SEQUENTIAL LEAST SQUARES ALGORITHMS FOR BLIND CO-CHANNEL SIGNAL SEPARATION

Inseon Jang and Seungjin Choi

Department of Computer Science & Engineering, POSTECH, Korea
jinsn@nova.postech.ac.kr, seungjin@postech.ac.kr

ABSTRACT

In this paper we consider a problem of blind co-channel signal separation, the goal of which is to estimate multiple co-channel digitally modulated signals using an antenna array. We consider the joint maximum likelihood estimation [1] and present a sequential algorithm, which is referred to as *sequential joint maximum likelihood* (SJML) algorithm. In addition we also apply the sequential least squares to ILSP [2] and the resulting algorithm is referred to as the *sequential least squares with projection* (SLSP). Useful behavior of these two algorithms are confirmed by simulations.

1. INTRODUCTION

Mobile communications are growing rapidly in the number of subscribers and in the range of services, but available radio frequency spectrum is limited. Increasing spectrum efficiency is an important challenging problem in signal processing and wireless communications. A promising solution to this lies in exploiting spatial diversity (via antenna arrays). Array processing techniques allows multiple co-channel users per cell in order to increase the capacity.

Blind co-channel signal separation aims at estimating multiple co-channel digitally modulated signals, given only observation vector (measured at an antenna array) which consists of a superposition of signal waveforms plus additive noise. Several methods have been developed so far, among which, we pay attention to two algorithms: (1) Iterative Least Squares with Projection (ILSP) [2]; (2) joint maximum likelihood estimation [1]. It seems that these two algorithms are more suitable in the task of blind co-channel signal separation than conventional gradient based ICA algorithms [3, 4, 5] which did not take the effect of additive noise into account. We apply the sequential least squares method to these two algorithms. Then resulting algorithms are referred to as SLSP and SJML. These algorithms converge to a solution much faster to the gradient-based ICA algorithms and shows better performance in the presence of additive white Gaussian noise. Algorithm derivation and some numerical examples are presented.

2. PROBLEM FORMULATION

Consider d narrowband signals entered at an array of m sensors with arbitrary characteristics. There are multiple reflected and diffracted paths from the source to the array in a wireless environment or channel. So they arrive at array sensors for different angles, and with different attenuations and time delays. Output of antenna array becomes

$$\boldsymbol{x}(t) = \sum_{k=1}^{d} \sum_{l=1}^{q_k} \boldsymbol{a}(\theta_{kl}) p_k \alpha_{kl} s_k(t - \tau_{kl}) + \boldsymbol{v}(t), \quad (1)$$

where $\boldsymbol{a}(\theta_{kl})$ is the array response vector to a signal from direction θ_{kl}, p_k is the amplitude of the k-th signal, $s_k(\cdot)$ is the k-th signal waveform, q_k is number of subpaths for the k-th signal. And α_{kl} and τ_{kl} are the attenuation and time delay corresponding to k-th subpath, $\boldsymbol{v}(\cdot)$ is white complex symmetric Gaussian noise. The antenna array output modeled as phase-shifts under the narrowband assumption. So data model can be written as

$$\boldsymbol{x}(t) = \sum_{k=1}^{d} p_k \boldsymbol{a}_k s_k(t) + \boldsymbol{v}(t), \quad (2)$$

where $\boldsymbol{a}_k$ is the total array response vector A

$$\boldsymbol{a}_k = \sum_{l=1}^{D_k} \alpha_{kl} \boldsymbol{e}^{-j\omega_c \tau_{kl}} \boldsymbol{a}(\theta_{kl}), \quad (3)$$

and ω_c is the carrier frequency. The source signal structure can be written as

$$s_k = \sum_{n=1}^{N} b_k(n) g(t - nT), \quad (4)$$

where N is the number of symbols in a data batch, $\{b_k(\cdot)\}$ is the symbol sequence of the k-th user, T is the symbol period. And $g(\cdot)$ is the unit-energy signal waveform of duration T. Assume that the signals are symbol-synchronous, we perform matched filtering on (2) over each symbol period T. We obtain the following equivalent discrete representation of the data

$$\boldsymbol{x}(n) = \sum_{k=1}^{d} p_k \boldsymbol{a}_k s_k(n) + \boldsymbol{v}(n). \quad (5)$$

