ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 143-148

Acceptability conditions for BSS problems

V. Vigneron[1,2] S. Lagrange[2], C. Jutten[2]

[1]MATISSE-SAMOS UMR 8595
90, rue de Tolbiac
75634 Paris cedex 13

[2]LIS
avenue Félix Viallet
38040 Grenoble cedex

Abstract. The Herault-Jutten (HJ) algorithm is a neuromimetic structure capable to perform *blind source separation* (BSS) of a linear mixture from an array of sensors without knowing the transmission characteristics of the channels, nor the inputs. We aim to show here how unused theoretic stability conditions can profit to BSS.

1 Introduction

Consider the HJ network presented in Figure 1 [2]. Suppose the observation signals $x_i(t)$ and $x_2(t)$ are linear combinations of the input signals (primary sources) $s_1(t)$ and $s_2(t)$, *i.e.*:

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t), \qquad x_2(t) = a_{21}s_1(t) + a_{22}s_2(t), \tag{1}$$

which can be written in a more compact matrix form $X(t) = AS(t)$, where $X(t) = (x_1(t), x_2(t)), S = (s_1(t), s_2(t))$ and $A \in \mathbb{R}^{2\times 2}$ is a regular constant matrix. The sources are assumed to be zero-mean, *i.e.* $\mathbb{E}\{s_1\} = \mathbb{E}\{s_2\} = 0$, stationary and independent. The outputs of the system are given (omitting time index t) by:

$$y_1 = x_1 - w_{12}y_2, \qquad y_2 = x_2 - w_{21}y_1. \tag{2}$$

The network contains adaptative weights w_{12} and w_{21} which must be adjusted in such a way that: (i) each output signal will be proportional to only one primary sources by canceling the influence of the other sources (ii) the output signals $y_1(t)$ and $y_2(t)$ are statistically independent (after the adaptation process). Eliminating x_1 and x_2 from equations (1), we obtain from Eq.(2):

$$\begin{aligned} y_1 &= \tfrac{1}{1-w_{12}w_{21}}((a_{11} - w_{12}a_{21})s_1 + (a_{12} - w_{12}a_{22})s_2), &(3)\\ y_2 &= \tfrac{1}{1-w_{12}w_{21}}((a_{21} - w_{21}a_{11})s_1 + (a_{22} - w_{21}a_{12})s_2). &(4)\end{aligned}$$

The independence of the signals mathematically means that these signals must be at least decorrelated,*i.e.* $\mathbb{E}\{s_1s_2\} = 0$ and therefore $\mathbb{E}\{y_1y_2\} = 0$.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 143-148

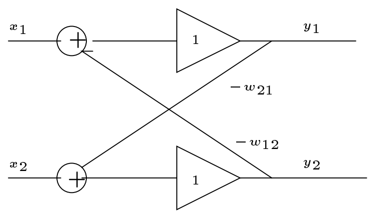

Figure 1: The two-channels Herault-Jutten network.

The latter condition is not sufficient to find the learning rule. Statistical independence of the output signals y_1 and y_2 implies $\mathbb{E}\{f(y_1)g(y_2)\} = 0$, where f and g are arbitrary (nonlinear) functions, typical examples are $f(x) = x^3, g(x) = x$. On the basis of the above requirements, Herault and Jutten have proposed to consider the following adaptation rule for the coefficients w_{ij} [2]

$$\frac{dw_{12}}{dt} = \mu y_1^3 y_2, \qquad \frac{dw_{21}}{dt} = \mu y_2^3 y_1, \tag{5}$$

where $\mu > 0$ is a constant. It is expected that $\mathbb{E}[\frac{dw_{ij}}{dt}] \to 0 \Rightarrow \mathbb{E}[f(y_i)g(y_j)] \to 0$ $(i \neq j)$. The functions $f(\cdot)$ and $g(\cdot)$ introduce higher-order moments.
The study is carried out first by examining stability conditions. An interval algorithm for BSS is then derived and equilibrium point zones are examined.

