Non-stationary Markovian Blind Source Separation

Solving Estimating Equations using an Equivariant Newton-Raphson Algorithm

Technical Report

Rima Guidara, Shahram Hosseini and Yannick Deville *

September 19, 2008

In this document, we present a method for solving estimating equations in a non-stationary Markovian blind source separation (BSS) context. The source and observation vectors are denoted, respectively, \( s(t) = [s_1(t), \ldots, s_K(t)]^T \) and \( x(t) = [x_1(t), \ldots, x_K(t)]^T \), and we denote the probability density function of a source \( s_i \) at time \( t \) by \( f_{s_i}(t) \).

Considering a linear instantaneous mixture model, our aim is to find a separating matrix \( B \) which is an estimate of the inverse of the mixing matrix up to classical BSS indeterminacies. To this end, we apply a maximum likelihood approach, where sources are supposed to be non-stationary \( q^h \)-order Markovian processes. Following the same steps as in [1], we finally obtain the following set of equations

\[
E_{N-q} \left[ \sum_{l=0}^{q} \psi_{s_i(t)}^l (s_i(t)|s_i(t-1), \ldots, s_i(t-q)) s_j(t-l) \right] = 0, \quad i \neq j = 1, \ldots, K
\]  

(1)

where \( E_{N-q} = \frac{1}{N-q} \sum_{t=q+1}^{N} \) is a temporal mean and \( \psi_{s_i(t)}^l (\cdot|\cdot) \) is the conditional score function of a source \( s_i \) at time \( t \) with respect to the source sample \( s_i(t-l) \), defined by

\[
\psi_{s_i(t)}^l (s_i(t)|s_i(t-1), \ldots, s_i(t-q)) = \frac{-\partial \log f_{s_i(t)}(s_i(t)|s_i(t-1), \ldots, s_i(t-q))}{\partial s_i(t-l)}, \quad \forall \quad 0 \leq l \leq q.
\]  

(2)

We here propose to solve equations (1) using an equivariant Newton-Raphson algorithm. The equivariance in BSS algorithms was defined in [2]. To simplify notations, we restrict our calculus to the case \( K = 2 \). However, extending the above results to more than 2 sources is straightforward.

Denoting \( \hat{B} \) the estimate of the separating matrix \( B \) at the current algorithm iteration, the new estimate \( \hat{B} \) is obtained by the updating formula \( \hat{B} = (I + \Delta) \hat{B} \). Post-multiplying this equation by the observation vector \( x \), the new source estimate can be written as \( \hat{s} = (I + \Delta) s \).

Denoting \( \Delta = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \), the above equation reads as follows for \( K = 2 \)

\[
\hat{s}_i(t) = \tilde{s}_i(t) + \delta_{ii} \tilde{s}_i(t) + \delta_{ij} \tilde{s}_j(t), \quad i \neq j = 1, 2
\]  

(3)

*Authors are with the Laboratory of Astrophysics Toulouse-Tarbes, University of Toulouse, CNRS, 14 av. Edouard Belin, 31400 Toulouse, France, \{contact: rguidara@ast.obs-mip.fr\}. 
The diagonal entries of Δ may be set to any small arbitrary value, due to scaling indeterminacy in the BSS problem. The off-diagonal terms are computed as follows.

To be independent, the estimated sources should satisfy the \( K(K - 1) \) estimating equations (1). Replacing the sources in (1) by their expressions (3), when \( K = 2 \), we obtain the following equations

\[
E_{N-q}\left[ \sum_{l=0}^{q} \psi_{s_{i}(t)}^{j}(\tilde{s}_{i}(t) + \delta_{ii}\tilde{s}_{i}(t) + \delta_{ij}\tilde{s}_{j}(t)|\tilde{s}_{i}(t - 1) + \delta_{ii}\tilde{s}_{i}(t - 1) + \delta_{ij}\tilde{s}_{j}(t - 1), \ldots, \tilde{s}_{i}(t - q) + \delta_{ii}\tilde{s}_{i}(t - q) + \delta_{ij}\tilde{s}_{j}(t - q)|\tilde{s}_{j}(t - l) + \delta_{ji}\tilde{s}_{i}(t - l) + \delta_{jj}\tilde{s}_{j}(t - l) \right] = 0
\]

\[ i \neq j = 1, 2 \quad (4) \]