In matrix form,

$$\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{s}(n) + \boldsymbol{v}(n). \quad (6)$$

The problem addressed in this paper is the estimation of $\boldsymbol{s}(n)$, given $\boldsymbol{x}(n)$, and a good estimate of $\boldsymbol{A}$, where source signal is $\boldsymbol{s}(n) = [s_1(n), \cdots, s_d(n)]^T$, $\boldsymbol{x}(n)$ is the matched filter output for array output, array response $\boldsymbol{A}$ is a matrix which dimension is $m \times d$ and $\boldsymbol{v}(n)$ is white Gaussian noise.

3. SEQUENTIAL LEAST SQUARES WITH PROJECTION

We assume that the noise $\boldsymbol{v}(t)$ is isotropic Gaussian with zero mean and variance σ^2. Then the log-likelihood function of matched filter output is given by

$$\log L(\boldsymbol{A}, \boldsymbol{s}(1), \cdots, \boldsymbol{s}(N)) \propto -mN \log \sigma^2 - \frac{1}{\sigma^2} \sum_{n=1}^{N} \|\boldsymbol{x}(n) - \boldsymbol{A}\boldsymbol{s}(n)\|_F^2, \quad (7)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The maximization of the log-likelihood function with respect to the unknown $\boldsymbol{A}$ and $\boldsymbol{s}(n), n = 1, \cdots, N$, becomes identical to the minimization of separable least squares (LS) problem, i.e.,

$$\min_{\boldsymbol{A}, \boldsymbol{S}(N) \in \Omega} \|\boldsymbol{X}(N) - \boldsymbol{A}\boldsymbol{S}(N)\|_F^2, \quad (8)$$

where $\boldsymbol{X}(N) = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(N)]$, $\boldsymbol{S}(N) = [\boldsymbol{s}(1), \ldots, \boldsymbol{s}(N)]$ and Ω is a set of finite alphabet (for example, $\Omega = \{1, -1\}$ for BPSK). The ILSP algorithm [2] finds a local minimum of the cost function (8) iteratively with projecting the estimated $S(N)$ onto its nearest alphabet.

Motivated by the PAST algorithm [6], we consider the exponentially weighted LS cost function

$$\mathcal{J} = \sum_{i=1}^{n} \beta^{n-i} \|\boldsymbol{x}(i) - \boldsymbol{A}\boldsymbol{s}(i)\|^2, \quad (9)$$

where $\boldsymbol{s}(n)$ belongs to a certain alphabet depending on its constellation. The minimization of (9) leads to

$$\begin{aligned} \widehat{\boldsymbol{s}}(n) &= \widehat{\boldsymbol{A}}^{\#} \boldsymbol{x}(n), & (10) \\ \boldsymbol{z}(n) &= \text{proj}(\widehat{\boldsymbol{s}}(n)), & \\ \widehat{\boldsymbol{A}}^{(n)} &= \widehat{\boldsymbol{C}}_{xz}(n) \widehat{\boldsymbol{C}}_{zz}^{-1}(n), & (11) \end{aligned}$$

where

$$\begin{aligned} \widehat{\boldsymbol{C}}_{xz}(n) &= \sum_{i=1}^{n} \beta^{n-i} \boldsymbol{x}(i) \boldsymbol{z}^T(i) \\ &= \beta \widehat{\boldsymbol{C}}_{xz}(n-1) + \boldsymbol{x}(n)\boldsymbol{z}^T(n), \quad (12) \end{aligned}$$

$$\begin{aligned} \widehat{\boldsymbol{C}}_{zz}(n) &= \sum_{i=1}^{n} \beta^{n-i} \boldsymbol{z}(i) \boldsymbol{z}^T(i) \\ &= \beta \widehat{\boldsymbol{C}}_{zz}(n-1) + \boldsymbol{z}(n)\boldsymbol{z}^T(n), \quad (13) \end{aligned}$$

and the superscript # denotes the pseudo-inverse and $\text{proj}(\cdot)$ means the projection onto its nearest alphabet. The standard sequential LS (also known as recursive LS) is applied to derive the SLSP that is summarized in Table 1.

