A Stability Result for RLS Adaptive Bilinear Filters

Junghsi Lee and V. John Mathews

Abstract—**This letter considers recursive least squares (RLS) adaptive nonlinear filtering using bilinear system models. It is shown that the extended RLS adaptive bilinear filter, as well as the equation-error RLS adaptive bilinear filter, are guaranteed to be stable in the sense that the time average of the squared estimation error is bounded whenever the underlying process that generates the input signals is stable in the same sense.**

I. Introduction

THE Volterra system model is extremely popular in adaptive nonlinear filtering. Although it is very useful in many situations, it often requires a large number of coefficients to characterize many nonlinear processes. One approach to alleviate this problem is to use recursive nonlinear models.

This letter considers a bilinear system whose input-output relationship is given by

$$y(n) = \sum_{i=0}^{l} a_i^o x(n-i) + \sum_{i=1}^{m} b_i^o y(n-i) + \sum_{i=0}^{r}\sum_{j=1}^{s} c_{i,j}^o x(n-i)y(n-j) \tag{1}$$

where a_i^o's, b_i^o's, and $c_{i,j}^o$'s represent the coefficients of the system. It has been shown under relatively mild conditions that a large class of nonlinear systems can be approximated with arbitrary precision using bilinear models with finite number of coefficients [2], [9]. In addition, many concepts associated with linear systems can be extended to the bilinear case. Several applications and properties of bilinear systems are reviewed in [3] and [9].

An important issue associated with the bilinear system model is that of its stability. It is possible to find bounded input signals that can cause the output of almost all bilinear systems to be unbounded. This is probably the main reason why only very limited work on the theory of adaptive bilinear filtering has been done. Some recent work in adaptive bilinear filtering uses the extended RLS adaptation algorithm or its variations [1], [5]. Although Fnaiech and Ljung [5] proposed considerably more complex variants of the recursive prediction error method (RPEM)—by employing extra Kalman filters—to handle the stability problem, Baik and Mathews [1] did not address this very critical stability issue. In this letter, we consider two RLS-type adaptive bilinear filters: the extended (output-error) RLS and the equation-error RLS bilinear filters. In spite of the recursive system model employed in these adaptive filters, we show that both algorithms are guaranteed to be stable in some sense whenever the underlying process that generates the input signals is stable in the same sense.

Manuscript received July 12, 1994; approved September 23, 1994. This work was supported by the National Science Foundation under Grant MIP-8922146, a Departmental Grant by the IBM Corporation, and by the Project No. 37H3100 conducted by the ITRI under sponsorship of the Minister of Economic Affairs, R.O.C. The associate editor coordinating the review of this letter and approving it for publication was Dr. P. Duhamel.

J. Lee is with the Advanced Technology Center, Computer and Communication Research Laboratories, Industrial Technology Research Institute, Chutung, Hsinchu, Taiwan, R.O.C.

V. J. Mathews is with the Department of Electrical Engineering, University of Utah, Salt Lake City, UT, 84112 USA.

IEEE Log Number 9407191.

II. The RLS Bilinear Filters

Let $d(n)$ and $x(n)$ represent the desired response signal and the input signal, respectively, to the adaptive filter. Let $\hat{d}(n)$ denote the output of the adaptive filter at time n, i.e.

$$\hat{d}(n) = \sum_{i=0}^{l} a_i(n)x(n-i) + \sum_{i=1}^{m} b_i(n)\tilde{d}(n-i) + \sum_{i=0}^{r}\sum_{j=1}^{s} c_{i,j}(n)x(n-i)\tilde{d}(n-j). \tag{2}$$

The equation-error approach simply uses samples of $d(k)$ for $\tilde{d}(k)$, whereas the output-error approach feeds back samples of $\hat{d}(k)$ for $\tilde{d}(k)$. Equation-error algorithms are straightforward to develop, and the mean-squared estimation error surface has a unique minimum. However, this minimum may not be at the correct solution to the problem in practice. The output-error algorithm is capable of providing unbiased estimates. However, since $d(k) - \hat{d}(k)$ is not a linear function of the filter coefficients, output-error algorithms can converge to a local minimum of the mean-squared error surface.

RLS bilinear filters minimize $J(n)$

$$J(n) = \sum_{k=0}^{n} \lambda^{n-k}(d(k) - \hat{d}_n(k))^2 \tag{3}$$

at each time n, where λ is the forgetting factor $(0 < \lambda < 1)$ that controls the rate at which the adaptive filter tracks time-varying environments. In (3), $\hat{d}_n(k)$ is defined as

$$\hat{d}_n(k) = \sum_{i=0}^{l} a_i(n)x(k-i) + \sum_{i=1}^{m} b_i(n)\tilde{d}(k-i) + \sum_{i=0}^{r}\sum_{j=1}^{s} c_{i,j}(n)x(k-i)\tilde{d}(k-j) \tag{4}$$

and is the estimate of $d(k)$ obtained using the coefficients of the adaptive filter at time n. Note that $\hat{d}(k) = \hat{d}_k(k)$.

