ESANN'2006 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2006, d-side publi., ISBN 2-930307-06-4.

Source separation with priors on the power spectrum of the sources

Jorge Igual[1], Raul Llinares[1] and Andres Camacho[1] *

1- Polytechnic University of Valencia - Dept of Communications, GATACA
Pza. Ferrandiz y Carbonell, 1, C.P. 03804, Alcoy - SPAIN

Abstract. A general approach introducing priors on the correlation function or equivalently power spectrum of the sources in the Blind Source Separation problem is presented. This prior modifies or constrains the contrast function that measures the independence of the recovered signals depending on its characteristics. Considering the case where the priors correspond to the sources that we are interested in recovering, the deflation approach is stated. This formulation is especially useful for those large-dimension problems where the ancillary sources are not needed to be estimated. We show its application to the biomedical problem of extracting the atrial activity from atrial fibrillation episodes, where discriminant information about the frequency content of the atrial activity with respect to the other components is available in advance.

1 Introduction

Blind Source Separation BSS consists of recovering the source signals from the observations obtained by mixtures of them. It is called blind because of nothing is assumed about the sources or the coefficients of the mixture but the statistical independence of the sources. However some assumptions are implicit in the model, such as: the sort of mixture (linear or non linear, instantaneous or convolutive, noisy or noise free), the dimensions of the problem (the size of the mixing matrix, i.e., the number of sources), at most one Gaussian source for algorithms based only on higher order statistics, non identical power spectra for algorithms based on second order statistics, i.e., exploiting the time structure, or the usual assumption of zero mean unit variance sources for simplification and fixing some indeterminacies.

In this paper we will follow the linear noise-free instantaneous mixing model:

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

where $\mathbf{x}(t) = \left[x_1(t), \ldots x_m(t)\right]^T$ is the vector of observed signals, i.e., the mixtures, $\mathbf{s}(t) = \left[s_1(t), \ldots, s_n(t)\right]^T$ the source vector and $\mathbf{A}_{mxn}$ the mixing matrix. The aim is recovering the sources from the only assumption of the statistical independence of the sources $p(\mathbf{s}) = \prod_{i=1}^{n} p(s_i)$:

$$\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t) \tag{2}$$

* This work is partially funded by Polytechnic University of Valencia and GVA. We also acknowledge BET group for the signals used in the results Section.

where $\mathbf{y}(t) = [y_1(t), \ldots, y_n(t)]^T$ are the recovered signals and **B** the demixing matrix.

When the mixing process (1) is not explicit, we call (2) the Independent Component Analysis ICA of $\mathbf{x}(t)$. In the square case $m = n$, the recovered sources are the original ones up to a permutation, sign and amplitude indeterminacy, $\mathbf{y} = \mathbf{Ps}$, where **P** is a matrix that has one and only one nonzero entry in each row and column.

When additional knowledge is included in the model, the problem is no more called *blind* and sometimes some restrictions can be relaxed, such as $m \geq n$. When information about the form of the densities is available, new approaches and *ad hoc* adaptations of classic algorithms can be obtained, e.g., when the sources are sparse [2]. We must note that some gradient based algorithms use component wise nonlinearities that include implicit priors about the pdf and the kurtosis of the sources, such as the Infomax algorithm [1]. But in this case we do not call them "priors"; if the algorithm fails, we simply say that the nonlinear function is not properly selected instead of talking about an incorrect prior. In fact, the sparseness is considered a prior in the literature but not the sign of the kurtosis in the selection of the nonlinearity, just because the former is formulated explicitly in the statement of the problem. Other usual source priors include its temporal structure [3] or the relaxation of the independence hypothesis [4]. Concerning the mixing matrix, priors can be modeled in a Bayesian approach [5], [6], or as constraints, being assumed some of the entries, e.g., due to the available information about the positions of sensors and sources, or being parameterized like in array signal processing.

We focus in this paper in the case where there is prior information about the power spectrum of some sources, at least one. We find many real applications such as in communications or biomedicine where this knowledge is available but not used by BSS algorithms. In addition, in many of these applications, we are interested in recovering only some few sources, so an algorithm that first extracts the interesting sources imposing the prior knowledge is necessary.

In Section 2, a general approach including the information about the frequency content of the sources is set. Depending on the prior, it includes the modification of the objective function that measures the independence or the restriction of the possible solutions. In Section 3, we present a biomedical example where the prior consists on that the power of the interesting source is concentrated in some frequencies, showing the results in Section 4.

2 Priors on the power spectrum of the sources

In basic ICA, solving (2) requires the use of higher order statistics in a direct way, such as cumulants, or in a more subtle way such as nonlinear functions. This is because only the decorrelation of the observations is not enough, remaining an orthogonal separating matrix for being estimated.

BSS algorithms are usually carried out in two steps. First, a PCA stage, consisting of whitening spatially the observations and reducing the dimension of the problem, $\mathbf{z}(t) = \mathbf{Vx}(t)$, where **V** is the *nxm* whitening matrix, so $E\{\mathbf{z}(t)\mathbf{z}(t)^T\} = \mathbf{I}$. Second, the remaining *nxn* orthogonal matrix **W**, $\mathbf{s}(t) = \mathbf{Wz}(t)$, with $\mathbf{WW}^T = \mathbf{I}$, is