Note that Pajunen and Karhunen [7] proposed a similar LS algorithm to our SLSP. Their algorithm is a nonlinear version of PAST, so it is referred to as *the nonlinear PAST*. The difference between the SLSP and the nonlinear PAST is that the former exploits the generative model, whereas the latter does the recognition model. As will be demonstrated in simulations, the SLSP is better in the presence of white Gaussian noise. The benefit of learning the generative model was also emphasized in [8].

$$\begin{aligned} \widehat{\boldsymbol{s}}(n) &= \widehat{\boldsymbol{A}}^{\#}(n-1)\boldsymbol{x}(n) \\ \boldsymbol{z}(n) &= \text{proj}(\widehat{\boldsymbol{s}}(n)) \\ \boldsymbol{h}(n) &= \boldsymbol{P}(n-1)\boldsymbol{z}(n) \\ \boldsymbol{g}(n) &= \boldsymbol{h}(n)/[\beta + \boldsymbol{z}^T(n)\boldsymbol{h}(n)] \\ \boldsymbol{P}(n) &= \tfrac{1}{\beta}\text{Tri}\left\{\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{h}^T(n)\right\} \\ \boldsymbol{e}(n) &= \boldsymbol{x}(n) - \widehat{\boldsymbol{A}}(n-1)\boldsymbol{z}(n) \\ \widehat{\boldsymbol{A}}(n) &= \widehat{\boldsymbol{A}}(n-1) + \boldsymbol{e}(n)\boldsymbol{g}^T(n) \end{aligned}$$

Table 1: Algorithm outline for SLSP. The operator $\text{Tri}\{\cdot\}$ indicates that only the upper (or lower) triangular part of $\boldsymbol{P}(n) = \boldsymbol{C}_{zz}^{-1}(n)$ is calculated and its transposed version is copied to the another lower (or upper) triangular part then $\boldsymbol{P}(n)$ is symmetric matrix.

4. SEQUENTIAL JOINT MAXIMUM LIKELIHOOD

For the case of noise-free data, the estimate of $\boldsymbol{s}$ is obtained by a linear transform, $\widehat{\boldsymbol{s}} = \boldsymbol{A}^{\#}\boldsymbol{x}$, given the estimate of $\boldsymbol{A}$. It was pointed out in [1] that the reconstruction of original sources requires a nonlinear transform in the presence of noise. See [1] for more details.

As in [1], we assume that all sources have identical distributions and the noise is isotropic white Gaussian with zero mean and variance σ^2. Then the MAP cost function is given by

$$\log L(\boldsymbol{A}, \boldsymbol{s}(1), \cdots, \boldsymbol{s}(N)) \propto -\sum_{n=1}^{N} \left[\frac{1}{2} \|\boldsymbol{A}\boldsymbol{s}(n) - \boldsymbol{x}(n)\|_{\Sigma^{-1}}^2 + \sum_{i=1}^{d} f_i(s_i(n)) \right], \quad (14)$$

where $\|\boldsymbol{e}\|_{\Sigma^{-1}}^2$ is defined as $\boldsymbol{e}^T \Sigma^{-1} \boldsymbol{e}$ and $f_i(\cdot) = -\log p_i(\cdot)$ ($p_i(\cdot)$ represent the probability density function of source s_i). All sources $\boldsymbol{s}_i$ are constrained to have unit variance. And the independent component s_i are here constrained to have unit variance.

