Blind System Identification Using Multiple Sensors

Yingbo Hua, Howard Yang, and Mingyong Zhou

Department of Electrical and Electronic Engineering
University of Melbourne
Parkville, Victoria 3052
Australia

Abstract

This paper presents a study of blind system identification using measurements of multiple sensors. The mathematical model we consider here consists of multiple FIR channels driven by an unknown common source. We first show an orthogonal complement system matrix which is very useful in understanding the multi-channel system and developing efficient identification techniques. We then show a fast maximum likelihood method, its relation to a cross-relation method, and a new method utilizing a minimum noise subspace. We also report a study of strict identifiability of the multi-channel system and its relation to existing concepts.

1. Introduction

Blind system identification using multipe sensors is a fundamental problem which arises from a wide range of applications. Recently, it has received a significant attention in the context of blind channel equalization for mobile communications. In this application, the fractionally sampled outputs from the (real) communication channel can be modelled as outputs of multiple (virtual) channels driven by input symbol sequence, and the task is to identify the input symbols with or without explicitly identifying the impulse responses of the multiple channels. Despite the recent attention on mobile communications, this problem is also useful in other applications such as multi-sensor seismic signal analysis and multi-sensor image restoration.

In this paper, we consider a model where the available data are outputs of multiple FIR channels which are driven by an unknown input. Assuming a stationary white input or a stationary colored input with known source correlations, this model was initially considered in [1] and [2]. The matrix pencil approach in [1] was then extended in [7] for colored input with an unknown weak correlation. The cyclostationary approach in [2] was then generalized in [11] for an arbitrary source correlation.

For arbitrary nonstationary input, work has been done in [3-6] and [8]. Liu-Xu et al [3] proposed a cross-relation (CR) method. Moulines et al [4] proposed a noise subspace (NS) method. Shao-Nikias [5] presented an estimate-maximize based iterative maximum likelihood (EMML) algorithm. Slock [6] showed an iterative quadratic maximum likelihood (IQML) method for two-channels case. Hua [8] independently developed a two-step maximum likelihood (TSML) method for any number of channels. This paper reviews some major results shown in [8] and presents some further developments.

In Section 2, the multi-channel system is formulated. In Section 3, a very useful orthogonal complement (OC) system matrix is shown. In Section 4, the TSML method is described, and its connection to the CR method is made. In Section 5, the performance of the CR and TSML methods are evaluated. In Section 6, a new version of the NS method is shown which utilizes a minimum noise subspace (MNS). In Section 7, a new concept of strict identifiability is introduced and related to existing ones.

2. The M-Channel System

We consider M parallel FIR channels of maximum order equal to L. The output vector of the ith channel can be written as

$$\mathbf{y}_i=\mathbf{H}_{(i)}\mathbf{s}+\mathbf{w}_i$$

where $\mathbf{y}_i$ contains N consecutive output samples of channel i, i.e., $\mathbf{y}_i=[y_i(0), ..., y_i(N-1)]^T$, s contains N+L consecutive input samples, i.e., $\mathbf{s}=[s(-L), ..., s(N-1)]^T$, and $\mathbf{H}_{(i)}$ is a N×(N+L) Hankel matrix, i.e.,

$$\mathbf{H}_{(i)}=\begin{bmatrix} h_i(L) & \cdots & h_i(0) & & \\ & \ddots & & \ddots & \\ & & h_i(L) & \cdots & h_i(0) \end{bmatrix}$$

and $\mathbf{w}_i$ is the noise vector. The output of the M-channel system can then be put into

$$\mathbf{y}=[y_1 \;\; y_2 \;\; \cdots \;\; \cdots \;\; y_M]^T$$

and then

$$y = H_M s + w$$

where

$$H_M = \begin{bmatrix} H_{(1)} \\ H_{(2)} \\ \vdots \\ H_{(M)} \end{bmatrix}$$

and w is the system noise vector. H_M is known as a generalized Sylvester (GS) matrix of order N. It has the following properties:

Property 1 [1,2,12]*:* H_M has a full column rank if and only if the channel transfer functions:

$$H_i(z) = \sum_{k=0}^{L} h_i(k) z^{-k} \quad \text{for } i=1,\ldots,M,$$

do not share a common zero, provided N≥L. Furthermore, rank(H_M)≤2N, i.e., the rank of a GS matrix of order N is no larger than 2N.

