ADAPTIVE IDENTIFICATION OF BILINEAR SYSTEMS

Zhiwen Zhu

Communications Research Laboratory
McMaster University
Hamilton, Ontario, Canada L8S 4K1
zhuz@sunburst.crl.mcmaster.ca

Henry Leung

Surface Radar Section
Defence Research Establishment Ottawa
Ottawa, Ontario, Canada K1A 0Z4
Henry.Leung@dreo.dnd.ca

ABSTRACT

The paper considers the adaptive identification of bilinear systems using the equation-error approach. An improved least squares (ILS) objective function is presented to reduce the bias of coefficient estimation in the case of large measurement noise when the standard least squares (LS) technique is used. An adaptive algorithm based on the ILS criterion is proposed for the identification of the bilinear system. Numerical simulations are given to demonstrate the effectiveness of the adaptive ILS algorithm. Compared with the least mean square (LMS) technique, the proposed algorithm has superior identification performance.

1. INTRODUCTION

System identification is an important issue in many areas including digital signal processing [1], process control [2] and communications [3]. It is concerned with characterizing an unknown system using measurements of the system's input and output signals. Linear system models have played a very crucial role in the development of various signal processing. But many real-life systems show nonlinear behavior [4]. As a result, it is important to consider nonlinear models in order to accurately characterize real-life phenomena. For example, high-speed wireless communication channels often need nonlinear equalizer for acceptable performance [5]. One particularly attractive model for nonlinear system identification is the bilinear model, which can characterize a wide class of nonlinear phenomena [6].

There are two different approaches to solve adaptive identification problems using the bilinear system model — equation-error and output-error approaches [9]. The mean square estimation error surface in the equation-error approach has a unique minimum. However, the recursive model leads to correlated residuals and thus biased coefficient estimations when the conventional least squares estimation technique is used. The output-error approach maybe less sensitive to additive noise in the desired signal than the equation-error approach. However, its error surface may have local minima and the adaptive identification may not converge to the global minimum, unless the system is initialized properly. The latter approach has been addressed by many researchers [7],[8].

In this paper, we consider the equation-error approach and introduce an improved objective function to replace the least squares (LS) function. While the LS criterion is to minimize the one-step-prediction error, the improved least squares (ILS) technique is to minimize the orthogonal Euclidean distance between every noisy output point and the hyper-surface defined by the bilinear system. The ILS technique can estimate the coefficients with a strongly reduced bias in the case of large amplitude measurement noise. An adaptive scheme based on the ILS criterion is then developed. Computer simulations are used to compare the performance of the proposed adaptive algorithm with that of the least mean squares (LMS) algorithm.

2. BILINEAR SYSTEM IDENTIFICATION

In this paper, we consider a bilinear system whose input-output relationship is governed by

$$\begin{aligned} x(n) &= \sum_{i=1}^{N} a_{1,i}x(n-i) + \sum_{j=0}^{N-1} a_{2,j}u(n-j) \\ &+ \sum_{i=1}^{N}\sum_{j=0}^{N-1} b_{i,j}x(n-i)u(n-j), \end{aligned} \tag{1}$$

and the measurement equation

$$y(n) = x(n) + e(n), \tag{2}$$

where

$u(n)$	input signal
$x(n)$	output signal
$y(n)$	measurement signal
$e(n)$	zero mean Gaussian white noise
$a_{1,i}, a_{2,i}, b_{i,j}$	coefficients

Using vector notation, we can rewrite (1) as

$$x(n) = \mathbf{w}_1^{\mathrm{T}}(n)\mathbf{x}(n-1) + \mathbf{w}_2^{\mathrm{T}}\mathbf{u}(n), \tag{3}$$

where

$$\mathbf{x}(n-1)=\begin{bmatrix} x(n-1)\\ x(n-2)\\ \cdots\\ x(n-N)\end{bmatrix},\ \mathbf{u}(n)=\begin{bmatrix} u(n)\\ u(n-1)\\ \cdots\\ u(n-N+1)\end{bmatrix},$$

$$\mathbf{w}_1(n)=\begin{bmatrix} a_{1,1}+\sum_{j=0}^{N-1} b_{1,j}u(n-j)\\ a_{1,2}+\sum_{j=0}^{N-1} b_{2,j}u(n-j)\\ \vdots\\ a_{1,N}+\sum_{j=0}^{N-1} b_{N,j}u(n-j)\end{bmatrix},\ \mathbf{w}_2=\begin{bmatrix} a_{2,0}\\ a_{2,1}\\ \cdots\\ a_{2,N-1}\end{bmatrix}.$$