The optimal nonlinear function h for reconstructing independent components s_i is given by

$$\widehat{\boldsymbol{s}} = h(\widehat{\boldsymbol{A}}^{\#}\boldsymbol{x}), \quad (15)$$

where

$$h^{-1}(u) = (1 - \sigma^2)u + \sigma^2 f'(u), \quad (16)$$

where $f'(u) = \frac{df(u)}{du}$.

Since we are dealing with digitally modulated communication signals, it is reasonable to assume that all sources have uniform distribution with zero mean and unit variance. Then the probability density function is given by

$$p_s(s_i) = \frac{1}{2\sqrt{3}} \left\{ u(s + \sqrt{3}) - u(s - \sqrt{3}) \right\}, \quad (17)$$

$$\widehat{\boldsymbol{u}}(n) = \widehat{\boldsymbol{A}}^{\#}(n-1)\boldsymbol{x}(n)$$

$$\widehat{\boldsymbol{s}}(n) = \text{sign}(\widehat{\boldsymbol{u}}(n)) \min\left(|\widehat{\boldsymbol{u}}(n)|, \sqrt{3}\right)$$

if n=1, $\boldsymbol{z}(n) = \frac{\widehat{\boldsymbol{s}}(n)}{\|\widehat{\boldsymbol{s}}(n)\|}$

else $\widehat{\boldsymbol{c}}(n) = [\mathbf{1}, \widehat{\boldsymbol{s}}(n)]$

$$\boldsymbol{z}(n) = \frac{\widehat{\boldsymbol{s}}(n)}{\|\widehat{\boldsymbol{c}}(n)\|}$$

$$\boldsymbol{h}(n) = \boldsymbol{P}(n-1)\boldsymbol{z}(n)$$

$$\boldsymbol{g}(n) = \boldsymbol{h}(n)/[\beta + \boldsymbol{z}^T(n)\boldsymbol{h}(n)]$$

$$\boldsymbol{P}(n) = \frac{1}{\beta}\text{Tri}\left\{\boldsymbol{P}(n-1) - \boldsymbol{g}(n)\boldsymbol{h}^T(n)\right\}$$

$$\boldsymbol{e}(n) = \boldsymbol{x}(n) - \widehat{\boldsymbol{A}}(n-1)\boldsymbol{z}(n)$$

$$\widehat{\boldsymbol{A}}(n) = \widehat{\boldsymbol{A}}(n-1) + \boldsymbol{e}(n)\boldsymbol{g}^T(n)$$

Table 2: Algorithm outline for SJML

where $u(\cdot)$ denotes the unit step function. From this we have

$$f'(s_i) = -\frac{\delta(s+\sqrt{3}) - \delta(s-\sqrt{3})}{u(s+\sqrt{3}) - u(s-\sqrt{3})}, \tag{18}$$

where $\delta(\cdot)$ represent the unit delta function.

We substitute (18) into (16) and assumes the noise variance σ^2 is very small to obtain the truncation operator

$$h(u) = \text{sign}(u)\min(|u|, \sqrt{3}). \tag{19}$$

The truncation operator in (19) clips the values outside the interval $[-\sqrt{3}, \sqrt{3}]$, since the uniformly distributed random variable with unit variance cannot exceed $\pm\sqrt{3}$.

In [1], the alternating variable method was applied to find a local maximum of (14). Here we apply the sequential LS to derive our SJML algorithm that is summarized in Table 2. The only difference between SLSP and SJML lies in the reconstruction of $\widehat{\boldsymbol{s}}$, given $\widehat{\boldsymbol{A}}$. In SLSP, we used a finite alphabet property so that the projection onto its nearest alphabet followed LS estimation (carried out by pseudo-inversion). In SJML, the optimal nonlinear reconstruction was calculated under the uniform density model. With the different choice of the nonlinear reconstruction function h, the SJML is applicable to the case where sources have super-Gaussian distribution. For the case of super-Gaussian distribution, the sparse-code shrinkage operator was shown to be efficient in the task of denoising [9].