Property 2 [4]*:* Let H_M' be constructed from $\{h'_i(k)\}$. Then range(H_M') ⊃ range(H_M) if and only if $h_i(k) = \alpha h'_i(k)$ for some constant α, provided that N≥L+1 and H_M has a full column rank.

The identification problem is to find the system impulse response $\{h_i(k)\}$ and the input $\{s(k)\}$ from $\{y_i(k)\}$.

3. The OC Matrix

The matrix H_M clearly reveals the structure in range(H_M). But it does not immediately show the structure in the orthogonal complement (OC) of range(H_M). To see the structure of the OC space, the following is useful. Define

$$G_2^H = \begin{bmatrix} -\bar{H}_{(2)} & \bar{H}_{(1)} \end{bmatrix}$$

$$G_M^H = \begin{bmatrix} \multicolumn{3}{c|}{G_{M-1}^H} & 0 \\ \hline -\bar{H}_{(M)} & & & -\bar{H}_{(1)} \\ & \ddots & & \vdots \\ & & -\bar{H}_{(M)} & -\bar{H}_{(M)} \end{bmatrix}$$

where $\bar{H}_{(i)}$ is the top-left (N-L)×N submatrix of $H_{(i)}$

Theorem 1 [8]:

Provided that all channels do not share a common zero and N≥2L (or N≥(L+1) for two channels case), an OC matrix of the system matrix H_M is G_M, i.e.,

$$P_G + P_H = I$$

where P_G and P_H denote the projection matrices onto range(G_M) and range(H_M), respectively

Note that for two channels case (M=2), the proof is straigtforward and a similar result was shown in [6]. But for M>2, the proof requires a significant effort and the theorem is novel. This theorem played a crucial role in developing the TSML method as described next.

4. The TSML Method

Define a data matix:

$$Y = Y_M = \begin{bmatrix} \multicolumn{3}{c|}{Y_{M-1}} & 0 \\ \hline \bar{Y}_{(M)} & & & -\bar{Y}_{(1)} \\ & \ddots & & \vdots \\ & & \bar{Y}_{(M)} & -\bar{Y}_{(M-1)} \end{bmatrix}$$

where

$$\bar{Y}_{(i)} = \begin{bmatrix} y_i(L) & \cdots & y_i(0) \\ \vdots & \vdots & \vdots \\ y_i(N-1) & \cdots & y_i(N-L-1) \end{bmatrix}$$

Then the TSML method is as follows:

Step 1: Minimize $h^H(Y^H Y)h$ with $||h||=1$ to yield h_c.

Step 2: Minimize $h^H(Y^H(G_c^H G_c)^+ Y)h$ with $||h||=1$ to yield h_e, where G_c is the same as G_M but constructed from h_c.

The TSML method consists of two major steps where each step minimizes a quadratic function, and hence it is computationally efficient compared to existing algorithms such as the EMML algorithm [5]. For two channels case, the TSML method coincides with the IQML method shown in [6]. Note that Step 1 of the TSML method interestingly coincides with the CR technique developed by Liu-Xu et al [3]. It can be shown that Step 1 yields the exact (consistent) estimate in the absence of noise [3] or for large data size if the noise is white [10]. It has also been shown [8] that Step 2 yields a statistically high-SNR efficient estimate whenever the OC matrix G_c is constructed from a consistent estimate of **h**.

5. Performance of the CR method

The CR method is simply the first step of the TSML method. It is shown in [10] that if the data length N is large or/and the SNR is large, the mean square error (MSE) of the CR estimate $\mathbf{h}_c$ is given by

$$E\left\{\|\mathbf{h}_c\|^2\right\}=\frac{1}{N-L}(\sigma^2 E_1+\sigma^4 E_2)$$

where σ^2 denotes the noise variance, and E_1 and E_2 are independent of σ and large N. To show an example of how well the CR method can perform, we consider a two-channel system where the transfer function of each channel is given by

$$H_i(z)=(1-e^{j\theta_i}z^{-1})(1-e^{-j\theta_i}z^{-1})$$

with i=1, 2, and $\theta_1=\pi/11$ and $\theta_2=\pi/11+\delta$. Note that δ denotes a separation between the zeros of the two channels. Let

$$SNR=20\log_{10}\left(\frac{E\{\|\mathbf{H}_M\mathbf{s}\|\}}{E\{\|\mathbf{w}\|\}}\right).$$

For SNR=45dB and N=30, Figure 1 shows the normalized-root-mean-square-error (NRMSE) of the CR estimate versus δ. Also shown in Figure 1 are the Cramer-Rao bound (CRB) and the NRMSE of the TSML method. This figure suggests that the CR method can achieve the CRB when the M-channel system is well conditioned (i.e., without closely located zeros), and the CR method does not achieve the CRB otherwise.