2 Stability conditions

A rigorous stability analysis of the network was given by Sorouchyari in terms of Lyapunov stability theory [6]. Equilibrium points of the H-J network are solutions of the system:

$$\begin{pmatrix} 1 & w_{12} \\ w_{21} & 1 \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{6}$$

The first solution reads $w_{12} = \frac{a_{12}}{a_{22}}$ and $w_{21} = \frac{a_{21}}{a_{11}}$, the second reads $w_{12} = \frac{a_{11}}{a_{21}}$ and $w_{21} = \frac{a_{22}}{a_{12}}$ [6]. It should be noted that the circuit Fig. 1 behaves as a feedback network which is *stable* under the condition $w_{12}w_{21} < 1$. Comon *et al.* [1] and Sorouchyari [6] investigate the convergence properties of the algorithm and perform a stability analysis for a 2 inputs/2 outputs network. They demonstrate that there are 4 paired equilibrium points in the sense that if the point (a, b) is a equilibrium point, then the point $(\frac{1}{a}, \frac{1}{b})$ is also a solution (see [6]).

Stability analysis performed by Sorouchyari consists to introduce a small perturbation on the stationnary points and to consider the further behaviour of the system around these points [4]. The stability conditions proposed are therefore:

$$w_{12}w_{21} < 1 \quad \text{and} \quad \mathbb{E}\{y_1^4\}\mathbb{E}\{y_2^4\} > 9(\mathbb{E}\{y_1^2 y_2^2\})^2 \tag{7}$$

Sorouchyari [6] shows that only one of the stationary points will be a stable separating solution.
The stability criteria in Eq. (7) which are based on the jacobian structure of the mixing matrix and which decide if the parameters w_{12} and w_{21} are *admissible*, have several drawbacks: (i) the choice of the initial value relies on guesswork, (ii) no guarantee of convergence to the global optimum can be provided, (iii) we are not interested in *the* optimal value, we rather like to characterize the set of *all* the acceptable values, (iv) uncertainty on the estimate is evaluated on the base of asymptotic assumptions, so no reliable evaluation is provided of the precision with which the estimated value is obtained.

3 Maximum *consistency*

This is why we shall look for the set of all models that are acceptable. The first step is then to list all the properties that the model should have to be acceptable. *Acceptability* will be defined here by a set of inequalities to be satisfied by the parameters. Once these conditions of acceptability have been defined, we wish to characterize – approximately but in a guaranteed way – the set of *all* values $\mathbb{W}$ (sometimes called *likelihood set*) of w_{12} and w_{21} that are *consistent*/acceptable given the data, *i.e.* the set of all values that are consistent with the prior feasible set and that satisfy all conditions of acceptability. This will be performed by the algorithm given in section 4.

Interval arythmetic An interval $[x]$ is a closed and connected subset of $\mathbb{R}$. The set of intervals of $\mathbb{R}$ will be denoted $\mathbb{IR}$. Let define $[x] = \{x \in \mathbb{R} \mid \overline{x} \leq x \leq \underline{x}, \overline{x} \in \mathbb{R}, \underline{x} \in \mathbb{R}\}$, where $\overline{x}$ and $\underline{x}$ are the upper and under bounds of $[x]$. A real number is an interval such that $\overline{x} = \underline{x} = x$. Basic operations on real numbers and vectors such as $+, -, *, /, \sin, \exp, \ldots$ extend to intervals in a natural way. For instance $[x] \times [y] = \{x \times y \mid x \in [x], y \in [y]\} = [\min(\underline{x} \cdot \underline{y}, \underline{x} \cdot \overline{y}, \overline{x} \cdot \underline{y}, \overline{x} \cdot \overline{y},), \max(\underline{x} \cdot \underline{y}, \underline{x} \cdot \overline{y}, \overline{x} \cdot \underline{y}, \overline{x} \cdot \overline{y},)]$. A *box* $[\boldsymbol{x}]$ of $\mathbb{R}^p$ is a cartesian product of p intervals, *i.e.* $[\boldsymbol{x}] = [x_1] \times \ldots \times [x_p]$.
$[f]$ is an *inclusion function* of the vector function f if, for any box $[\boldsymbol{x}]$, $[f]([\boldsymbol{x}])$ is also a box such that

$$\forall [\boldsymbol{x}], f([\boldsymbol{x}]) \triangleq \{f(\boldsymbol{x}) \mid \boldsymbol{x} \in [\boldsymbol{x}]\} \subset [f]([\boldsymbol{x}]). \tag{8}$$