The basic idea of equation-error approach is to use samples of the input signal $u(n)$ and the measurement signal $y(n)$ to obtain an adaptive identification system as

$$\hat{x}(n)=\hat{\mathbf{w}}_1^{\mathrm{T}}(n)\mathbf{y}(n-1)+\hat{\mathbf{w}}_2^{\mathrm{T}}\mathbf{u}(n), \tag{4}$$

where

$$\hat{\mathbf{w}}_1(n)=\begin{bmatrix} a_{1,1}(n)+\sum_{j=0}^{N-1} b_{1,j}(n)u(n-j)\\ a_{1,2}(n)+\sum_{j=0}^{N-1} b_{2,j}(n)u(n-j)\\ \vdots\\ a_{1,N}(n)+\sum_{j=0}^{N-1} b_{N,j}(n)u(n-j)\end{bmatrix},\ \hat{\mathbf{w}}_2=$$

$$\begin{bmatrix} a_{2,0}(n)\\ a_{2,1}(n)\\ \vdots\\ a_{2,N-1}(n)\end{bmatrix},\ \mathbf{y}(n-1)=\begin{bmatrix} y(n-1)\\ y(n-2)\\ \vdots\\ y(n-N+1)\end{bmatrix},\ \text{and}$$

$a_{1,i}(n),\ a_{2,i}(n)$, and $b_{i,j}(n)$ $(1\le i\le N,\ 0\le j\le N-1)$ are coefficients of the adaptive bilinear system at time n.

The conventional identification method considers the LS criterion

$$V(\hat{\mathbf{w}}_1(n),\hat{\mathbf{w}}_2)=E[(y(n)-\hat{x}(n))^2], \tag{5}$$

where E denotes mathematical expectation. A stochastic gradient update equation for minimizing the mean square estimation error $V(\hat{\mathbf{w}}_1(n),\hat{\mathbf{w}}_2)$ is

$$\begin{aligned} a_{1,i}(n+1)&=a_{1,i}(n)+\mu_{1,i}\varepsilon(n)y(n-i),\\ a_{2,j}(n+1)&=a_{2,j}(n)+\mu_{2,j}\varepsilon(n)u(n-i),\\ b_{i,j}(n+1)&=b_{i,j}(n)+\nu_{i,j}\varepsilon(n)y(n-i)u(n-j),\\ i&=1,\ 2,\ \cdots,\ N\ \text{ and } j=0,\ 1,\ \cdots,\ N-1,\end{aligned} \tag{6}$$

where $\varepsilon(n)=y(n)-\hat{x}(n)$, $\mu_{1,i}$, $\mu_{1,j}$, and $\nu_{i,j}$ are small positive constants (stepsize) that control stability and rate of convergence of the adaptive algorithm. (6) is the conventional adaptive identification method — LMS algorithm.

Note that the LS technique seeks to minimize the mean squares regression for $y(n)$ with respect to plane $(\mathbf{y}(n-1),\ \mathbf{u}(n))$. If $y(n)$ contains noise, the statistics of the input signal to the adaptive identification system will be biased from the statistics of the ideal desired response signal and this will result in biased estimations [10]. One appropriate minimization objective function to tackle the problem is given by

$$V_b(\mathbf{w}_1(n),\mathbf{w}_2)=E[d^2(n)], \tag{7}$$

$$\begin{aligned} d^2(n)=\min_{\mathbf{x}(n-1)}&[\|\mathbf{x}(n-1)-\mathbf{y}(n-1)\|^2+\\ &(y(n)-\mathbf{w}_1^{\mathrm{T}}(n)\mathbf{x}(n-1)-\mathbf{w}_2^{\mathrm{T}}\mathbf{u}(n))^2],\end{aligned} \tag{8}$$

where $\|\bullet\|$ denotes L_2 norm. The objective function represents a minimization of the orthogonal Euclidean distance between every noisy point $(y(n),\ \mathbf{y}(n-1))$ and the hypersurface defined by (1).