Let the input vector Z_n and the coefficient vector W_n at time n be defined as

$$Z_n = [x(n), \ldots, x(n-l), \tilde{d}(n-1), \ldots, \tilde{d}(n-m), x(n)\tilde{d}(n-1), \ldots, x(n)\tilde{d}(n-s), \ldots, x(n-r)\tilde{d}(n-s)]^T \tag{5}$$

and

$$W_n = [a_o(n), \dots, a_l(n), b_1(n), \dots, b_m(n), \\ c_{0,1}(n), \dots, c_{0,s}(n), \dots, c_{r,s}(n)]^T \quad (6)$$

respectively. The optimal solution W_n, which minimizes $J(n)$ defined in (3), is given by

$$\hat{W}_n = \Omega_n^{-1} P_n, \quad (7)$$

where

$$\Omega_n = \sum_{k=0}^{n} \lambda^{n-k} Z_k Z_k^T \quad (8)$$

and

$$P_n = \sum_{k=0}^{n} \lambda^{n-k} Z_k d(k). \quad (9)$$

Here, Ω_n is the LS autocorrelation matrix of the input vector Z_n, and P_n is the LS cross-correlation vector between the input vector Z_n and the desired response $d(n)$. We will assume that Ω_n is nonsingular at all times. The implications of this assumption will be discussed later. Recursive implementations of the above solutions can be easily derived [1], [5]. The following theorem, which is an extension to the one given in [7] for linear system models, is the main result of this letter.

III. The Stability Theorem

Theorem: Let $e(k)$ be the *a posteriori* estimation error of the adaptive bilinear filter, i.e., $e(k) = d(k) - \hat{d}(k)$. The RLS adaptive bilinear filtering algorithms considered in this letter provide a stable output whenever the desired response signal is stable in the sense that $\frac{1}{n+1}\sum_{k=0}^{n} e^2(k)$ is bounded whenever $\frac{1}{n+1}\sum_{k=0}^{n} d^2(k)$ is bounded.

Proof: The theorem will be proved by showing that

$$\frac{1}{n+1}\sum_{k=0}^{n} e^2(k) \le \frac{1}{n+1}\sum_{k=0}^{n} d^2(k) + \frac{\lambda}{n+1}\hat{W}_{-1}^T \Omega_{-1} \hat{W}_{-1}. \quad (10)$$

Let $\alpha(k)$ denote the *a priori* estimation error, i.e., $\alpha(k) = d(k) - \hat{W}_{k-1}^T Z_k$. It is easy to show that

$$P_k = \lambda P_{k-1} + Z_k d(k), \quad (11)$$

$$\hat{W}_k = \lambda \Omega_k^{-1} P_{k-1} + \Omega_k^{-1} Z_k d(k), \quad (12)$$

$$\hat{W}_k = \hat{W}_{k-1} + \Omega_k^{-1} Z_k \alpha(k) \quad (13)$$

and

$$\alpha(k) = e(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k). \quad (14)$$

By using (7), (13), and (11), it follows that

$$\begin{aligned} \hat{W}_k^T \Omega_k \hat{W}_k &= \hat{W}_k^T P_k \\ &= (\hat{W}_{k-1}^T + Z_k^T \Omega_k^{-1} \alpha(k))(\lambda P_{k-1} + Z_k d(k)) \\ &= \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1} + \lambda \alpha(k) Z_k^T \Omega_k^{-1} P_{k-1} \\ &\quad + \hat{W}_{k-1}^T Z_k d(k) + \alpha(k) Z_k^T \Omega_k^{-1} Z_k d(k). \quad (15) \end{aligned}$$

From (12) and the definition of $e(k)$, we can show that

$$\begin{aligned} &\lambda \alpha(k) Z_k^T \Omega_k^{-1} P_{k-1} + \alpha(k) Z_k^T \Omega_k^{-1} Z_k d(k) \\ &= \alpha(k) Z_k^T (\lambda \Omega_k^{-1} P_{k-1} + \Omega_k^{-1} Z_k d(k)) \\ &= \alpha(k) Z_k^T \hat{W}_k \\ &= \alpha(k)(d(k) - e(k)). \quad (16) \end{aligned}$$

