ADAPTIVE BILINEAR PREDICTORS

Junghsi Lee
Advanced Technology Center
Computer and Communication Research Laboratories
Industrial Technology Research Institute
Chutung, Hsinchu, Taiwan, R.O.C.

V. John Mathews
Department of Electrical Engineering
University of Utah
Salt Lake City, Utah 84112

ABSTRACT

This paper considers an extended recursive least squares (RLS) adaptive bilinear predictor. It is shown that the extended RLS adaptive bilinear predictor is guaranteed to be stable in the sense that the time average of the squared *a-posteriori* prediction error signal is bounded whenever the input signal is bounded in the same sense. It also shows that the *a-priori* prediction error itself is bounded whenever the desired signal is bounded. This paper also contains simulation results to demonstrate the usefulness of the extended RLS adaptive bilinear predictor.

1. INTRODUCTION

Linear prediction has found a variety of applications in many areas including those in spectral estimation, speech analysis/synthesis, control theory, image processing, communications, and economic forecasting [8, 15]. While linear models provide adequate performance in many applications, there are several instances where they can give misleading results since many systems are inherently nonlinear. In such cases, there may be advantages in considering nonlinear predictors. Consequently, there has been a large amount of activity in the recent past on fitting signals with nonlinear models. Examples of nonlinear models include threshold linear models, polynomial models and others [12]. The Volterra system model has built up some popularity in many areas of nonlinear filtering [6, 9, 11, 13]. Another polynomial model that has attracted some attention recently is the bilinear system model [4, 6-7, 9-10].

In this paper, we consider the prediction of a bilinear time series satisfying the following difference equation

$$\begin{aligned} x(n) &= \sum_{i=1}^{h} a_i \eta(n-i) + \sum_{i=1}^{m} b_i x(n-i) \\ &+ \sum_{i=1}^{r}\sum_{j=1}^{s} c_{i,j}(n)\eta(n-i)x(n-j), \qquad (1) \end{aligned}$$

where η(n) belongs to a stationary and zero-mean white process. In statistical literature, the above model is often referred to as the BL(m,h,s,r) model [14]. This model is particularly attractive since it shares several features with linear system models. In particular, if $c_{i,j} = 0$ for all i's and j's, we have the familiar ARMA model.

In spite of the simplicity of bilinear system models, there is a large class of nonlinear systems that can be adequately modeled as bilinear systems. For example, it has been shown under fairly mild conditions that a bilinear system with finite number of coefficients can be used to approximate any Volterra system with arbitrary precision [2]. Consequently, such system models have found a variety of applications including those in control systems, population models, biological systems, economic models, etc. An overview of continuous-time bilinear models and their applications can be found in [10].

An important issue associated with the bilinear system model is that of its stability. Due to the nonlinear structure, the stability problem is much more complicated than that associated with recursive linear systems. For example, it is possible to find bounded input signals that can cause the output of almost all bilinear systems to be unbounded. Some recent work in adaptive bilinear filtering that used the extended RLS adaptation of algorithm either did not address this very critical stability issue [1] or proposed considerably more complex variants of the recursive prediction error method - by employing extra Kalman filters - to handle the stability problem [4]. It was only very recently that the extended RLS adaptive bilinear filter was shown to be stable in the sense that the time average of the squared residual is bounded whenever the desired signal is bounded in the same sense [6-7].

In this paper, we provide some stability results for the extended RLS adaptive bilinear predictor. This result differs from the one in [6-7] in that the input signal that generates the desired signal is not observable in our case. The rest of the paper is organized as follows. Section II presents an extended RLS adaptive bilinear predictor. In Section III, we provide two stability results for the bilinear predictor. Section IV contains simulation results that demonstrate the usefulness of the extended RLS bilinear predictor. The concluding remarks are made in Section V.

2. THE EXTENDED RLS BILINEAR PREDICTOR

The task of the adaptive bilinear predictor is to adaptively predict the input signal $x(n)$, based on delayed versions of $x(n)$, using the bilinear system model. We consider a bilin-

