A New Step-Adaptive Natural Gradient Algorithm for Blind Source Separation

Huan Tao, Jian-yun Zhang, and Lin Yu

Information Engineering Dept.,
Electronic Engineering Institute,
230037 HeFei, China
e_end@sina.com.cn

Abstract. The main differences between the mixed signals and origin signals: Gaussian probability density function, statistical independence and temporal predictability. The proposed BSS algorithms mainly derived from the Gaussian probability density function and statistical independence. A new adaptive method is proposed in the paper. The method uses the temporal predictability as cost function which is not studied as much as other generic differences between the properties of signals and their mixtures. Step-adaptive nature gradient algorithm is proposed to separate signals, which is more robust and effective. Compared to fixed step natural gradient algorithm, Simulations show a good performance of the algorithm.

1 Introduction

The goal of blind signal separation (BSS) is to recover mutually independent signals from their mixture. The problem has recently attracted a lot of interest because of its wide number of applications in diverse fields and some effective methods have been proposed such as in [2],[5],[6]. BSS can be very computationally demanding if the number of source signals is large. Application call for a BSS method which is computationally affordable, fast convergent, stable and reasonable accurate. BSS can be decomposed into two-steps: the observations are first pre-whitened by a whitening matrix, and then an orthogonal matrix can be separately calculated by constraining the source separation with different cost function. There are three main differences between the mixed signals and origin signals: Gaussian probability density function, statistical independence and temporal predictability. The proposed BSS algorithms mainly derived from the Gaussian probability density function and statistical independence. A new cost function based on temporal predictability is proposed by Reference [1] and proved to be effective which found a new way for blind signal separation. But it has limitations. The added calculation of temporal predictability makes the separate algorithm based on nature gradient time-consuming and instable.

D.-S. Huang, K. Li, and G.W. Irwin (Eds.): ICIC 2006, LNCIS 344, pp. 35 – 40, 2006.

In this paper, a new BSS method based on maximizing temporal predictability of signal mixtures has been introduced which has better separation performance.

2 Preliminaries

2.1 BSS Model

As a general model for BSS let L observed signals be related to N independent source signals $s_i(t)(t=1,\cdots,N)$ by $L\times N$ unknown channel matrix A:

$$x(t)=As(t)+n(t). \tag{1}$$

Where $s(t)=[s_1(t),s_2(t),\cdots,s_N(t)]^T$ and A is full-column rank matrix. $n(t)=(n_1(t),n_2(t),\cdots,n_m(t))^T$ is the vector of additive noise. Without loss of generality we assume in the derivation that signal are real-valued, $L=N$ and no noise. The BSS can operate into two steps. The first step is to pre-whiten the observations according to a whitening matrix B which results in a set of uncorrelated and normalized signals. Pre-whitening can be carried out in any of known methods and it is not dealt with here. After pre-whitened, an appropriate cost function based on high-order statistics can separate the sources by forcing their independence.

2.2 Temporal Predictability

In Reference [1], the definition of signal predictability F is:

$$F(W_{i,}x)=\log\frac{V(W_i,x)}{U(W_i,x)}=\log\frac{V_i}{U_i}=\log\frac{\sum_{i=1}^{n}(\bar{y}_i-y_i)^2}{\sum_{i=1}^{n}(\tilde{y}_i-y_i)^2}. \tag{2}$$

Where $y_i=W_ix_i$ is the value of signal y at time i. The term U_i reflects the extent to which y_i is predict by a short-term 'moving average' $\tilde{y}_i$ of values in y. In contrast, the term V_i is a measure of the overall variability of in y. As measured by the extend to which y_i is predicted by a long term 'moving average' $\bar{y}_i$ of values in y. The predicted value $\tilde{y}_i$ and $\bar{y}_i$ of y_i are both exponentially weighted sums of signal values measured up to time $i-1$,such that recent values have a larger weighting than those in the distant past:

$$\begin{aligned} \tilde{y}_i &= \lambda_S \tilde{y}_{i-1} + (1-\lambda_S) y_{i-1} \quad 0 \le \lambda_S \le 1 \\ \bar{y}_i &= \lambda_L \bar{y}_{i-1} + (1-\lambda_L) y_{i-1} \quad 0 \le \lambda_L \le 1 \end{aligned} . \tag{3}$$

The half-life h_L of λ_L is much longer than the corresponding half-life h_S of λ_S. The relation between a half-life h and parameter λ is defined as $\lambda = 2^{-1/h}$.

3 Step-Adaptive Nature Gradient Algorithm

Equation (2) can be rewritten as:

$$F = \log \frac{W_i \bar{C} W_i^t}{W_i \tilde{C} W_i^t} . \tag{4}$$

Where $\tilde{C}$ is long-term covariance between signal mixtures, and $\bar{C}$ is short-term covariance, which can be expressed as:

Reference [3] proved gradient ascent on F with respect to W_i could be used to maximize F .

$$\nabla F = \frac{2W_i}{V_i} \bar{C} - \frac{2W_i}{V_i} \tilde{C} . \tag{5}$$

Iteratively updating W_i until a maximum of F is located:

$$W_{i+1} = W_i + \mu \nabla F . \tag{6}$$

Commonly, the convergence time and stability is based upon the properly selection of step μ . Reference [4] has analyzed the stability conditions of nature gradient algorithm. A step-adaptive algorithm is desirable.

