ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

A Competitive Learning Algorithm for Separating Binary Sources

Petteri Pajunen

Helsinki University of Technology
Laboratory of Computer and Information Science
Rakentajanaukio 2 C, FIN-02150 Espoo, FINLAND

Abstract. A neural algorithm for separating binary sources from their linear unknown mixtures is presented. The a priori knowledge of binary sources is utilized by using competitive learning. With the algorithm it is possible to handle the difficult case of separating more sources than sensors. When the mixtures are noisy, it is possible to recover the sources exactly when the noise level is low.

1. Introduction

Most source separation algorithms are derived under the following assumptions:

- The sources are mutually statistically independent
- There are at least as many sensor as there are sources
- At most one source is Gaussian
- The mixing matrix is non-singular

However, utilizing a priori knowledge when it is available we can often drop some of the above assumptions [4]. In this paper, we assume that the sources are binary and that their linear combinations are different. These assumptions hold for example in the non-fading CDMA model [3]. Since there are only finitely many different binary source vectors, there are also finitely many mixtures. We require that all these mixtures are different. Note that this does not require that the columns of the mixing matrix are linearly independent. It follows that it is possible to separate more sources than sensors.

We briefly review some previous work where separation of binary or n-valued sources is considered. In [6] a geometrical approach was given for separating n-valued sources from linear mixtures. However, this work concentrated on the case when there are as many sources as sensors. In [2] the self-organizing map was applied to the less sensors than sources case. The results were not completely satisfactory, which is due to not fully taking advantage of the linear mixing model. In [1] the EM-algorithm was used to separate n-valued sources from less sensors than sources.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

The algorithm presented in this paper allows separation of binary sources from less sensors than sources. Both the estimation of the basis vectors and the separation of sources utilize competitive learning.

2. Separation of Binary Sources

For simplicity, let us assume that the binary source signals take values in $\{-1, 1\}$. Writing the ICA model (see e.g. [5]) as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s} = \sum_{i=1}^{m} s_i(t)\mathbf{a}_i, \qquad (1)$$

where $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m]$ and the $s_i(t)$ are the binary source signals, we see that each $\mathbf{x}(t)$ is a binary linear combination of the real-valued ICA basis vectors $\mathbf{a}_i$. There are 2^m different binary source vectors $\mathbf{s}(t) = [s_1(t), \ldots, s_m(t)]^T$ and therefore at most 2^m different vectors $\mathbf{x}(t)$. If all the possible vectors $\mathbf{x}(t)$ are different, then it is possible in theory to separate the sources since the mapping from the sources to the mixtures is one-to-one, hence reversible. To find the required conditions on matrix $\mathbf{A}$, consider the case when two mixtures are the same:

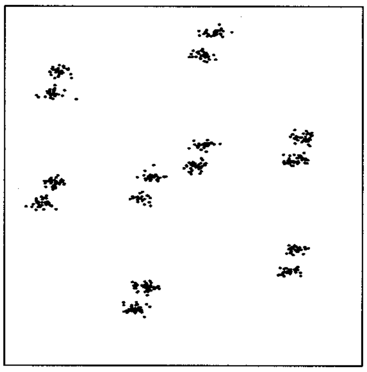

Figure 1: Two noisy mixtures of four binary sources are illustrated. The $2^4 = 16$ different clusters are clearly visible.

$$\begin{aligned} \mathbf{A}\mathbf{s}_1 = \mathbf{A}\mathbf{s}_2 &\iff \\ \mathbf{A}(\mathbf{s}_1 - \mathbf{s}_2) = \mathbf{0} &\iff \\ \mathbf{A}\mathbf{v} = \mathbf{0}, &\ \forall i\ v_i \in \{-2, 0, 2\} \end{aligned}$$

The last line shows that any binary sum of any nonempty subset of $\{\mathbf{a}_i\}$ must be nonzero.

It remains to label each observed mixture vector with a corresponding source vector. Note that to actually observe all the possible mixture vectors $\mathbf{x}$, we need to require a sort of independence of the source signals.