Through derivation, it is known that when the vector $\mathbf{x}(n-1)$ is chosen as

$$\hat{\mathbf{x}}(n-1)=\frac{\mathbf{w}_1(n)(y(n)-\mathbf{w}_2^{\mathrm{T}}\mathbf{u}(n))+\mathbf{y}(n-1)}{1+\mathbf{w}_1^{\mathrm{T}}(n)\mathbf{w}_1(n)}, \tag{9}$$

the term in the bracket of (8) reaches its minimum, and

$$d^2(n)=\frac{(y(n)-\mathbf{w}_1(n)^{\mathrm{T}}\mathbf{x}(n-1)-\mathbf{w}_2^{\mathrm{T}}\mathbf{u}(n))^2}{1+\mathbf{w}_1^{\mathrm{T}}(n)\mathbf{w}_1(n)}. \tag{10}$$

The objective function in (7) can be written as

$$\begin{aligned} &V_b(\mathbf{w}_1(n),\mathbf{w}_2)\\ &=E\left[\frac{(y(n)-\mathbf{w}_1(n)^{\mathrm{T}}\mathbf{x}(n-1)-\mathbf{w}_2^{\mathrm{T}}\mathbf{u}(n))^2}{1+\mathbf{w}_1^{\mathrm{T}}(n)\mathbf{w}_1(n)}\right]\end{aligned} \tag{11}$$

Similar to the LMS technique, an updating algorithm for the coefficients $a_{1,i}$, $a_{2,i}$, and $b_{i,j}$ $(1\le i\le N,\ 0\le j\le N-1)$, which attempts to minimize the objective function in (11), can be obtained by applying the stochastic gradient method

$$\begin{aligned} a_{1,i}(n+1)&=a_{1,i}(n)-\frac{\mu_{1,i}}{2}\nabla_{a_{1,i}}\hat{V}_b(\mathbf{w}_1(n),\mathbf{w}_2),\\ a_{2,j}(n+1)&=a_{2,j}(n)-\frac{\mu_{2,j}}{2}\nabla_{a_{2,j}}\hat{V}_b(\mathbf{w}_1(n),\mathbf{w}_2),\\ b_{i,j}(n+1)&=b_{i,j}(n)-\frac{\nu_{i,j}}{2}\nabla_{b_{i,j}}\hat{V}_b(\mathbf{w}_1(n),\mathbf{w}_2)\\ i&=1,\ 2,\ \cdots,\ N\ \text{ and } j=0,\ 1,\ \cdots,\ N-1,\end{aligned} \tag{12}$$

where

$$\hat{V}_b(\hat{\mathbf{w}}_1(n),\hat{\mathbf{w}}_2)=\frac{(y(n)-\hat{\mathbf{w}}_1(n)^{\mathrm{T}}\mathbf{x}(n-1)-\hat{\mathbf{w}}_2^{\mathrm{T}}\mathbf{u}(n))^2}{1+\hat{\mathbf{w}}_1^{\mathrm{T}}(n)\hat{\mathbf{w}}_1(n)}.$$

Then an adaptive algorithm (12) is obtained to identify the coefficients of the bilinear system based on the new objective function (11). For simplicity, (12) is called as adaptive ILS algorithm.

3. COMPUTER SIMULATIONS

In this section, we present some simulation results comparing the performance of the adaptive ILS algorithm with that of the LMS algorithm. The problem considered here is to estimate the coefficients of a bilinear system, which is governed by the following equation

$$\begin{aligned} x(n) &= 1.5x(n-1) - 0.7x(n-2) + 0.8u(n) \\ &\quad +0.5u(n-1) + 0.24x(n-1)u(n), \\ y(n) &= x(n) + e(n). \end{aligned}$$

Assume that the input signal $u(n)$ belongs to a stationary and zero-mean white Gaussian process with variance 0.02. The measurement noise is also a Gaussian process $N(0, \sigma_e^2)$. σ_e^2 is a constant for varying SNR.

The adaptive ILS algorithm and the LMS algorithm are used to estimate the coefficients in the above bilinear system. In our study, the bias and standard derivation (STD) of the estimated coefficients are ensemble averages over 50 Monte Carlo trials, in which the estimation error is measured by the time-averaging in the range $[10001, 20000]$. For simplicity, we assume the stepsizes corresponding to different coefficients are the same. That is, $\mu_{1,i} = \mu_{2,j} = \nu_{i,j}$, for $i = 1, 2, \cdots, N$ and $j = 0, 2, \cdots, N-1$ in (6) and (12).