5. SIMULATIONS

We demonstrate the useful behavior of SLSP and SJML that are summarized in Table 1 and 2. and compare their performance to the nonlinear PAST [7] and the conventional natural gradient ICA algorithm [3, 4, 5].

We assume a uniform linear 3-element antenna array with each element being half wavelength spaced. We consider two digitally modulated QPSK sources with angles of arrival, 10° and 30°. Randomly-chosen initial value is assigned to $\boldsymbol{A}(0)$ or $\boldsymbol{W}(0)$. The identity matrix is assigned to $\boldsymbol{P}(0)$. For SJML, SLSP, and the nonlinear PAST, we used the forgetting factor $\beta = 0.99$ and for the ICA algorithm (with $\varphi_i(\widehat{s}_i) = |\widehat{s}_i|^2\widehat{s}_i$), we used the learning rate $\eta = 0.001$. At each SNR, we carried out 5 independent runs and calculated averaged BER (see Fig. 1). As shown in Fig. 1, our algorithms, SLSP and SJML outperforms the nonlinear PAST and the ICA algorithm in the presence of white Gaussian noise.

Besides the BER performance, we also evaluated the performance of algorithms in terms of the performance index (PI) that is defined by

$$\begin{aligned} \text{PI} \quad = \quad & \frac{1}{n(n-1)}\sum_{i=1}^{n}\left\{\left(\sum_{k=1}^{n}\frac{|g_{ik}|}{\max_j |g_{ij}|} - 1\right)\right. \\ & \left. + \left(\sum_{k=1}^{n}\frac{|g_{ki}|}{\max_j |g_{ji}|} - 1\right)\right\}, \end{aligned} \tag{20}$$

where g_{ij} is the (i,j)-element of the global system matrix $\boldsymbol{G} = \widehat{\boldsymbol{A}}^{\#}\boldsymbol{A} = \boldsymbol{W}\boldsymbol{A}$. The smaller value of PI, the better performance.

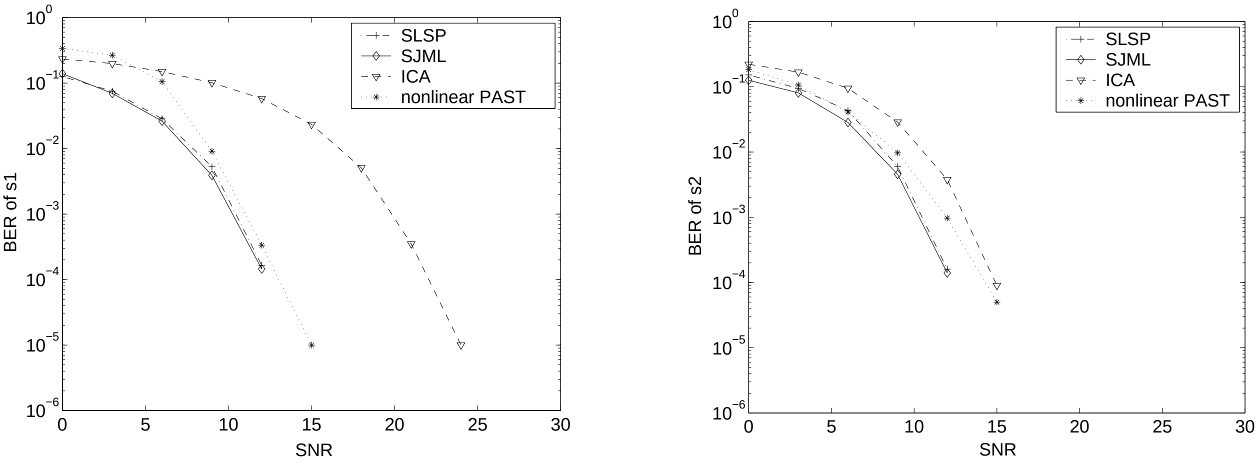
The convergence of all these algorithms are shown in Fig. 2. Since SLSP, SJML, and the nonlinear PAST employ the sequential least squares, they converge to a solution much faster than the gradient based ICA algorithm. In Fig. 2, the SLSP shows the fastest convergence because it exploits the finite alphabet property. However the computational complexity of SLSP for M-QAM ($M = 8, 16, 64, ...$) will be increased compared to the SJML, because the SLSP needs more search to project the data into its nearest constellation. The SJML and SLSP are based on learning generative model, whereas the nonlinear PAST exploits the recognition model. In the presence of noise, the recognition model does not take the additive noise effect into account. This is one of the reason why SJML and SLSP outperforms the nonlinear PAST, although they employ the same sequential least squares algorithm.