We can also allow additive zero-mean noise in the observed mixture vectors. When the noise level is low, this results in a set of clusters in the mixture space (Fig. 1). The algorithm for labeling the observed clusters is a competitive learning algorithm for updating the estimates of the basis vectors $\mathbf{a}_i$. The idea behind the algorithm follows from the cluster-like structure of the mixture space. Each binary linear combination of the estimates of the basis vectors $\mathbf{a}_i$ should correspond to one of the clusters. For each sample mixture $\mathbf{x}(t)$ a best-matching unit is searched among all the linear combinations of basis vectors. The winner is updated towards the sample mixture and the update is divided among all the basis vectors. The estimated parameters are the basis vectors $\mathbf{a}_i$.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

One iteration of the learning rule is as follows:

1. Obtain a mixture vector $\mathbf{x}(t)$.
2. Compute the best-matching unit, i.e. find such numbers $b_i \in \{-1, 1\}$ that the distance $||\mathbf{x}(t) - \sum_{i=1}^{m} b_i \mathbf{a}_i||$ is minimized.
3. Update the basis vectors by $\Delta \mathbf{a}_i = \mu(t)(b_i/m)\mathbf{e}$ where $\mathbf{e} = \mathbf{x}(t) - \sum_i b_i \mathbf{a}_i$

To avoid the possible dead units, i.e. vectors that never win, a conscience learning scheme is introduced. This can be implemented by multiplying the learning rate by $k(b_1, \ldots, b_m)$, the number of iterations passed since the best-matching unit had coefficients $b_1, \ldots, b_m$, yielding $\mu'(t) = \mu(t) * 2^{-m} * k(b_1, \ldots, b_m)$ as the new learning rate. The multiplier 2^{-m} is added for scaling. For stability reasons this should be limited to a constant less than one by $\mu'(t) = \min[\mu(t), c]$ where the value $c = 0.5$ has been used in the simulations of this paper.

3. Theory

We have not shown the convergence of the algorithm presented above. However, to convince the reader that binary sources can be separated from less sensors than sources, we show a result concerning two noise-free linear mixtures of binary sources:

Theorem. *Assume that we observe two noise-free linear mixtures of m independent binary sources taking values in $\{-1, 1\}$ and that the ICA basis vectors $\mathbf{a}_i$ are pairwise linearly independent. Then the ICA basis vectors multipied by 2 are edges of the convex hull of the set of mixtures. Furthermore, each basis vector appears twice in the convex hull and there are no other edges.*

Proof.

Take any column of $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_m]$, e.g. $\mathbf{a}_1$. We define the right half plane H of the vector $\mathbf{a}_1 = [a_{11}, a_{21}]^T$ as the set of vectors

$$H = \{\mathbf{b} \mid [a_{21}, -a_{11}]^T \mathbf{b} > 0\}$$

We can arrange so that each basis vector $\mathbf{a}_i, i = 2, \ldots, m$ is in H by changing the sign of the basis vectors and the corresponding source signals as necessary.

Now consider the vectors $\mathbf{x}_j = \mathbf{A}\mathbf{s}_j, j = -1, 1$ where $\mathbf{s}_j = [j, -1, \ldots, -1]^T$. Clearly $\mathbf{x}_1 - \mathbf{x}_{-1} = 2\mathbf{a}_1$. Take any source vector $\mathbf{s} = [s_1, \ldots, s_m]$ and the corresponding mixture $\mathbf{x} = \mathbf{A}\mathbf{s}$. We show that $\mathbf{d} = \mathbf{x} - \mathbf{x}_{-1} \in H$. Compute

$$[a_{21}, -a_{11}]^T \mathbf{d} = \sum_{i=1}^{m} [a_{21}, -a_{11}]^T (\mathbf{a}_i s_i - \mathbf{a}_i(-1)) = \sum_{i=1}^{m} [a_{21}, -a_{11}]^T \mathbf{a}_i (s_i + 1)$$

In the last sum each inner product $[a_{21}, -a_{11}]^T \mathbf{a}_i$ is positive for $i > 1$ and zero for $i = 1$. Since $(s_i + 1) \geq 0$, it follows that the sum is nonnegative. If it is zero,

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

then $\mathbf{s} = [j, -1, \ldots, -1]^T$, i.e. $\mathbf{s} = \mathbf{x}_j, j = -1$ or 1. This shows that the line passing through $\mathbf{x}_{-1}$ and $\mathbf{x}_1$ defines a half-plane containing all other mixtures. Therefore the edge with endpoints $\mathbf{x}_{-1}$ and $\mathbf{x}_1$ is contained in the convex hull of $\{\mathbf{x} = \mathbf{A}\mathbf{s} \mid \mathbf{s} = [s_1, \ldots, s_m]^T, s_i \in \{-1, 1\}\}$.