Table 1 presents the performance of the two identification algorithms with different stepsizes as SNR=20dB. When the stepsize is chosen as 5×10^{-2}, the LMS algorithm diverges, and the corresponding results are not given. It is noted that the LMS algorithm produces a substantial estimation bias, and the adaptive ILS algorithm can largely reduce it. The improvement degrees are different depending on coefficients. Compared with the LMS method, while it reduces the estimation bias twice for coefficients a_{20} and b_{10}, the adaptive ILS algorithm reduce the bias about one hundred times for coefficients a_{11}, a_{12} and a_{21}. The choice of stepsize affects the performance of the adaptive identification system. The small stepsize results in small bias and STD. However, the convergence time becomes long in this case. When stepsizes in the adaptive ILS algorithm and in the LMS algorithm are chosen as 5×10^{-3} and 10^{-2}, respectively, both have the same convergence time. Figure 1 gives an example to illustrate the evolution of adaptive identification coefficients.

We also evaluate the performance of the adaptive ILS algorithm and the LMS algorithm with different SNR and present the result in Figure 2. To make a reasonable comparison, we choose the stepsizes as 5×10^{-3} (LMS) and 10^{-2} (ILS) to let their convergence times approximately equal. From the figure, it is observed that the bias and the STD of the estimated coefficients via the adaptive ILS algorithm are generally smaller that those via the LMS algorithm, especially the estimation bias. When SNR is high, the superiority of the adaptive ILS algorithm to the LMS algorithm is not so much obvious for some coefficients, even the performance of the former is worse than that of the latter. When SNR is low, the adaptive ILS algorithm apparently outperforms the LMS method.

4. CONCLUSIONS

Bilinear models represent a class of nonlinear recursive systems and have been used in a variety of applications. However, the identification performance of the bilinear model can be limited in practice when conventional least squares techniques are used, as this leads to biased coefficient estimations. In this paper, a new algorithm called adaptive ILS algorithm is developed to estimate the coefficients of the bilinear systems. Compared with the LMS algorithm, the proposed algorithm can strongly reduce the estimation bias when large amplitude measurement noise exists and has superior identification performance.

5. REFERENCES

[1] T. Söderström and P. Stoica, “On some system identification techniques for adaptive filtering,” *IEEE Trans. Circuits Syst.*, vol. 35, pp.457-461, 1988.

[2] G. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting, and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[3] M. Maqusi, “Performance of baseband digital data transmission in nonlinear channels with memory,” *IEEE Trans. Commun.*, vol. 33, pp. 715-719, 1985.

[4] T. C. Tsai and D. Anastassiou, “Nonlinear recursive filters and applications to image processing,” *Proc. ICASSP'88*, pp. 828-831, 1988.

[5] M. B. Mattews and G. S. Moschytz, “The identification of nonlinear discrete-time fading-memory systems using neural network models,” *IEEE Trans. Circuits Syst.-II*, vol. 41, pp. 740-751, 1994.

[6] R. R. Mohler W. J. Kolodziej, “An overview of bilinear system theory and applications,” *IEEE Trans. Systems, Man, and Cybernetics*, vol. 10, pp. 683-688, 1980.

[7] F. Fnaiech and L. Ljung, “Recursive identification of bilinear systems,” *Int. J. Control*, vol. 45, pp. 453-470, 1987.

[8] J. Lee and V. J. Mattews, “Output-error LMS bilinear filters with stability monitoring,” *Proc. ICASSP'95*, pp. 965-968, 1987.

[9] V. J. Mattews, “Adaptive Polynomial Filters,” *IEEE Signal Processing Magazine*, pp. 10-26, July, 1991.

[10] W. A. Fuller, *Measurement Error Models*, New York: John Wiley, 1987.

Table 1: Bias, Standard Derivation (STD) and Convergence Time (T) of Estimated Coefficients in the Bilinear System.

