Least Mean Square Algorithm for Unbiased Impulse Response Estimation

H. C. So

Department of Computer Engineering & Information Technology
City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
e-mail: ithcso@cityu.edu.hk

Abstract - In this paper, a least mean square (LMS) type algorithm is devised for unbiased system identification in the presence of white input and output noise, assuming that the ratio of the noise powers is known. The proposed approach aims to minimize the mean square value of the equation-error function under a constant-norm constraint, and is equivalent to minimizing a modified mean square error function. Analysis of the algorithm shows that the estimates will converge to the true values in the mean sense. Computer simulations are included to corroborate the theoretical development and to evaluate the impulse response estimation performance of the LMS algorithm.

I. Introduction

Estimation of the impulse response of an unknown system from its measured input and output has found many applications in signal processing, communications and control [1]-[3]. Many estimation algorithms based on least squares (LS) and least mean square (LMS) techniques [4]-[9] have been developed for accurate system parameter estimation if additive noise exists only at the output. However, noise-free input measurements are difficult to obtain in many practical situations because these signals are also corrupted by quantization noise, self-noise and/or other additive noise [10], and neglecting the presence of input noise will lead to biased impulse response estimates.

The total least squares (TLS) approach [11] seems to be an appropriate choice for system identification with noisy input and output because it considers both stimulation and response errors. Computationally attractive adaptive impulse response estimation schemes employing the TLS criterion include the recursive total least squares (RTLS) [10], constrained anti-Hebbian (CAH) [12], and total least mean square (TLMS) [13] algorithms. Nevertheless, these methods are restricted for parameter estimation of finite impulse response (FIR) systems, and assume that the input and output noise powers are identical. For the general case of infinite impulse response (IIR) system modeling, the Koopmans-

Levin method [14], which is based on spectral decomposition of a covariance matrix formed using the input-output data and a modified LS algorithm [15] that requires estimation of the measurement noise variances, can be used to provide reliable estimates but they involve intensive computations. In this paper, we develop an LMS approach for unbiased impulse response estimation for which real-time implementation is allowed, assuming that the power ratio of the input and output interference is known [14].

The organization of the paper is as follows. In Section II, the problem of impulse response estimation using noisy input and output measurements is formulated. It is shown that bias-free estimation is possible if the minimization is subject to a constant-norm constraint. The constrained problem is then converted into an unconstrained minimization of a weighted mean square error function by an LMS algorithm. Learning characteristics of the impulse response estimates are also derived. Simulation results are presented in Section III to corroborate the theoretical analyses and to evaluate the system identification performance of the new approach. Finally, conclusions are drawn in Section IV.

II. Bias-Free LMS Algorithm

The block diagram for adaptive equation-error IIR filtering is depicted in Figure 1. Without loss of generality, we consider that the unknown system $H(z)$ is stable as well as causal and has the form

$$H(z) = \frac{B(z)}{A(z)} \quad (1)$$

where

$$A(z) = 1 + \sum_{i=1}^{M-1} a_i z^{-i} \quad (2)$$
$$B(z) = \sum_{j=0}^{N-1} b_j z^{-j}$$

The orders of the denominator and numerator polynomials, M and N , respectively, are assumed to be known. The observed system input and output are

$$x(k) = s(k) + n_i(k) \quad (3)$$
$$r(k) = d(k) + n_o(k)$$

where $s(k)$ and $d(k)$ denote the noise-free input and output, respectively, while $n_i(k)$ and $n_o(k)$ represent the input and output measurement noise which are independent of $s(k)$. It is assumed that $n_i(k)$ and $n_o(k)$ are uncorrelated white processes with variances $\sigma_{n_i}^2$ and $\sigma_{n_o}^2$, respectively, and the ratio of the noise powers, say, $\gamma = \sigma_{n_i}^2 / \sigma_{n_o}^2$, is available. Our goal is to find a_i , $i = 1, 2, \dots, M-1$ and b_j , $j = 0, 1, \dots, N-1$, from $x(k)$ and $r(k)$ using the adaptive FIR filters $\hat{A}(z) = 1 + \sum_{i=1}^{M-1} \hat{a}_i z^{-i}$ and $\hat{B}(z) = \sum_{j=0}^{N-1} \hat{b}_j z^{-j}$, where $\{\hat{a}_i\}$ and $\{\hat{b}_j\}$ are the parameter estimates.