By considering points $\mathbf{x}'_j = [j, 1, 1, \ldots, 1]^T, j = -1, 1$ we can repeat the above reasoning to find that each basis vector appears twice in the convex hull.

To see that there are no other edges in the convex hull except the basis vectors $\mathbf{a}_i$, we consider an edge of the convex hull composed of more than one basis vector. Without loss of generality we can assume that this edge vector can be written $\mathbf{e} = \sum_{i=1}^{p} 2\mathbf{a}_i$. Since we assume that $\mathbf{e}$ is an edge of the convex hull, then it defines a half plane in which all the mixtures must lie in. Since the vectors $\mathbf{a}_i, i = 1, \ldots, p$ are pairwise linearly independent, at most one of them can be parallel to $\mathbf{e}$ and the rest must be in the same half-plane. But if this is the case, then their sum cannot be equal to $\mathbf{e}$ and we have a contradiction. $\square$

4. Simulations

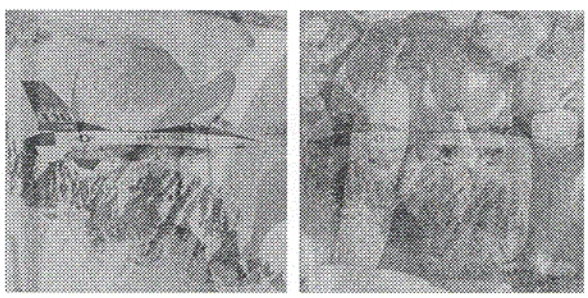

Figure 2: Two linear, noisy mixtures of four binary images.

An experiment was made where four binary 512×512 images were taken as the source signals. Two linear mixtures of the images were generated by multiplying the source vectors with the randomly chosen mixing matrix

$$\mathbf{A} = \begin{pmatrix} 1.1650 & 0.0751 & -0.6965 & 0.0591 \\ 0.6268 & 0.3516 & 1.6961 & 1.7971 \end{pmatrix}$$

and adding Gaussian noise to the mixtures. The standard deviation of the noise was 0.1. The original and separated images are compared in Fig. 3. It is observed that the images are recovered with small amount of noise. The two mixture images are shown in Fig. 2.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

Figure 3: Left: Original images. Right: Separated images.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 255-260

5. Discussion

The algorithm presented is able to separate linearly mixed independent binary source signals when there are less sensors than sources. The algorithm is somewhat heuristically justified but experiments verify its applicability to the problem.

The convergence of the algorithm has not yet been proven and some care must be taken when choosing appropriate values for the learning rate $\mu(t)$. Poorly chosen learning rates can lead to wrong results. Initial learning rate should be large enough for convergence but the learning rate should be decreased slowly for achieving good final accuracy.

References

[1] A. Belouchrani and J-F. Cardoso. Maximum likelihood source separation by the expectation-maximization technique: Deterministic and stochastic implementation. In *Proc. of the 1995 Int. Symposium on Nonlinear Theory and its Applications (NOLTA'95)*, pages 49–53, Las Vegas, USA, December 1995.

[2] M. Herrmann and H.H. Yang. Perspectives and limitations of self-organizing maps in blind separation of sources. In S. Amari et al., editors, *Progress in Neural Information Processing (ICONIP-96)*. Springer, 1996.

[3] J. Joutsensalo and P. Pajunen. Blind binary symbol estimation in CDMA system. Submitted to *IEEE Symposium on Personal and Indoor Mobile Radio Communications (PIMRC'97)*, September 1997.

[4] C. Jutten and J-F. Cardoso. Separation of sources: Really blind? In *Proc. of the 1995 Int. Symposium on Nonlinear Theory and its Applications (NOLTA'95)*, volume 1, pages 79–84, Las Vegas, USA, December 1995.

[5] J. Karhunen. Neural approaches to independent component analysis and source separation. In *Proc. of the 4th European Symposium on Artificial Neural Networks (ESANN'96)*, pages 249–266, Bruges, Belgium, April 1996.

[6] C. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, and J. Ortega. Separation of sources: A geometry-based procedure for reconstruction of n-valued signals. *Signal Processing*, 46:267–284, 1995.