6. CONCLUSION

In this paper we proposed two sequential algorithms (SLSP, SJML) for blind co-channel signal separation. The key ingredient in the derivation of these algorithms was the sequential LS method. The algorithms are much faster than the gradient based source separation algorithms and are free of learning rate. The SLSP and SJML differs only in the part of reconstruction of sources, given the estimate of $\boldsymbol{A}$. The SLSP exploited the finite property, whereas the SJML employed the optimal nonlinear reconstruction function under the uniform density model. Thus the SJML can be applied to the case where sources have arbitrary distributions. Simulations verified the high performance of our algorithms.

7. ACKNOWLEDGMENT

This work was supported by Korea Ministry of Science and Technology under International Cooperative Research Program and by POSTECH BK21.



(a) (b)

Figure 1: BER performance of SLSP, SJML, the nonlinear PAST and the ICA algorithm for: (a) source 1; (b) source 2.

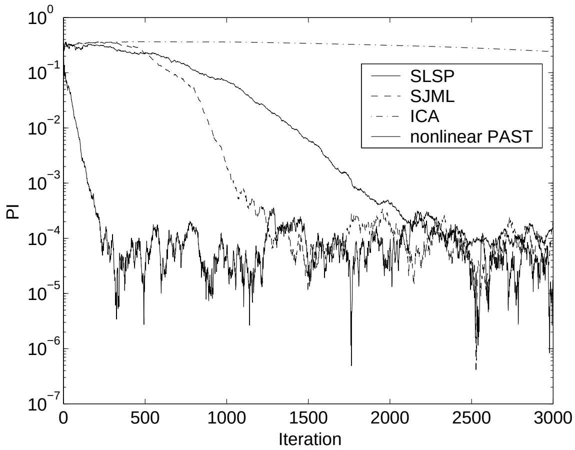


Figure 2: The convergence comparison for SLSP, SJML, the nonlinear PAST and the ICA algorithm.

8. REFERENCES

[1] A. Hyvärinen, "Independent component analysis in the presence of gaussian noise by maximizing joint likelihood," *Neurocomputing*, vol. 22, pp. 49-62, 1998.

[2] S. Talwar, M. Viberg and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 29-31, 1994.

[3] S. Amari and A. Cichocki, "Adaptive blidn signal processing - neural network approaches," *Proc. of the IEEE, Special Issue on Blind Identification and Estimation*, vol. 86, no. 10, pp. 2026-2048, Oct. 1998.

[4] J. F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 3017-3030, Dec. 1996.

[5] S. Choi, A. Cichocki and S. Amari, "Flexible independent component analysis," *Journal of VLSI Signal Processing*, vol. 26, pp. 25-38, Aug. 2000.

[6] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. on Signal Processing*, vol. 43, no. 1, pp. 95-107, 1995.

[7] P. Pajunen and J. Karhunen, "Least-squares methods for blind source separation based on nonlinear PCA," *Int. Journal of Neural Systems*, vol. 8, pp. 601-612, 1998.

[8] S. Choi, A. Cichocki, L. Zhang and S. Amari, "Approximate maximum likelihood source separation using the natural gradient," *Proc. IEEE Workshop on Signal Processing Advances in Vireless Communications*, (Taoyuan, Taiwan), pp. 235-238, 2001.

[9] A. Hyvärinen, "Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation", *Neural Computation*, vol. 11, no. 7, pp. 1739-1768, 1999.