The interval function$[f]([\boldsymbol{x}])$ is thus a box that contains $f(\boldsymbol{x})$, *i.e.* the enveloping box of $f([\boldsymbol{x}])$, see Fig. 2.a. The inclusion function $[f]$ for f is miniml if for any $[\boldsymbol{x}]$, $[f]([x])$ is the smallest box that contains $f(\boldsymbol{x})$. Interval computation makes it possible to obtain inclusion functions of a large class of nonlinear functions.
The interval union $[\boldsymbol{x}] \sqcup [\boldsymbol{y}]$ is the smallest box which contains the union of two boxes $[\boldsymbol{x}] \cup [\boldsymbol{y}]$.
In the HJ neural-like system, we aim at characterizing the parameter set $\mathbb{W} \subset \mathbb{R}^2$ such that $[\mathbb{W}, \mathbb{X}_1, \mathbb{X}_2] = \{(w_{12}, w_{21})^T \in \mathbb{W}, (x_1, x_2)^T \in \mathbb{X}_1 \times \mathbb{X}_2 | x_1 - w_{12} y_2 \in \mathbb{Y}_1, x_2 - w_{21} y_1 \in \mathbb{Y}_2\}$ from the knowledge of the set of pairs datum

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 143-148

$\mathbb{Y} = \mathbb{Y}_1 \times \mathbb{Y}_2 = \{(y_{1i}, y_{2i})\}_{i=1}^{T} \in \mathbb{R}^2$ and the vector function (suposedly inversible) f. Let $\mathbb{X}$ be the set of unknown primary sources. Then the BSS problem is formulated here as a problem of set inversion, which must be solved globally. The analysis of the parameter space of interest (w_{12}, w_{21}) will be performed by building sets of non-overlapping boxes with nonzero width. Exploration is limited to an initial box of interest, say $[\boldsymbol{w}]^{(0)}$, which is split by the algorithm into smaller boxes whenever needed until either a conclusion can be reached or the width of the box considered becomes smaller than some tolerance parameter ϵ. Interval analysis provide us two basic tests for deciding whether the given box $[f]([\boldsymbol{x}], [\boldsymbol{w}])$ is included in $\mathbb{Y}$:

$$[f]([\boldsymbol{x}], [\boldsymbol{w}]) \subset [\boldsymbol{y}] \Rightarrow [\boldsymbol{w}] \in \mathbb{W} \quad \text{i.e. } [\boldsymbol{w}] \quad \text{is feasible,} \tag{9}$$

$$[f]([\boldsymbol{x}], [\boldsymbol{w}]) \cap [\boldsymbol{y}] = \varnothing \Rightarrow [\boldsymbol{w}] \cap \mathbb{W} = \varnothing \quad \text{i.e. } [\boldsymbol{w}] \quad \text{is unfeasible.} \tag{10}$$

In all other cases, $[\boldsymbol{w}]$ is *indeterminate* and has a width *greater* than the precision parameter ϵ. Then it should be bisected into two subpavings namely W^- containing all boxes that were proved *feasible*, and W^+ consisting of all *indeterminate* boxes and the test should be *recursively* applied to these newly generated boxes (see Figure 2.b). From these subpavings, it is easy to bracket the portion of $W \subset \mathbb{W}$ contained in $[\boldsymbol{w}](0)$ as: $W^- \subset [\boldsymbol{w}](0) \cap W \subset W_{\text{out}} := W^- \cup W^+$. W_{out} is a finite union of boxes guaranteed to contain the portion