0-7803-1775-0/94 $3.00 © 1994 IEEE

ear predictor satisfying the following difference equation:

$$\hat{x}(n) = \sum_{i=1}^{h} a_i e(n-i) + \sum_{i=1}^{m} b_i x(n-i) + \sum_{i=1}^{r}\sum_{j=1}^{s} c_{i,j}(n) e(n-i) x(n-j), \quad (2)$$

where $e(n)$ is the *a-posteriori* prediction error signal defined as

$$e(n) = x(n) - W_n^T Z_n. \quad (3)$$

Let us also define $\alpha(n)$, the *a-priori* prediction error signal as

$$\alpha(n) = x(n) - W_{n-1}^T Z_n. \quad (4)$$

In (3) and (4), coefficient vector W_n and data vector Z_n are defined as

$$W_n = [a_1(n), ..., a_h(n), b_1(n), ..., b_m(n), c_{1,1}(n), c_{1,2}(n), ..., c_{r,s}(n)]^T, \quad (5)$$

and

$$Z_n = [e(n-1), ..., e(n-h), x(n-1), ..., x(n-m), e(n-1)x(n-1), e(n-1)x(n-2), ..., e(n-r)x(n-s)]^T, \quad (6)$$

respectively.

The adaptive bilinear predictor may be implemented using several algorithms currently available in literature. In this paper, we will focus on the extended RLS bilinear predictor, mainly because of its good convergence properties and its guaranteed stability, which will be discussed in Section III. In the extended RLS approach, we solve for the coefficient vector W_n at each time instant, so that a weighted sum of the squared prediction errors, given by

$$J(n) = \sum_{k=0}^{n} \lambda^{n-k}(x(k) - W_n^T Z_k)^2 \quad (7)$$

is minimized. In (7), λ is the forgetting factor which is bounded above and below by 1 and 0, respectively. In real applications, we always set λ slightly less than 1 so that it can track slowly varying parameters in the input signal statistics.

It is well known that the optimum solution to the above problem is given by

$$\hat{W}_n = \Omega_n^{-1} P_n, \quad (8)$$

where

$$\Omega_n = \sum_{k=0}^{n} \lambda^{n-k} Z_k Z_k^T, \quad (9)$$

and

$$P_n = \sum_{k=0}^{n} \lambda^{n-k} Z_k x(k). \quad (10)$$

Recursive implementation of the above solution is straightforward. Much more efficient implementations of the above scheme are also possible.

The extended RLS adaptive bilinear predictor is an iterative realization of the above solution, and may be summarized as follows:

For each instant of time, $n = 1, 2, ...$

1. $\hat{x}(n) = W_{n-1}^T Z_n$
2. $\alpha(n) = x(n) - \hat{x}(n)$
3. $K_n = \frac{\lambda^{-1}\Omega_{n-1}^{-1} Z_n}{1+\lambda^{-1} Z_n^T \Omega_{n-1}^{-1} Z_n}$
4. $W_n = W_{n-1} + K_n \alpha(n)$
5. $\Omega_n^{-1} = \lambda^{-1}\Omega_{n-1}^{-1} - \lambda^{-1} K_n Z_n^T \Omega_{n-1}^{-1}$

The applicability of the extended RLS algorithm requires that Ω_n is nonsingular all the time. It is known that Ω_n may become singular if the input signal is not persistently exciting. Dasgupta has studied the persistent excitation issue of bilinear systems [3]. In this paper, we assume that there exists $\sigma_1, \sigma_2 > 0$, such that

$$\sigma_1 I < \Omega_n < \sigma_2 I \quad (11)$$

for all n.

As previously mentioned, adaptive bilinear predictors are not guaranteed to be stable. While not very much can be stated about the overall stability of the bilinear models generated by the extended RLS approach, in the next section, we present two stability theorems. The first theorem is an extension to the one given in [7].

3. STABILITY RESULTS

Theorem 1 *The extended RLS adaptive bilinear predictor is stable in the sense that* $\frac{1}{n}\sum_{k=0}^{n} e^2(k)$ *is bounded whenever* $\frac{1}{n}\sum_{k=0}^{n} x^2(k)$ *is bounded.*

Although the construction of data vector Z_n is slightly different from that in [7], we can still use techniques similar to those used in [7] to prove theorem 1. The extension of the proofs in [7] is straightforward, and is omitted here.

In the following, we present another stability result under the situation that the forgetting factor is less than 1.