Using a first-order Taylor expansion of the score function \( \psi_{s_{i}(t)}^{j} \), at the estimate \( \tilde{s}_{i}(t) \), the above equation can be written as

\[
E_{N-q}\left[ \sum_{l=0}^{q} \left[ \psi_{s_{i}(t)}^{j}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q)) + \sum_{n=0}^{q} \frac{\partial \psi_{s_{i}(t)}^{j}}{\partial \tilde{s}_{i}(t - n)}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q)) \right] \left( \delta_{ii}\tilde{s}_{i}(t - n) + \delta_{ij}\tilde{s}_{j}(t - n) \right) \left( \tilde{s}_{j}(t - l) + \delta_{ji}\tilde{s}_{i}(t - l) + \delta_{jj}\tilde{s}_{j}(t - l) \right) \right] = 0, \quad i \neq j = 1, 2 \quad (5) \]

If we neglect second-order terms in the above equation, we obtain a linear equation with respect to the entries of the matrix Δ, that reads

\[
(1 + \delta_{jj})J_{1} + \delta_{ij}J_{2} + \delta_{ii}J_{3} + \delta_{ij}J_{4} = 0, \quad i \neq j = 1, 2 \quad (6) \]

with

\[
J_{1} = E_{N-q}\left[ \sum_{l=0}^{q} \psi_{s_{i}(t)}^{j}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q))\tilde{s}_{j}(t - l) \right]
\]

\[
J_{2} = E_{N-q}\left[ \sum_{l=0}^{q} \psi_{s_{i}(t)}^{j}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q))\tilde{s}_{i}(t - l) \right]
\]

\[
J_{3} = E_{N-q}\left[ \sum_{l=0}^{q} \left\{ \sum_{n=0}^{q} \frac{\partial \psi_{s_{i}(t)}^{j}}{\partial \tilde{s}_{i}(t - n)}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q))\tilde{s}_{i}(t - n) \right\} \tilde{s}_{j}(t - l) \right]
\]

\[
J_{4} = E_{N-q}\left[ \sum_{l=0}^{q} \left\{ \sum_{n=0}^{q} \frac{\partial \psi_{s_{i}(t)}^{j}}{\partial \tilde{s}_{i}(t - n)}(\tilde{s}_{i}(t)|\tilde{s}_{i}(t - 1), \ldots, \tilde{s}_{i}(t - q))\tilde{s}_{i}(t - n) \right\} \tilde{s}_{j}(t - l) \right]
\]

We neglect \( \delta_{jj} \) with respect to 1 in Eq. (6). In the vicinity of the solution, the estimated sources may be assumed to be nearly independent and centered, so that for any function \( \Phi, E\left[ \Phi(\tilde{s}_{i}(t - l)) \cdot \tilde{s}_{j}(t - n) \right] \simeq E\left[ \Phi(\tilde{s}_{i}(n - l)) \right] \cdot E\left[ \tilde{s}_{j}(t - n) \right] \) is small, which means that \( \delta_{ii}J_{3} \) is negligible with respect to the other terms in (6).
These simplifications finally yield a linear set of equations defined by

\[
E_{N-q} \left[ \sum_{l=0}^{q} \psi_{s_i(t)}^{l}(\tilde{s}_i(t)|\tilde{s}_i(t-1), \ldots, \tilde{s}_i(t-q), \tilde{s}_i(t-l)) \delta_{ji} \right]
\]

\[
+ E_{N-q} \left[ \sum_{l=0}^{q} \left\{ \sum_{n=0}^{q} \frac{\partial \psi_{s_i(t)}^{l}}{\partial \tilde{s}_i(t-n)}(\tilde{s}_i(t)|\tilde{s}_i(t-1), \ldots, \tilde{s}_i(t-q), \tilde{s}_j(t-n)), \tilde{s}_j(t-l) \right\} \delta_{ij} \right]
\]

\[
= -E_{N-q} \left[ \sum_{l=0}^{q} \psi_{s_i(t)}^{l}(\tilde{s}_i(t)|\tilde{s}_i(t-1), \ldots, \tilde{s}_i(t-q)), \tilde{s}_j(t-l) \right], \quad i \neq j = 1, 2
\]

References