In the equation-error formulation, the output error function $e(k)$ is computed from

$$e(k) = r(k) + \sum_{i=1}^{M-1} \hat{a}_i r(k-i) - \sum_{j=0}^{N-1} \hat{b}_j x(k-j) \quad (4)$$

It is easy to show that the mean square value of $e(k)$, $E\{e^2(k)\}$, is

$$\begin{aligned} E\{e^2(k)\} &= E \left\{ \left(d(k) + \sum_{i=1}^{M-1} \hat{a}_i d(k-i) - \sum_{j=0}^{N-1} \hat{b}_j s(k-j) \right)^2 \right\} \\ &\quad + \sigma_{n_o}^2 \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right) \end{aligned} \quad (5)$$

In the presence of input and/or output noise, it is obvious that minimizing $E\{e^2(k)\}$ with respect to $\{\hat{a}_i\}$ and $\{\hat{b}_j\}$ cannot give unbiased impulse response estimates because the noise components of (5) are functions of $\{a_i\}$ and $\{b_j\}$. Using the idea of [6], if we minimize $E\{e^2(k)\}$ subject to the constraint that $\left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2\right)$ is a positive constant, say, G , then unbiased solution can be attained. This optimization problem is equivalent to minimize

$$\frac{E\{e^2(k)\}}{G}$$

or

$$\begin{aligned} \frac{1}{G} E \left\{ \left(d(k) + \sum_{i=1}^{M-1} \hat{a}_i d(k-i) - \sum_{j=0}^{N-1} \hat{b}_j s(k-j) \right)^2 \right\} \\ + \frac{\sigma_{n_o}^2}{G} \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right) \\ \text{subject to} \quad \frac{1}{G} \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2 \right) = 1 \end{aligned} \quad (6)$$

Since the variables $\{a_i\}$ and $\{b_j\}$ in (6) appear in the first and second degrees only, the error surface is quadratic and thus it is unimodal [5]. This implies that gradient search approach can be utilized for the constrained minimization. From (6),

$$G = \frac{1}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2} \quad (7)$$

As a result, (6) is identical to an unconstrained optimization of minimizing a modified error function, $E\{\zeta^2(k)\}$, of the form

$$E\{\zeta^2(k)\} = \frac{E\{e^2(k)\}}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2 + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2} \quad (8)$$

In our study, the computationally attractive LMS algorithm is used to estimate $H(z)$ on a sample-by-sample basis. Based on (8), the instantaneous value of $\zeta(k)$ is

$$\zeta(k) = \frac{e(k)}{\sqrt{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)}} \quad (9)$$

where $\hat{a}_i(k)$, $i = 1, 2, \dots, M-1$, $\hat{b}_j(k)$, $j = 0, 1, \dots, N-1$, denote the estimates of $\{a_i\}$ and $\{b_j\}$ at time k , while $e(k)$ is calculated using the instantaneous parameter estimates. The bias-free least mean square (BFLMS) updating equations for unbiased system identification are then derived as

$$\begin{aligned} \hat{a}_i(k+1) &= \hat{a}_i(k) - \frac{\mu_a \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k) \right)}{2} \frac{\partial \zeta^2(k)}{\partial \hat{a}_i(k)} \\ &= \hat{a}_i(k) - \mu_a e(k) \cdot \left[r(k-i) - \hat{a}_i(k) \frac{e(k)}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)} \right], \\ &\quad i = 1, 2, \dots, M-1 \end{aligned} \quad (10)$$

and

$$\begin{aligned} \hat{b}_j(k+1) &= \hat{b}_j(k) - \frac{\mu_b \left(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k) \right)}{2} \frac{\partial \zeta^2(k)}{\partial \hat{b}_j(k)} \\ &= \hat{b}_j(k) + \mu_b e(k) \cdot \left[x(k-j) + \gamma \hat{b}_j(k) \frac{e(k)}{1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)} \right], \\ &\quad j = 0, 1, \dots, N-1 \end{aligned} \quad (11)$$

where μ_a and μ_b are positive scalars that govern convergence rate and stability of the adaptive algorithm. We note that (10) and (11) are a modification of the equation-error scheme where bias-removal terms are added to the update of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$. At each sampling interval, only $(4M + 4N - 2)$ multiplications, 1 division and $(4M + 4N - 4)$ additions are required in the BFLMS algorithm.