Theorem 2 *The extended RLS adaptive bilinear predictor that employs a forgetting factor less than 1 is stable in the sense that the a-priori prediction error* $\alpha(k)$ *is bounded whenever the input signal* $x(k)$ *is bounded.*

Proof We can again extend a result in [7] to show that

$$e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k) = x^2(k) + \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1} - \hat{W}_k^T \Omega_k \hat{W}_k. \quad (12)$$

Because Ω_k is nonnegative definite, the left-hand-side, and hence the right-hand-side of (12) is positive. Therefore, we have the following inequality.

$$\hat{W}_k^T \Omega_k \hat{W}_k \le x^2(k) + \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1}. \quad (13)$$

Let D denote the least upper bound of $x^2(k)$. By applying the above inequality iteratively k times, we get the following result:

$$\begin{aligned}\hat{W}_k^T \Omega_k \hat{W}_k &\leq D + \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1} \\ &\leq (1 + \lambda + \ldots + \lambda^{k-1}) D. \qquad (14)\end{aligned}$$

Here, without loss of generality, we assume that $\hat{W}_{-1}$ is a zero vector, i.e., the coefficient vector is initialized with zero values. Because, $\hat{W}_k^T \Omega_k \hat{W}_k$, the last term on the right-hand-side of (12) is nonnegative, we have

$$\begin{aligned}e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k) &\leq x^2(k) + \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1} \\ &\leq D + \lambda(1 + \lambda + \ldots + \lambda^{k-2}) D \\ &\leq \left\{\frac{1}{1-\lambda}\right\} D. \qquad (15)\end{aligned}$$

Since we have assumed that λ is less than 1, it follows that $e^2(k)$, and hence $e(k)$ is finite.

It is straightforward to show that

$$\alpha(k) = e(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k). \qquad (16)$$

Since $1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k$ is finite [5], it follows that $\alpha(k)$ also is bounded. This completes the proof.

4. SIMULATION RESULTS

In this section, we present some simulation results that demonstrate the usefulness of the extended RLS adaptive bilinear predictor. The problem considered is that of estimating the parameters of a bilinear time series using the extended RLS algorithm. The bilinear time series model is governed by the following equation

$$\begin{aligned}x(n) &= \sum_{i=1}^{3} a_i^0 \eta(n-i) + \sum_{i=1}^{2} b_i^0 x(n-i) \\ &+ \sum_{i=1}^{3} \sum_{j=1}^{2} c_{i,j}^0 \eta(n-i) x(n-j). \qquad (17)\end{aligned}$$

The coefficients are given in Table 1. In (17), $\eta(n)$ belongs to a stationary and zero-mean white Gaussian process with variance 0.05. The adaptive bilinear predictor was run with the same structure and the same number of coefficients as that of (17). All the results presented are ensemble averages over 50 independent runs. The steady state squared prediction error and the predictor coefficients were measured by time-averaging the corresponding ensemble average in the range [17001, 20000] and were given in Table 1. These are then compared to the true values of the coefficients. Note that the mean values of the estimated parameters show very good match with the actual parameters. We also illustrate the evolution of some predictor coefficients in Figure 1. Note that coefficients a_i's and $c_{i,j}$'s converge slower than the coefficients b_i's. This is so because we use the prediction error signals in place of the unobservable input $\eta(n)$.

We also evaluated the performance of the adaptive bilinear predictor for the case when $\eta(n)$ belongs to a uniformly distributed random process. Results comparable to those presented here were obtained in these experiments also.

5. CONCLUDING REMARKS

In this paper, we presented an extended RLS adaptive bilinear predictor. We showed that the predictor is stable in the sense that the time average of the squared *a-posteriori* prediction error is bounded whenever the input signal is bounded in the same sense. We also showed that the *a-priori* prediction error is bounded if the input signal is bounded. This paper also provided simulation results that demonstrate the usefulness of the adaptive bilinear predictor.

ACKNOWLEDGEMENTS

This work was supported in part by the Project No. 37H3100 conducted by the ITRI under sponsorship of the Minister of Economic Affairs, R.O.C. and by the National Science Foundation under Grant MIP -8922146.

REFERENCES

[1] H. K. Baik and V. J. Mathews, "Adaptive lattice bilinear filters," *IEEE Transactions on Signal Processing*, Vol. 41, No. 6, pp. 2033-2046, June 1993.

[2] R. W. Brockett, "Volterra series and geometric control theory," *Automation*, Vol. 12, pp. 167-176, 1976.

[3] S. Dasgupta, Y. Shrivastava and G. Krenzer, "Persistent excitation in bilinear systems," *IEEE Transactions on Automatic Control*, Vol. 36, No. 3, pp. 305-313, March 1991.

[4] F. Fnaiech and L. Ljung, "Recursive identification of bilinear systems," *International Journal of Control*, Vol. 45, No. 2, pp. 453-470, 1987.

[5] S. Haykin, *Adaptive filter theory*, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[6] J. Lee, *Adaptive polynomial filtering algorithms*, Ph.D. Dissertation, University of Utah, 1992.

[7] J. Lee and V. J. Mathews, "On the extended RLS adaptive bilinear filters," in *Proceedings 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, Minnesota, 1993, pp. III 428-431.