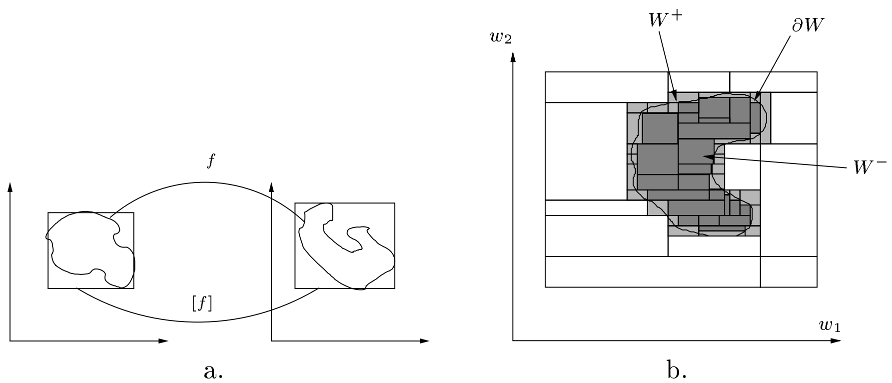

Figure 2: (a) Minimal inclusion function $[f]$ of a function f. (b) Feasibility of boxes.

of $\mathbb{W}$ of interest. A stack will be used to store the boxes still under considerations. Initialisation is performed by setting stack $= \varnothing, W^+ = \varnothing$ The algorithm requires a very large search box $[\boldsymbol{w}](0) =]-\infty; \infty[$ to which W is guaranteed to belong.

Upon completion of this algorithm, the consistency of W given the data is maximized and no indeterminate box will have a width larger than ϵ [3]. Under a few realistic technical conditions W^+ and W_{out} will tend to W (respectively from within and from without) when $\epsilon \to 0$ [5]. At the end, the set W^- of all boxes that have been proved to be feasible can be plotted in the parameter space (see Figure 2.b).

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 143-148

How conditions of stability of BSS solutions can be used as an optimization tool ? We suppose that the residuals between the data $(y_1, y_2)^T$ and the corresponding model output lie betwen known bounds, *i.e.* the confidence interval attached to individual measurements. Such a criterion is realized when minimizing the following nonlinear correlations (see section 2): $f([y_{1i}])g([y_{2i}]) \subset]0;\infty[, \forall i \subset \{1,\ldots,T\}$. As a consequence, a reduced domain for $[w_{12}],[w_{21}]$ (w.r.t. the initial one) is computed. Moreover, a necessary and sufficient condition for the HJ model to be asymptotically stable is that equations (7) are satisfied, iif $\left(9\overline{[y_1]^2[y_2]^2}^2 - \overline{[y_1]}^4\overline{[y_2]}^4\right) \subset]-\infty;0[$ and $(1-[w_{12}][w_{21}]) \subset]0;\infty[$, where $\overline{[y_1]^2[y_2]^2} = \frac{1}{T}\sum_i^T [y_{1i}]^2[y_{2i}]^2$, $\overline{[y_2]}^4 = \frac{1}{T}\sum_i^T [y_{1i}]^4$ and $\overline{[y_2]}^4 = \frac{1}{T}\sum_i^T [y_{2i}]^4$. This set of conditions defined so-called *acceptability conditions.*

4 Test simulation

As an illustration, consider the discrete time model in which the data have been generated by simulating for $k = 1,\ldots,500$: $\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} = \begin{pmatrix} 1 & 0,6 \\ 0,3 & 1 \end{pmatrix} \begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix}$, where $s_1(k) = \sin(7,3kT_e)$, $s_2(k) = \sin(4kT_e)$, $T_e = 0,2$ and by adding to $(x_1(k), x_2(k))$ a random white noise with a uniform distribution in the interval $[-10^{-2};10^{-2}]$.

```
CONSISTENCY ALGORITHM
inputs:  ([x1(k)],[x2(k)]), i = 1,...,T;  (assumed to be zero-mean))
         [y1_i] := [y2_i] :=]-∞;∞[, i = 1,...,T;
         [w12] := [w21] :=]-5,5[;
         [δ+] := [0;∞];  [δ-] := [-∞;0];  [κ] := [0;∞];
REPEAT
         FOR i := 1 TO T
  1         [y1_i] := ([x1_i] - [w12] × [y2_i]);
  2         [y2_i] := ([x2_i] - [w21] × [y1_i]);
  3         [δ+] := [δ+] ∩ (1 - [w12] × [w21]);
  4         [κ] := [κ] ∩ ([y1_i] × [y1_i] × [y1_i] × [y2_i]);
         ENDFOR
  5      [δ-] := [δ-] ∩ (‾[y1]‾^4 ‾[y2]‾^4 - 9 ‾[y1]^2[y2]^2‾^2)
WHILE the contraction is significant
output:  [w12],[w21];
```

