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, $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^T$ and $\mathbf{x}(t) = [x_1(t), \ldots, x_K(t)]^T$, 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 $\mathbf{B}$ 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^{th}-order Markovian processes. Following the same steps as in [1], we finally obtain the following set of equations

$$E_{N-q}\Big[\sum_{l=0}^{q} \psi^l_{s_i(t)}(s_i(t)|s_i(t-1), \ldots, s_i(t-q))s_j(t-l)\Big] = 0, \quad i \neq j = 1, \ldots, K \tag{1}$$

where $E_{N-q} = \frac{1}{N-q}\sum_{t=q+1}^{N}$ is a temporal mean and $\psi^l_{s_i(t)}(.|.)$ 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

$$\psi^l_{s_i(t)}(s_i(t)|s_i(t-1), \ldots, s_i(t-q)) = \frac{-\partial \log f_{s_i(t)}(s_i(t)|s_i(t-1), \ldots, s_i(t-q))}{\partial s_i(t-l)}, \quad \forall \quad 0 \leq l \leq q. \tag{2}$$

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 $\tilde{\mathbf{B}}$ the estimate of the separating matrix $\mathbf{B}$ at the current algorithm iteration, the new estimate $\hat{\mathbf{B}}$ is obtained by the updating formula $\hat{\mathbf{B}} = (\mathbf{I} + \mathbf{\Delta})\tilde{\mathbf{B}}$. Post-multiplying this equation by the observation vector $\mathbf{x}$, the new source estimate can be written as $\hat{\mathbf{s}} = (\mathbf{I} + \mathbf{\Delta})\tilde{\mathbf{s}}$.
Denoting $\mathbf{\Delta} = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix}$, the above equation reads as follows for $K = 2$

$$\widehat{s}_i(t) = \widetilde{s}_i(t) + \delta_{ii}\widetilde{s}_i(t) + \delta_{ij}\widetilde{s}_j(t), \quad i \neq j = 1, 2 \tag{3}$$

*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 $\boldsymbol{\Delta}$ 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

$$E_{N-q}\Big[\sum_{l=0}^{q}\psi^{l}_{s_i(t)}(\tilde{s}_i(t)+\delta_{ii}\tilde{s}_i(t)+\delta_{ij}\tilde{s}_j(t)|\tilde{s}_i(t-1)+\delta_{ii}\tilde{s}_i(t-1)+\delta_{ij}\tilde{s}_j(t-1),\dots$$
$$,\tilde{s}_i(t-q)+\delta_{ii}\tilde{s}_i(t-q)+\delta_{ij}\tilde{s}_j(t-q))\{\tilde{s}_j(t-l)+\delta_{ji}\tilde{s}_i(t-l)+\delta_{jj}\tilde{s}_j(t-l)\}\Big]=0$$
$$i\neq j=1,2 \quad (4)$$

Using a first-order Taylor expansion of the score function $\psi^{l}_{s_i(t)}$, at the estimate $\tilde{s}_i(t)$, the above equation can be written as

$$E_{N-q}\Big[\sum_{l=0}^{q}\big[\psi^{l}_{s_i(t)}(\tilde{s}_i(t)|\tilde{s}_i(t-1),\dots,\tilde{s}_i(t-q))+\sum_{n=0}^{q}\frac{\partial\psi^{l}_{s_i(t)}}{\partial s_i(t-n)}(\tilde{s}_i(t)|\tilde{s}_i(t-1),\dots,\tilde{s}_i(t-q))$$
$$(\delta_{ii}\tilde{s}_i(t-n)+\delta_{ij}\tilde{s}_j(t-n))\big].\{\tilde{s}_j(t-l)+\delta_{ji}\tilde{s}_i(t-l)+\delta_{jj}\tilde{s}_j(t-l)\}\Big]=0,\quad i\neq j=1,2 \quad (5)$$

If we neglect second-order terms in the above equation, we obtain a linear equation with respect to the entries of the matrix $\boldsymbol{\Delta}$, that reads

$$(1+\delta_{jj})\mathbf{J}_1+\delta_{ji}\mathbf{J}_2+\delta_{ii}\mathbf{J}_3+\delta_{ij}\mathbf{J}_4=0,\quad i\neq j=1,2 \quad (6)$$

with

$$\mathbf{J}_1 = E_{N-q}\Big[\sum_{l=0}^{q}\psi^{l}_{s_i(t)}\big(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)\big)\widetilde{s}_j(t-l)\Big]$$
$$\mathbf{J}_2 = E_{N-q}\Big[\sum_{l=0}^{q}\psi^{l}_{s_i(t)}\big(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)\big)\widetilde{s}_i(t-l)\Big]$$
$$\mathbf{J}_3 = E_{N-q}\Big[\sum_{l=0}^{q}\Big\{\sum_{n=0}^{q}\frac{\partial\psi^{l}_{s_i(t)}}{\partial s_i(t-n)}\big(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)\big)\widetilde{s}_i(t-n)\Big\}\widetilde{s}_j(t-l)\Big]$$
$$\mathbf{J}_4 = E_{N-q}\Big[\sum_{l=0}^{q}\Big\{\sum_{n=0}^{q}\frac{\partial\psi^{l}_{s_i(t)}}{\partial s_i(t-n)}\big(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)\big)\widetilde{s}_j(t-n)\Big\}\widetilde{s}_j(t-l)\Big]$$

We neglect δ_{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 Φ, $E\Big[\Phi(\widetilde{s}_i(t-l))\cdot\widetilde{s}_j(t-n)\Big]\simeq E\Big[\Phi(\widetilde{s}_i(n-l))\Big]\cdot E\Big[\widetilde{s}_j(t-n)\Big]$ is small, which means that $\delta_{ii}\mathbf{J}_3$ is negligible with respect to the other terms in (6).

These simplifications finally yield a linear set of equations defined by

$$
\begin{aligned}
& E_{N-q}\Big[\sum_{l=0}^{q} \psi_{s_i(t)}^{l}(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)).\tilde{s}_i(t-l)\Big]\delta_{ji} \\
&\quad + E_{N-q}\Big[\sum_{l=0}^{q}\Big\{\sum_{n=0}^{q} \frac{\partial \psi_{s_i(t)}^{l}}{\partial \tilde{s}_i(t-n)}(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q))\tilde{s}_j(t-n)\Big\}.\tilde{s}_j(t-l)\Big]\delta_{ij} \\
&\quad\quad = -E_{N-q}\Big[\sum_{l=0}^{q} \psi_{s_i(t)}^{l}(\widetilde{s}_i(t)|\widetilde{s}_i(t-1),\dots,\widetilde{s}_i(t-q)).\tilde{s}_j(t-l)\Big], \quad i \neq j = 1,2
\end{aligned}
$$

References

[1] S. Hosseini, C. Jutten and D. T. Pham, Markovian Source Separation, *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3009-3019, 2003.

[2] J. F. Cardoso and B. Laheld, Equivariant adaptive source separation, *IEEE Transactions on Signal Processing*, vol. 44, no. 12, pp. 3017-3030, 1996.