For our applications:

\[
\sum_{\ell=1}^{L} a(m, \ell) = 1 \quad \forall m \in \{1, \ldots, M\}
\]  

Assumptions:

- \(X, A\) and \(S\) are non-negative
- \(A\) is a full column rank matrix
- The number of sources \(L\) is known
- Sparsity assumption of sources (presence of single source observations)
- The sum of mixing coefficients is unconstrained
Geometrical framework

- Each observed vector $x_m$ is represented as an element of a $\mathbb{R}^N$ vector space.

$$x_m = As_m$$  \hspace{1cm} (2)

- $A$ and $S$ being non-negative and the column $a_\ell$ being linearly independent:

- The set:

$$C_A = \{x_m \mid x_m = As_m, \ s_m \in \mathbb{R}_+^L\}$$  \hspace{1cm} (3)

  is a simplicial cone (the convex hull spanned by the non negative linear combination of the columns of $A$).

- Each column vector $a_\ell$ of $A$ spans an edge $E_\ell$ of the simplicial cone $C_A$:

$$E_\ell = \{c \mid c = \alpha a_\ell, \ \alpha \in \mathbb{R}_+\}$$  \hspace{1cm} (4)
Geometrical framework

- 2-dimension example:

![Figure: Scatter plot of mixed data and edges of the simplicial cone.](image)

**Figure**: Scatter plot of mixed data and edges of the simplicial cone.

- If the sparsity assumption of sources is valid:
  
  \[ C_A = C_X \]  

- Identifying the column of \( A \) → finding the columns of \( X \) which are **furthest apart in the angular sense**.
Maximum Angle Source Separation (MASS)

- **Estimation of mixing matrix $A$:** L-1 steps
  - Columns of $X$ are normalize to unit length.
  - Illustration of the algorithm in 3-dimension with a mixture of 3 sources:

![Scatter plot of the observed data.](image)

**Figure:** Scatter plot of the observed data.
Identify the first 2 columns of $\hat{A}$ by selecting the two columns of $X$ that have the largest angle:

$$(m_1, m_2) = \arg\max_{i,j} \cos^{-1}(x_i^T x_j) \quad \forall i, j \in \{1, \ldots, M\}. \quad (6)$$

$$(m_1, m_2) = \arg\min_{i,j} x_i^T x_j \quad \forall i, j \in \{1, \ldots, M\}. \quad (7)$$

$$\tilde{A} = [x_{m_1}, x_{m_2}] \quad (8)$$
Identify the column which has the largest angle with $x_{m_1}$ and $x_{m_2}$:

- maximum angle between the column and its orthogonal projection on the simplicial cone spanned by the columns of $\tilde{A}$:

$$\Pi_{\tilde{A}}(X) = \tilde{A}(\tilde{A}^T \tilde{A})^{-1} \tilde{A}^T X.$$  \hspace{1cm} (9)

$$m_3 = \arg\min_i x_i^T \pi_i \quad \forall i \in \{1, \ldots, M\}$$  \hspace{1cm} (10)

$$\tilde{A} = [x_{m_1}, x_{m_2}, x_{m_3}]$$  \hspace{1cm} (11)
This projection and identification procedure is then repeated to identify the $L$ columns of the mixing matrix.

\[ \hat{A} = [x_{m1}, \ldots, x_{mL}] \] (12)

Source matrix reconstruction:

$X$, $A$ and $S$ are non-negative $\rightarrow$ Non-Negative Least Square algorithm (NNLS):

\[ J(\hat{s}_m) = \frac{1}{2} \| x_m - \hat{A}\hat{s}_m \|_2^2 \quad s.t. \; \hat{s}_m \geq 0, \; \forall m \in \{1, \ldots, M\} \] (13)
Hyperspectral unmixing in astrophysics

- Application field of MASS (non-negativity, correlated sources without sum-to-one constraint) is very common in astrophysics.
- Area observed at high spectral resolution → each pixel corresponds to an emission spectrum of a portion of the area.

Figure: NGC7023

Problem: The observed spectra are generally constituted by a mixture of elementary spectra (components of the gas cloud containing different chemical species).
Experimental results

(a) Source 1  (b) Source 2  (c) Source 3  (d) Source 4

Figure: Extracted spectra in NGC7023-NW: NMF $\rightarrow$ blue spectra, MASS $\rightarrow$ red spectra.

- Similar results for the extracted spectra
- Very fast algorithm compared to the NMF
- Uniqueness of the solution
- MASS is able to identify a weak signal present in only some observation