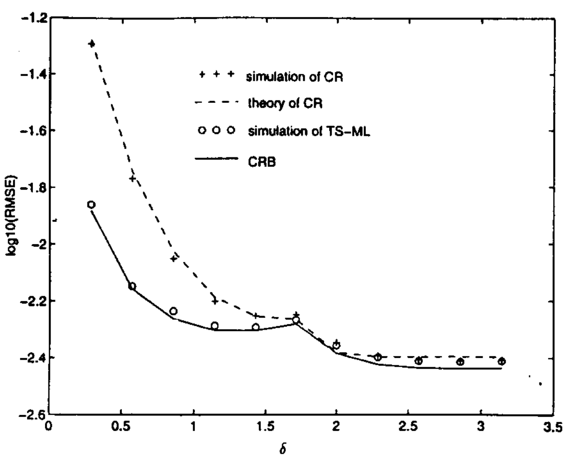

Figure 1

6. The MNS Method

Based on property 2 of the GS matrix $\mathbf{H}_M$, it was shown in [4] that the system impulse response can be uniquely determined by the noise vectors of the data covariance matrix:

$$\mathbf{R}_y=E\left\{\mathbf{y}\mathbf{y}^H\right\}=\mathbf{H}_M\mathbf{R}_s\mathbf{H}_M+\mathbf{R}_w$$

where

$$\mathbf{R}_s=E\left\{\mathbf{s}\mathbf{s}^H\right\}$$

provided that the source covariance matrix $\mathbf{R}_s$ is nonsingular, the noise covariance matrix $\mathbf{R}_w$ is known up to a constant, and the conditions of property 2 are met. It was also shown in [4] that for two channels case, any single noise vector of $\mathbf{R}_y$ is sufficient to yield the unique solution asymptotically. But for more than two channels, it was conjectured in [4] that a single noise vector of $\mathbf{R}_y$ can also yield unique estimate. With the introduction of the OC matrix $\mathbf{G}_M$, this conjecture can be shown unfortunately to be wrong.

Theroem 2: No single noise vector of $\mathbf{R}_y$ can asymptotically uniquely determine the system impulse response for more than two channels case.

Proof: It suffices to consider the case where each of $\mathbf{H}_M$ and $\mathbf{R}_s$ has a full column rank and $\mathbf{R}_w$ is absent. A noise vector $\mathbf{v}$ of $\mathbf{R}_y$ must be a linear combination of columns of $\mathbf{G}_M$, i.e., $\mathbf{v}=\mathbf{G}_M\mathbf{a}$ for some vector $\mathbf{a}$. If $\mathbf{v}$ determines $\{h_i(k)\}$ uniquely, $\{h_i(k)\}$ must be the unique solution to (based on the NS principle)

$$\mathbf{v}^H\mathbf{H}_M=\mathbf{0}$$

This equation can then be expressed (after a lightly tedious manipulation) as

$$\mathbf{h}^T\mathbf{Q}=\mathbf{0}$$

where $\mathbf{h}$ is the M(L+1)×1 vector of $\{h_i(k)\}$ and $\mathbf{Q}$ is a M(L+1)×(N+L) GS matrix of order L+1. From Property 1, rank(Q)≤2(L+1) and hence the above equation has no unique solution for M>2.

