論文 95-12-12-25

AN UNBIASED IIR LMS ALGORITHM

Seung Hyon Nam* Regular Member

바이어스가 없는 IIR LMS 알고리즘

正會員 南 承 鉉*

ABSTRACT

The equation error formulation in the adaptive IIR filtering provides convergence to a global minimum regardless a local minimum with a large stability margin. However, the equation error formulation suffers from the bias in the coefficient estimates. In this paper, a new algorithm, which does not require a prespecification of the noise variance, is proposed for the equation error formulation. This algorithm is based on the equation error smoothing and provides an unbiased parameter estimate in the presence of white noise. Through simulations, it is demonstrated that the algoritm eliminates the bias in the parameter estimate while retaining good properties of the equation error formulation such as fast convergence speed. the large stability margin, and global convergence.

要 約

IIR 적응 휠터의 공식오차 방식은 지역 최소값에 관계없이 전역 최소값에 수렴하며 안정성이 높다. 그러나 공식오차 방식은 입력 신호에 잡음이 섞여 있는 경우 예측계수가 바이어스되는 문제가 있다. 본 논문에서는 사전에 잡음에 대한 지식이 없이 바이어스가 없는 예측계수를 얻을 수 있는 새로운 공식오차 방식을 위한 알고리즘을 제안한다. 이 알고리즘은 공식오차를 스므딩하는 방식을 이용하여 입력에 추가되는 잡음이 백색잡음인 경우 바이어스가 없는 계수를 예측할 수 있다. 시뮬레이션을 통해 새로운 알고리즘이 공식오차의 주요한 장점인 빠른 수렴속도와 안정성, 그리고 전역 최소값으로의 수렴 특성등을 유지하며 바이어스를 효율적으로 제거함을 볼수 있다.

* 배재대학교 응용수학과
論文番號：95256-0802
接受日字：1995年 8月 2日

Ⅰ. INTRODUCTION

Over the past decade adaptive IIR filtering has been studied extensively [1]-[6] and many application areas have been considered [7]-[9]. Adaptive IIR filtering has a pole-zero structure whereas adaptive FIR filtering has only an all-zero structure. Inclusion of poles in adaptive filtering changes the filtering problem in many ways and adaptive IIR filtering has many advantages over its adaptive FIR counterpart. For example, in channel equalization problems, communication channels are modeled to have zeros. Thus it is necessary for an adaptive filter to have poles to remove distortion caused by the channel. In adaptive FIR filtering, however, poles are approximated using zeros since FIR filters do not have poles. Hence, the number of taps must be large enough to get satisfactory results. On the other hand, if we use adaptive IIR filtering, the zeros in the channel model can be exactly countered using poles in the adaptive filter. Therefore, the number of taps in the adaptive filter is significantly reduced and the computational burden is greatly relieved.

Although adaptive IIR filtering provides better performances in many applications, it has a stability problem. In adaptive FIR filtering, instability occurs when the coefficients get larger without bound due to a large step size beyond the upper bound. In adaptive IIR filtering, however, instability can occur without the coefficients blow-up since the poles outside the unit circle produce unbounded output. Hence, there are two sources of instability in adaptive IIR filtering. Despite this stability problem, adaptive IIR filtering would be an ultimate way to adaptive filtering since it is promising in many signal processing applications. In general, there are two approaches to adaptive IIR filtering that minimizes the mean square error (MSE). One is the equation error formulation and the other is the output error formulation.

In the output error formulation, an adaptive algorithm updates the feedback coefficient directly. Hence it is considered as a natural generalization of adaptive FIR filtering. The output-error adaptive filter is in a recursive form in that the filter output is fed back to the input. Due to this feedback, the filter output is a nonlinear function of the filter coefficients and the MSE surface is not quadratic. This nonlinearity makes the output error formulation complicated than the equation error counterpart. Further, the MSE surface may have local minima as well as a global minimum in some cases. The MSE surface for the output error formulation has been investigated [10]-[12]. The MSE surface has local minima when the order of the adaptive filter is less than that of the signal model. Moreover, even in the exact order case, the MSE surface may have local minima when the input is colored [11]. Some sufficient conditions for the error surface not to have local minima has been investigated [12]. The necessary and sufficient conditions are, however, not known to date. Convergence to a global minimum also depends on the specific algorithm. No algorithm is known to converge to the global minimum for all cases. The nonquadratic nature of the MSE surface also makes adaptive algorithms slower than their equation error counterparts.

In the equation error formulation, an adaptive algorithm updates the feedback coefficients in an all-zero, nonrecursive form. The feedback coefficients are then copied to a second filter implemented in an all-pole form.

Therefore, the equation error formulation uses an adaptive FIR filtering technique directly. Algorithms like RLS and LMS can be employed to update filter coefficients. Adaptive algorithms based on the equation error formulation have many properties in common with the corresponding adaptive FIR algorithm. The MSE surface is quadratic with respect to the filter coefficients. Hence, it has a global minimum without local minima. Convergence speed is faster than that of the output error counterpart. Adaptive algorithm is in a simple form due to nonrecursive nature of the adaptive filter. Moreover, it has the self-stabilizing feature in that the adaptive filter can remain stable even if poles are outside the unit circle over a half of the time [13]. Unfortunately, however, its filter coefficients are biased in the presence of additive noise [14]. Thus the equation error formulation has been limited to those where the bias is not a significant problem.