Here we proposed a new step-adaptive algorithm. Intuitively, we can use the distance between the separation matrix and optimal separation matrix to adjust the step adaptively. But the optimal separation matrix is unknown before the signals separated. An alternation we use

$$\Delta W(k) = \left\| W_{i+1} - W_i \right\|_F^2 . \tag{7}$$

To smooth the $\Delta W(k)$, $E(\Delta W(k))$ is used to perform step-adaptive adjustment. In the process of adaptation, the increasing of $E(\Delta W(k))$ means the fluctuation of algorithm, so a smaller step is desirable; on the contrary, the decreasing of $E(\Delta W(k))$ means a larger step is wanted to accelerate the convergence of algorithm. The updating expression of step is:

$$\mu(k+1) = \alpha(k)\mu(k). \tag{8}$$

α can be expressed as follows:

$$\alpha(k) = \begin{cases} 1+\gamma E(\Delta W(k)), E(\Delta W(k)) < E(\Delta W(k-1)) \\ \dfrac{1}{1+\beta E(\Delta W(k))}, E(\Delta W(k)) > E(\Delta W(k-1) \\ 1, \qquad else \end{cases} . \tag{9}$$

Where $0 < \beta < 1, 0 < \gamma < 1$. γ is in charge of the convergence speed and β controls the steady error when convergence.

$E(\Delta W(k))$ can be get form:

$$E(\Delta W(k+1)) = \frac{k}{k+1} E(\Delta W(k)) + \frac{1}{k+1} \Delta W(k+1). \tag{10}$$

4 Simulation and Performance

Compared to fixed step nature gradient algorithm, the performance of the step-adaptive nature gradient algorithm is evaluated through simulations.

Here we use three source signals with the sample of 5000 points. The mixing matrix A is generated randomly. The simulation parameters are as follows: $\lambda_L = 0.9, \lambda_S = 0.004, \mu_0 = 0.001, \beta = 0.5, \gamma = 0.06$. After separation, the separated result is: $\begin{bmatrix} 0.02 & 0.04 & 1 \\ 0.06 & 1 & 0.1 \\ 1 & 0.008 & 0.003 \end{bmatrix}$.

To evaluate the separation performance and the convergence speed of different algorithms, we use the correlation coefficiency between the original signals and recovered signals. The definition of correlation coefficiency is defined by (11).

$$\rho_{ij} = \frac{\mathrm{cov}(s_i, s_j)}{\sqrt{\mathrm{cov}(s_i, s_i)\,\mathrm{cov}(s_j, s_j)}}. \tag{11}$$

A comparison between fixed step nature-gradient algorithm with different steps and step-adaptive nature-gradient algorithm is done based on correlation coefficiency. The result is depicted in Fig. 1.

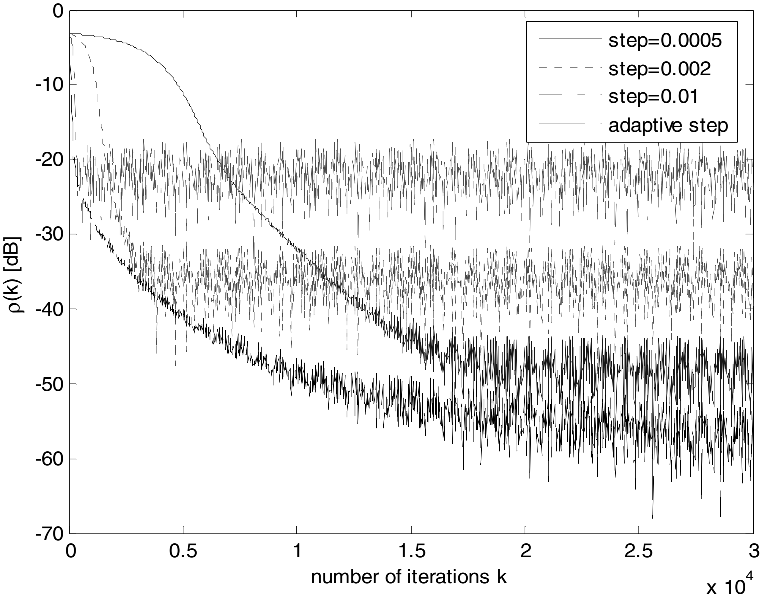

Fig. 1. Comparison of step-adaptation with fixed steps of 0.0005,0.002 and 0.01

From Fig.1, we can clearly see the step-adaptive nature-gradient algorithm is superior to fixed step nature-gradient algorithm in convergence speed.

5 Conclusion

A new adaptive separation algorithm is proposed based on maximizing temporal predictability of signal which is not studied as much as other generic differences between the properties of signals and their mixtures. The algorithm is step-adaptive. So it is more robust compare to fixed step natural gradient. Simulations show that it is effective and can get good separation precision. The step-adaptive nature gradient algorithm can also be used to other BSS method based on different cost function.

References

1. James, V. Stone.: Blind Source Separation Using Temporal Predictability, Neural computation(in press) (2001) 1196-1199
2. Belouchrani, A., Abed-Meraim, K., Cardoso, J.F.: A Blind Source Separation Using Second Order Statistics. IEEE Trans. On signal processing,Feb. Vol.45 (1997)434-444
3. Amari, S I.:Natural Gradient Works Efficiently in Learning. Neural Computation(1998) 251-276

4. Amari, S. I., Chen, T. P.,Cichocki, A.: Stability Analysis of Adaptive Blind Source Separation, Neural Networks(1997) 1345-1351.
5. Sergio, A., Cruces-Alvarez, Andrzej Cichocki, Shun-Ichi Amari.: On A New Blind Signal Extraction Algorithm: Different Criteria and Stability Analysis.. IEEE SIGNAL PROCESSING LETTERS, VOL.9,NO.8,AUGUST (2002)
6. Yan, Li, Peng Wen, David Powers.: Methods for The Blind Signal Separation Problem. IEEE Int. Conf. Neural Networks&Signal Processing. December (2002)