[8] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, Vol. 63, No. 4, pp. 561-580, April 1975.

[9] V. J. Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Magazine*, pp. 10-26, July 1991.

[10] R. R. Mohler and W. J. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-10, No. 10, pp. 683-688, 1980.

[11] E. Mumolo and D. Francescato, "Adaptive predictive coding of speech by means of Volterra predictors," in *Proceedings 1993 IEEE Winter Workshop on Nonlinear Digital Signal Processing*, Tampere, Finland, 1993, pp.2.1 4.1-4.4.

[12] M. B. Priestley, *Non-linear and Non-stationary Time Series Analysis*, Academic Press, London, 1988.

[13] G. L. Sicuranza and G. Ramponi, "Adaptive nonlinear prediction of TV image sequences," *IEEE Electronics Letters*, Vol. 25, No. 8, pp. 526-527, April 1989.

[14] T. Subba Rao, "On the theory of bilinear time series models," *Journal of Royal Statistical Society (b)*, Vol. 43,No. 2, pp. 244-255, 1981.

[15] J. R. Zeidler, "Performance analysis of LMS adaptive prediction filters," *Proceedings of the IEEE* , Vol. 78, No. 12, pp. 1781-1806, December 1990.

Parameter	λ=0.999		λ=0.9995		λ=0.9999	
	mean	variance (10^{-3})	mean	variance (10^{-3})	mean	variance (10^{-3})
$a_1 = 0.7$	0.6797	0.665	0.6882	0.427	0.6833	0.218
$a_2 = 0.5$	0.4848	0.345	0.4916	0.223	0.4877	0.086
$a_3 = -0.3$	-0.2998	0.503	-0.3011	0.337	-0.3143	0.168
$b_1 = -1.4$	-1.3874	0.646	-1.3923	0.413	-1.3879	0.186
$b_2 = -0.48$	-0.4676	0.592	-0.4727	0.382	-0.4698	0.176
$c_{1,1} = 0.5$	0.4949	2.004	0.4988	1.318	0.5017	0.518
$c_{1,2} = 0.5$	0.4964	2.257	0.4999	1.526	0.5008	0.613
$c_{2,1} = 0.7$	0.6780	5.584	0.6929	3.626	0.6785	1.069
$c_{2,2} = 0.7$	0.6769	5.145	0.6910	3.278	0.6780	0.926
$c_{3,1} = -0.6$	-0.6015	1.946	-0.5994	1.256	-0.6101	0.686
$c_{3,2} = -0.6$	-0.6007	1.677	-0.5992	1.126	-0.6125	0.706
$\sigma^2 = 0.05$	0.0513	0.00176	0.0506	0.00170	0.0502	0.00161

Table 1. Steady state predictor coefficients and squared *a-priori* prediction error.

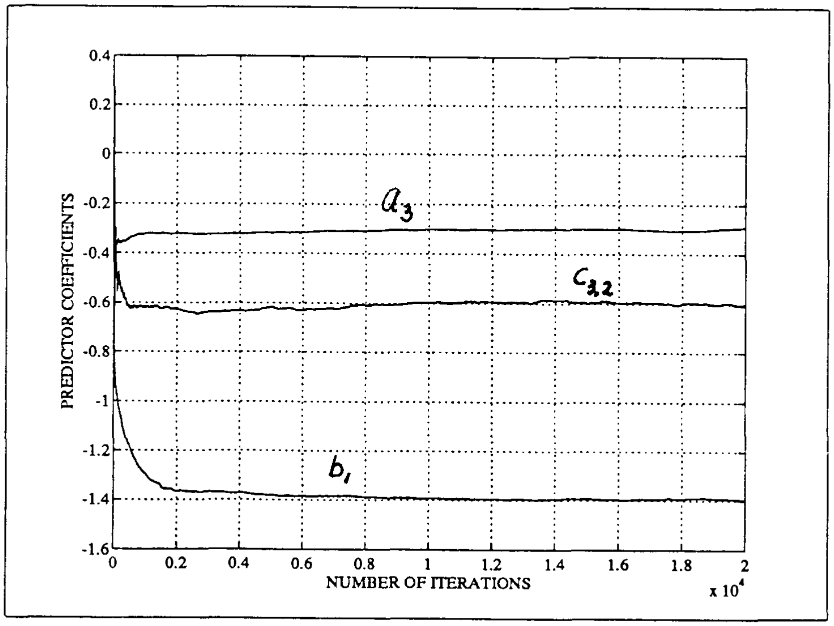

Figure 1: Evolution of predictor coefficients.