We are thus certain that the interval data $([x_1(k)],[x_2(k)])$ contain the unknown true data. The prior domains (frame boxes) are $[w_{12}](0) := [w_{21}](0) :=]-5,5[$ and for the $[y_i](k)$'s are all taken equal to $[-10,10]^2$. The problem to solve is: given conditions of acceptability of the HJ system, compute accurate interval enclosure for the unknown true values for the y_{1i}, y_{2i}'s and w_{12}, w_{21}'s.

First results Fig. 3.a gives the true values of the stationary solutions. Only one of both is stable. After completion, the contracted intervals in the Figures 3.b and 3.c include the true values of the parameters. Fig. 3.b is obtained performing only steps 1-4 of the consistency algorithm (993 boxes). 3 zones are furnished by this algorithm.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 143-148

In Fig. 3.c performing the complete algorithm (steps 1-5), 2 zones on 3 are cancelled (probably non stable solutions). The computing time is about 0,5 seconds for both cases. Only black boxes are domain solutions.

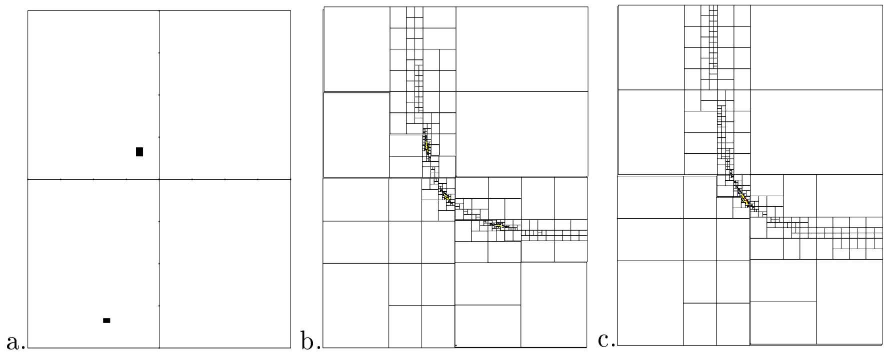

Figure 3: (a) True values of the parameters w_{12}, w_{21}. (b-c) Contracted bounded domains for $([w_{12}], [w_{21}])$ including stability conditions (c) or not (b).

5 Conclusion

A new method for analyzing the stability of the seminal Herault-Jutten algorithm for BSS is derived using interval analysis. We develop an algorithm capable to give, in a error bounded context, all the feasible solutions of the separating matrix. All these solutions are stable stationary solutions satisfying stability constraints. It is not computationally feasible for small-scale problems only.

References

[1] P. Comon, C. Jutten and J. Herault, "blind separation of sources, Part 2: Problems statement", *Signal Processing*, 24, pp. 11-20, 1991.

[2] J. Herault and C. Jutten, "Space ot time adaptative signal processing by neural network models", Neural networks for computing, Snowbird, 1986.

[3] L. Jaulin and E. Walter, "Guaranteed nonlinear parameter estimation from bound-error data via interval analysis", Mathematics and Computers in Simuation, **5**, pp. 123-137,1993.

[4] N. Minorsky, "Nonlinear oscillations", Van Nostrand, Reinhold, New York, 1962.

[5] J.P. Norton, "Special issue on bounded-error estimation: Issue 1". In *International Symposium Control and Signal Processing*, **8**(1), pp. 1-118, 1994.

[6] E. Sorouchyari, "Blind source separation, Part 3: Stabiity analysis", *Signal Processing*, 24, pp. 21-29, 1991.