Bias in the equation error formulation was investigated by Mendel [14]. The bias is a function of the signal to noise (SNR) ratio. That is, the smaller the SNR the larger the bias. He devised an unbiased algorithm assuming that the variance of noise is available. Similar approach was used by Treichler for frequency estimation of a noisy sinusoid corrupted by additive white noise [15]. However, these approaches require a priori knowledge on the variance of additive noise, which is not available in many real situations. Recently, two bias removal techniques have been proposed: One is the BRLE algorithm by Lin and Unbehauen [16] and the other is the QCEE algorithm by Ho and Chan [17]. The BRLE algorithm, however, cannot remove the bias completely due to stability problem for some cases. On the other hand, the QCEE algorithm has ambiguity in the sign of the coefficient estimates.

This paper explores the bias problem and proposes an algorithm that provides unbiased coefficient estimates. The algorithm is based on the equation error smoothing via a FIR filter (or smoother) whose coefficients are derived from the on-line adaptive filter coefficient to be estimated. It is shown that the smoother can eliminate the bias completely in the presence of additive white noise. Simulations are performed to demonstrate that the new algorithm eliminates the bias while retaining the nice properties of the equation error formulation such as high speed, global convergence, and self-stabilizing feature.

Ⅱ. ADAPTIVE IIR FILTERING

Output Error Formulation

Assume that the unknown system in Fig. 1 is described by the difference equation

$$y(k) = \sum_{i=1}^{N} y(k-i) + \sum_{j=0}^{M} x(k-j) \qquad (1)$$

where a_i's and b_j's are coefficients to be estimated. Throughout the paper, the subscript "a" in Fig. 1 will be replaced by others for convience. For example, e_o denotes the output error and e_e the equation error.

The output-error adaptive filter is described in Fig. 2. The output-error adaptive filter that approximates the unknown system is governed by the difference equation

$$y_o(k) = \sum_{i=1}^{N} \hat{a}_i(k) y(k-i) + \sum_{j=0}^{M} \hat{b}_j(k) x(k-j). \qquad (2)$$

The output error is then given by

$$e_o(k) = d(k) - y_o(k). \qquad (3)$$

Here the desired signal $d(k)$ is the contaminated version of the output of the unknown system by additive noise $v(k)$ and is given by

$$d(k) = y(k) + v(k) = \sum_{i=1}^{N} a_i y(k-i) + \sum_{j=0}^{M} b_j x(k-j) + v(k). \qquad (4)$$

A coefficient adaptation rule that minimizes the mean square output error (MSOE)

$$J_o(k) = E\{e_o^2(k)\} = E\{d(k) - y_o(k)\}^2 \qquad (5)$$

may be obtained by differentiating the MSOE with respect to the filter coefficient vector $\hat{\theta}(k)$ defined by

$$\hat{\theta}(k) = [\ \hat{a}_1(k) \cdots \hat{a}_N(k) \quad \hat{b}_0(k) \cdots \hat{b}_M(k)]^T. \qquad (6)$$

Equation Error Formulation

Figure 3 shows the adaptive IIR filter based on the equation error formulation. It is also called a pole-zero adaptive filter since it consists of two adaptive FIR filters, one for poles and the other for zeros. The equation-error adaptive filter is governed by the difference equation

$$y_e(k) = \sum_{i=1}^{N} \hat{a}_i(k) d(k-i) + \sum_{j=0}^{M} \hat{b}_j(k) x(k-j) \qquad (7)$$

which is nonrecursive. By defining

$$\phi_e(k) = [d(k-1) \cdots d(k-N) \ \ x(k) \cdots x(k-M)]^T, \qquad (8)$$

the output $y_e(k)$ in (7) can be rewritten in a linear regression form as

$$y_e(k) = \hat{\theta}^T(k) \phi_e(k). \qquad (9)$$

The equation error is then given by

$$e_e(k) = d(k) - y_e(k) = d(k) - \hat{\theta}^T(k) \phi_e(k) \qquad (10)$$

and the mean square equation error (MSEE) to be minimized is given by

$$J_e(k) = E\{e_e^2(k)\} = E\{d(k) - y_e(k)\}^2. \qquad (11)$$

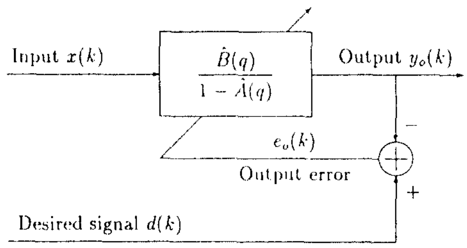

Fig. 2. Adaptive IIR filter based on the output error formulation.