Parameter	Method	stepsize=5 $\times 10^{-3}$			stepsize=10^{-2}			stepsize=5 $\times 10^{-2}$		
		Bias	STD	T	Bias	STD	T	Bias	STD	T
a_{11}	LMS	0.1309	0.0119	2100	0.1351	0.0217	1200			
(1.5)	ILS	0.0009	0.0021	5800	0.0012	0.0039	2800	0.0315	0.0681	600
a_{12}	LMS	0.1195	0.0117	3900	0.1198	0.0223	1800			
(-0.7)	ILS	0.0001	0.0016	9000	0.0005	0.0040	4200	0.0280	0.0613	1500
a_{20}	LMS	0.0031	0.0143	2000	0.0048	0.0158	800			
(0.5)	ILS	0.0049	0.0138	2400	0.0022	0.0114	2000	0.0054	0.0270	400
a_{21}	LMS	0.1111	0.0163	1900	0.1131	0.0242	500			
(0.8)	ILS	0.0040	0.0045	2200	0.0072	0.0127	800	0.0184	0.0423	300
b_{10}	LMS	0.0084	0.0171	1800	0.0094	0.0316	1000			
(0.24)	ILS	0.0009	0.0050	4800	0.0058	0.0126	2800	0.0192	0.0717	500

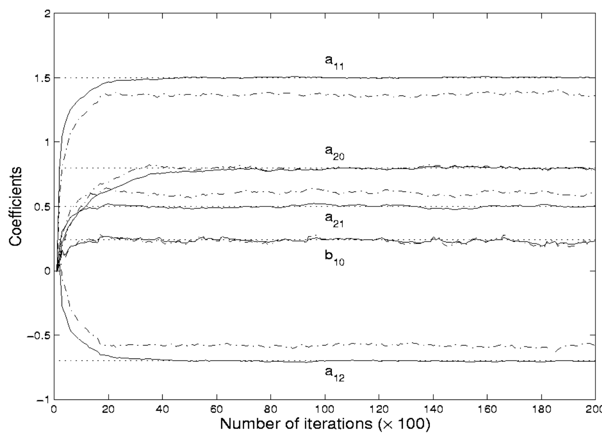
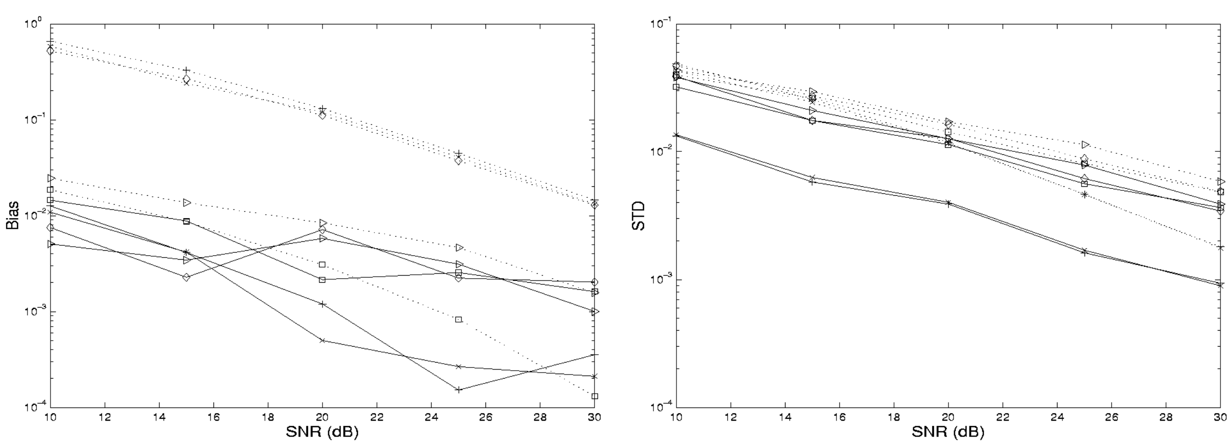


Figure 1: Evolution of Coefficients in the Adaptive Bilinear System (— : ILS, $\cdots$: LMS).



(a) (b)

Figure 2: Comparison of Bias and Standard Deviation Between the Adaptive ILS Algorithm and LMS Algorithm: (a) Bias; (b) Standard Deviation (— : ILS, $\cdots$: LMS; $+ : a_{11}$, $\times : a_{12}$, $\Box : a_{20}$, $\diamond : a_{11}$, $\triangleright : b_{10}$).