Assume the fluctuations in the coefficient errors are small so that the elements of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ can be replaced by their true values and $E\{e^2(k)\} \approx e^2(k)$. Then $e^2(k)/(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)) \approx \sigma_{n_o}^2$. Together with the independence assumption [5] and taking expectations on both sides of (10) and (11), the learning behaviour of the BFLMS estimates can be obtained from

$$\begin{aligned} & E\{\hat{a}_i(k+1)\} \\ &= E\{\hat{a}_i(k)\} - \mu_a \sum_{l=1}^{M-1} (E\{\hat{a}_l(k)\} - a_l) r_{dd}(l-i) + \\ & \quad \mu_a \sum_{l=i}^{N-1} (E\{\hat{b}_l(k)\} - b_l) r_{ds}(l-i), \\ & \quad i = 1, 2, \dots, M-1 \end{aligned} \quad (12)$$

and

$$\begin{aligned} & E\{\hat{b}_j(k+1)\} \\ &= E\{\hat{b}_j(k)\} + \mu_b \sum_{l=1}^j (E\{\hat{a}_l(k)\} - a_l) r_{ds}(j-l) - \\ & \quad \mu_b \sum_{l=0}^{N-1} (E\{\hat{b}_l(k)\} - b_l) r_{ss}(j-l), \\ & \quad j = 0, 1, \dots, N-1 \end{aligned} \quad (13)$$

where $r_{ss}(\tau) = E\{s(k)s(k-\tau)\}$, $r_{dd}(\tau) = E\{d(k)d(k-\tau)\}$ and $r_{ds}(\tau) = E\{d(k)s(k-\tau)\}$ represent the autocorrelation functions of $s(k)$ and $d(k)$ and their cross-correlation function, respectively. It can be seen that closed form solutions for $E\{\hat{a}_i(k)\}$ and $E\{\hat{b}_j(k)\}$ are not generally available because the convergence characteristic of each parameter estimate is dependent on all autocorrelation/crosscorrelation functions, although they can be determined by brute force if the signal statistics are known. Nevertheless, we can see that $\hat{a}_i(k) \rightarrow a_i$, $i = 1, 2, \dots, M-1$, and $\hat{b}_j(k) \rightarrow b_j$, $j = 1, 2, \dots, N-1$, as $k \rightarrow \infty$ in the mean sense, provided that $0 < \mu_a < 2/r_{dd}(0)$ and $0 < \mu_b < 2/r_{ss}(0)$, that is, there is convergence.

III. Simulation Results

Computer simulations had been conducted to corroborate the theoretical derivations and to evaluate the

performance of the BFLMS algorithm for estimating IIR system parameters using noisy measurements. The noise-free input $s(k)$ was a white Gaussian process of unity power and the measurement noises, $n_i(k)$ and $n_o(k)$, were independent white uniform random variables. All simulation results provided were averages of 200 independent runs.

We evaluated the system identification performance of the BFLMS algorithm for an IIR system with $M = 3$ and $N = 2$. The system parameters were given by $a_1 = -0.3$, $a_2 = -0.54$, $b_0 = 1.0$ and $b_1 = -2.4$. All filter coefficients were set to zero initially and $\mu_a = \mu_b = 0.002$ were assigned. Figure 2 and Figure 3 show the learning trajectories of $\{\hat{a}_i(k)\}$ and $\{\hat{b}_j(k)\}$ at $\sigma_{n_i}^2 = 0.1$ and $\sigma_{n_o}^2 = 1.0$. It is observed that $\hat{a}_1(k) \rightarrow -0.3$ and $\hat{b}_0(k) \rightarrow 1.0$ at approximately the 2000th iteration while $\hat{a}_2(k)$ and $\hat{b}_1(k)$ converged to their desired values at the 1000th and the 3000th iteration, respectively. The learning characteristics of the parameter estimates also conformed to their predicted trajectories except for the transient behaviour of $\hat{a}_2(k)$. This discrepancy was due to the assumption of $e^2(k)/(1 + \sum_{i=1}^{M-1} \hat{a}_i^2(k) + \gamma \sum_{j=0}^{N-1} \hat{b}_j^2(k)) \approx \sigma_{n_o}^2$ in deriving (12) and (13), which possibly could not be applicable to all parameter estimates at the beginning of the adaptation.