We can conjecture, however, that for the M-channel system the minimum number of noise vectors required is M-1. The M-1 noise vectors must be chosen properly. In fact, each of the M-1 noise vectors can be found by computing the least eigenvector of a covariance matrix corresponding to a distinct pair of channels. Since each such noise vector can determine the impulse responses of a distinct pair of channels uniquely up to a constant, the M-1 noise vectors can determine the impulse responses of all M channels uniquely up to a constant. The M-1 noise vectors can be obtained in parallel as shown below (for M=3):

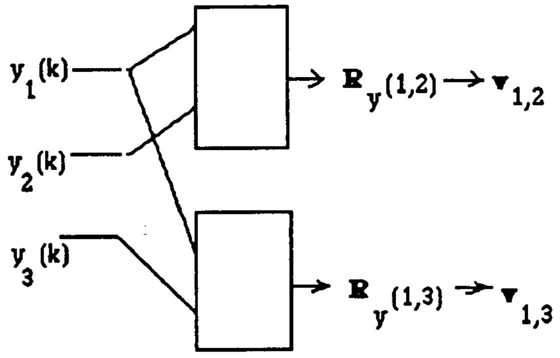

Figure 2

Also note that each such noise vector is computed from a smaller covariance matrix than $\mathbf{R}_y$. This MNS (minimum noise subspace) method is more efficient in computation than the original NS method [4]. A detailed performance analysis of this approach will be shown in our future communications.

7. Strict Identifiability

Channel identifiability conditions have been recently investigated by several researchers. Assuming that the input to all channels is white, stationary and infinitely long, Tong et al [1] and Li-Ding [2] studied the channel identifiability conditions based on the second order statistics of the channel outputs. Assuming that the input is an unknown deterministic sequence, Liu-Xu et al [3] showed a number of sufficient and necessary identifiability conditions based on the cross-relation (CR) equation. This problem was further considered by Hua [8] where the channel identifiability conditions were analyzed based on a Fisher information (FI) matrix. An M-channel system is defined to be FI identifiable if and only if the FI matrix has a desired rank.

More recently, we studied the identifiability of the M-channel FIR system in a strict sense. An M-channel FIR system is defined to be strictly identifiable if the given channel outputs, in the absence of noise, can only be realized by a unique system impulse response and a unique input sequence. In contrast to several existing definitions of identifiability, the strict identifiability is independent of any identification technique or any preprocessing on the channel outputs. Among several fundamental observations made in [9], we have found a surprising fact that the CR identifiability, the FI identifiability and the strict identifiability are equivalent to each other provided that the number of the output samples of each channel is no less than twice the maximum order of the FIR channels.

References

[1] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Trans. Info. Theory*, April 1994.

[2] Y. Li and Z. Ding, "Blind channel identification based on second-order cyclostationary statistics," *Proc. of IEEE ICASSP'93*, pp.81-84, Minneapolis, MN, April 1993.

[3] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind channel identification of multi-channel FIR systems," *Proc of IEEE ICASSP'94*, Vol. 4, pp. 581-584, Australia, April 1994.

[4] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *Proc. of IEEE ICASSP'94*, Vol. 4, pp. 573-576, Australia, April 1994.

[5] M. Shao and C. L. Nikias, "An ML/MMSE estimation approach to blind equalization," *Proc of IEEE ICASSP'94*, Vol. 4, pp. 569-572, Australia, April 1994.

[6] D. T. M. Slock, "Blind fractionally-space equalization, perfect-reconstruction filter bank and multichannel linear prediction," *Proc. of IEEE ICASSP'94*, Vol. 4, pp. 585-588, Australia, April 1994.

[7] Y. Hua, H. Yang, and W. Qiu, "Source correlation compensation for blind channel identification based on second-order statistics," *IEEE Signal Processing Letters*, Vol. 1, No. 8, pp. 119-120, August 1994

[8] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," submitted. Also presented at *28th Annual Asilomar Conference on SSC*, CA, Nov 1994.

[9] Y. Hua and M. Wax, "Strict identifiability of multiple FIR channels driven by an unknown arbitrary input," submitted.

[10] H. Yang and Y. Hua, "On performance of cross-relation method for blind channel identification," submitted.

[11] G. Giannakis, "A linear cyclic correlation approach for blind identification of FIR channels," presented at *28th Annual Asilomar Conference on SSC*, CA, Nov 1994.

[12] R. R. Bitmead, S.-Y. Kung, B. D. O. Anderson, and T. Kailath, "Greatest common division via generalized Sylvester and Bezout matrices," *IEEE Transactions on Automatic Control*, Vol. 23, No. 6, pp. 1043-1047, Dec. 1978.