Similarly

$$\hat{W}_{k-1}^T Z_k d(k) = (d(k) - \alpha(k)) d(k). \quad (17)$$

Substituting (16), (17), and (14) into (15) gives

$$\begin{aligned} \hat{W}_k^T \Omega_k \hat{W}_k &= \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1} + d^2(k) \\ &\quad - e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k). \quad (18) \end{aligned}$$

Rewrite (18) as

$$\begin{aligned} &e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k) \\ &= d^2(k) - \hat{W}_k^T \Omega_k \hat{W}_k + \lambda \hat{W}_{k-1}^T \Omega_{k-1} \hat{W}_{k-1}. \quad (19) \end{aligned}$$

Summing both sides of (19) from $k = 0$ to $k = n$ and then dividing them by $n + 1$ gives

$$\begin{aligned} &\frac{1}{n+1}\sum_{k=0}^{n} \{e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k)\} \\ &= \frac{1}{n+1}\sum_{k=0}^{n} d^2(k) + \frac{\lambda}{n+1}\hat{W}_{-1}^T \Omega_{-1} \hat{W}_{-1} \\ &\quad - \left\{\frac{(1-\lambda)}{n+1}\sum_{k=0}^{n-1} \hat{W}_k^T \Omega_k \hat{W}_k + \frac{1}{n+1}\hat{W}_n^T \Omega_n \hat{W}_n\right\}. \quad (20) \end{aligned}$$

Because Ω_n is positive definite and λ is no greater than unity, the last term on the right-hand-side of (20) is nonnegative. It is then easy to see that

$$\begin{aligned} \frac{1}{n+1}\sum_{k=0}^{n} e^2(k) &\le \frac{1}{n+1}\sum_{k=0}^{n} \{e^2(k)(1 + \lambda^{-1} Z_k^T \Omega_{k-1}^{-1} Z_k)\} \\ &\le \frac{1}{n+1}\sum_{k=0}^{n} d^2(k) + \frac{\lambda}{n+1}\hat{W}_{-1}^T \Omega_{-1} \hat{W}_{-1}. \quad (21) \end{aligned}$$

This completes the proof.

Remark: The derivation assumes that the LS autocorrelation matrix Ω_k is nonsingular at all times. This means that the input signal $x(k)$ satisfies certain persistence of excitation conditions. For system identification problems, Dasgupta *et al.* [4] have derived conditions on input signals that guarantees persistent excitation of bilinear systems. In an application where the users do not have control over the selection of the input signals, we may modify the adaptive filtering algorithms as in [7] to ensure positive definiteness of Ω_k.

IV. Concluding Remarks

This letter showed that the extended RLS bilinear filter, as well as the equation-error RLS bilinear filter, are stable in the sense that the time average of the sqaured estimation error is bounded whenever the desired response signal is stable in the same sense. The modeling efficiency and the guaranteed stability of the extended RLS adaptive bilinear filters should make them attractive choices in nonlinear filtering applications.

The stability result described in this letter can easily be extended to problems involving adaptive bilinear prediction [6]. With some simple additional assumptions, we can also show that extended RLS bilinear predictors employed in predictive coding applications will operate in a stable manner [8].

References

[1] H. K. Baik and V. J. Mathews, "Adaptive lattice bilinear filters," *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2033–2047, June 1993.

[2] R. W. Brockett, "Volterra series and geometric control theory," *Automatica*, vol. 12, pp. 167–176, 1976.

[3] C. Bruni, G. Di Pillo, and G. Koch, "Bilinear systems: An appealing class of 'nearly linear systems' in theory and applications," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 4, pp. 334–348, Aug. 1974.

[4] S. Dasgupta, Y. Shrivastava, and G. Krenzer, "Persistent excitation in bilinear systems," *IEEE Trans. Automat. Contr.*, vol. 36, no. 3, pp. 305–313, Mar. 1991.

[5] F. Fnaiech and L. Ljung, "Recursive identification of bilinear systems," *Int. J. Contr.*, vol. 45, no. 2, pp. 453–470, 1987.

[6] J. Lee and V. J. Mathews, "Adaptive bilinear predictors," in *Proc. 1994 IEEE Int. Conf. Acoust., Speech, Signal Processing* (Adelaide, South Australia), Apr. 1994, pp. 489–492.

[7] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification.* Cambridge, MA: MIT Press, 1983.

[8] V. J. Mathews and J. Lee, "Techniques for nonlinear time series analysis," in *Proc. 27th Ann. Asilomar Conf. Syst., Signals Comput.* (Pacific Grove, CA), Nov. 1993, pp. 1016–1020.

[9] R. R. Mohler and W. J. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Trans. Syst., Man Cybern.* 1980.