IV. Conclusions

An LMS-based algorithm, called the BFLMS method, has been developed for accurate system identification in the presence of input and output interference, given the noise power ratio. The idea is to minimize the mean square value of the equation-error function subject to a constant-norm constraint and it is proved that the constrained optimization can be converted into an unconstrained optimization of minimizing a scaled mean square error function. Learning behaviour of the estimated parameters are derived and verified by computer simulations.

References

- [1] T.Soderstrom and P.Stoica, *System Identification*, Englewood Cliffs, NJ: Prentice-Hall, 1989
- [2] G.E.P.Box, G.M.Jenkins and G.C.Reinsel, *Time Series Analysis, Forecasting and Control*, Englewood Cliffs, NJ: Prentice-Hall, 1994
- [3] T.Kailath, A.H.Sayed and B.Hassibi, *Linear Estimation*, Upper Saddle River, NJ: Prentice-Hall, 2000
- [4] S.Sagara and K.Wada, "On-line modified least-squares parameter estimation of linear discrete dynamic systems," *Int. J. Contr.*, vol.25, no.3, pp.329-343, 1977

- [5] B.Widrow and S.D.Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985
- [6] P.A.Regalia, "An unbiased equation error identifier and reducer-order approximation," *IEEE Trans. Signal Processing*, vol.42, no.6, pp.1397-1412, June 1994
- [7] K.C.Ho and Y.T.Chan, "Bias removal in equation-error adaptive IIR filters," *IEEE Trans. Signal Processing*, vol. 43, no.1, pp.51-62, Jan. 1995
- [8] S.C.Douglas and M.Rupp, "On bias removal and unit norm constraints in equation error adaptive IIR filters," *Proc. 30th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, pp.1093-1097, Nov. 1996
- [9] H.N.Kim and W.J.Song, "Un biased equation-error adaptive IIR filtering based on monic normalization," *IEEE Sig. Proc. Lett.*, vol.6, no.2, pp.35-37, Feb. 1999
- [10] C.E.Davila, "An efficient recursive total least squares algorithm for FIR adaptive filtering," *IEEE Trans. Signal Processing*, vol. 42, no.2, pp.268-280, Feb. 1994
- [11] G.H.Golub and C.F.Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no.6, pp. 883-893, 1980
- [12] K.Gao, M.O.Ahmad and M.N.Swamy, "A constrained anti-Hebbian learning algorithm for total least-squares estimation with applications to adaptive FIR and IIR filtering," *IEEE Trans. Circuits Syst. II*, vol. 41, no. 11, pp.718-729, November 1994
- [13] D.Z.Feng, Z.Bao and L.C.Jiao, "Total least mean squares algorithm," *IEEE Trans. Signal Processing*, vol. 46, no.8, pp.2122-2130, August 1998
- [14] K.V.Frenando and H.Nicholson, "Identification of linear systems with input and output noise: the Koopmans-Levin method," *IEE Proc. Pt. D*, vol.132, no.1, pp.30-36, 1985
- [15] W.X.Zheng, "Adaptive IIR filtering for noisy input-output systems," *Proc. Int. Symp. Signal Processing and its Applications*, pp.375-378, Aug. 1999

Figure 1: System block diagram for equation-error IIR filtering

Figure 2: Estimates of $\{a_i\}$ for BFLMS algorithm in IIR filtering

Figure 3: Estimates of $\{b_j\}$ for BFLMS algorithm in IIR filtering