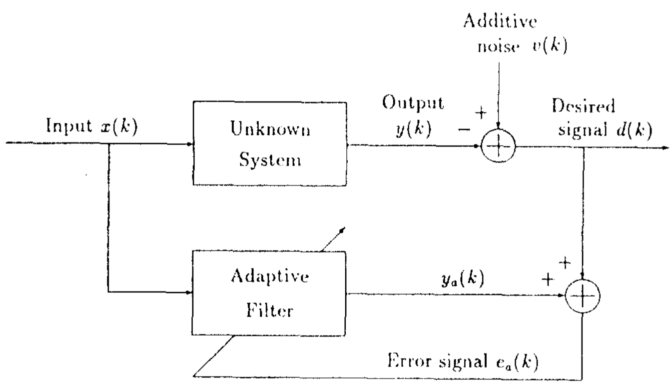

Fig. 1. System identification configuration.

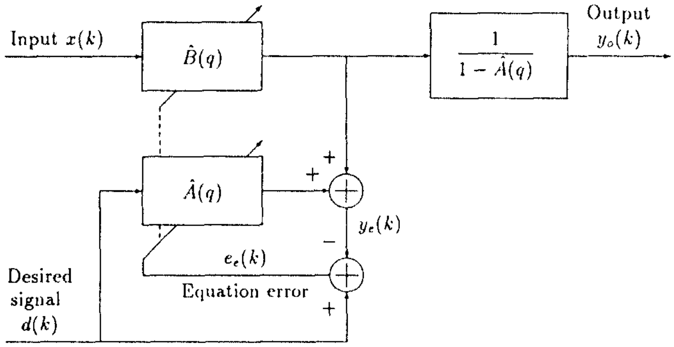

Fig. 3. Adaptive IIR filter based on the equation error formulation.

Note that the output $y_e(k)$ is a linear function of the filter coefficients since $d(k)$ and $x(k)$are independent of the filter coefficients. Hence the equation error is also linear in the filter coefficients and the MSEE is quadratic in the filter coefficients. The MSEE surface has a global minimum without local minima.

Using the instantaneous gradients, the coefficients adaptation rule can be found to be

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \mu e_e(k)\phi_e(k). \tag{12}$$

Equations (10) and (12) constitute the LMS algorithm for the IIR case. Throughout this paper, this algorithm will be referred to as the LMS with equation error (LMS-EE) algorithm.

The LMS-EE algorithm is in fact a direct extension of the LMS algorithm. Thus many properties are common to both algorithms. The LMS-EE algorithm converges faster than its output-error counterpart due to the quadratic nature of the MSEE surface. Further, it has the self-stabilizing feature in that the adaptive filter can remain stable even though poles lie outside of the unit circle over a part of the time. The main reason for this self-stabilizing feature is the nonrecursive nature of the equation-error adaptive filter. Despite these advantages of the equation error formulation, it has a major drawback. The filter coefficient vector $\theta(k)$ is biased in general in the presence of additive noise.

Ⅲ. BIAS ELIMINATION FOR THE EQUATION ERRROR FORMULATION

Begin to introduce the bias elmination technique by defining polynomials

$$\begin{aligned}
1-A(q) &= 1-a_1q^{-1} - a_2q^{-2} \cdots - a_Nq^{-N} \\
1-\hat{A}(q) &= 1-\hat{a}_1(k)q^{-1} - \hat{a}_2(k)q^{-2} \cdots - \hat{a}_N(k)\, q^{-N} \\
B(q) &= b_0+b_1q^{-1}+b_2q^{-2}+\cdots+b_Mq^{-M} \\
\hat{B}(q) &= \hat{b}_0(k)+\hat{b}_1(k)q^{-1}+\hat{b}_2(k)q^{-2}+\cdots+\hat{b}_M(k)\, q^{-M}
\end{aligned} \tag{13}$$

where q^{-1} is a delay operator for a time varying process. Then the output error $e_o(k)$ in (3) is given by

$$\begin{aligned}
e_o(k) &= d(k) - y_o(k) \\
&= d(k) - \hat{y}_o(k) - \hat{B}(q)x(k)
\end{aligned} \tag{14}$$

and the equation error in (10) by

$$\begin{aligned}
e_o(k) &= d(k) - y_e(k) \\
&= d(k) - \hat{A}(q)d(k) - \hat{B}(q)x(k).
\end{aligned} \tag{15}$$

Thus

$$e_e(k)-e_o(k) = -\hat{A}(q)[d(k)-y_o(k)] = -\hat{A}(q)e_o(k) \quad (16)$$

or

$$\begin{aligned} e_e(k) &= [1-\hat{A}(q)]e_o(k) \\ &= [1-\hat{A}(q)][d(k)-y_o(k)]. \end{aligned} \quad (17)$$

This is the well-known relationship between the equation error and the output error.

In adaptive filtering, the ultimate goal is to minimize the MSOE. However, the equation error formulation takes a round-about approach. As a result, in the equation error formulation, the adaptive filter should minimize the MSEE, which is a function of both the output error and the feedback coefficients. In other word, the equation-error adaptive filter should minimize the filtered version of the MSOE. This filtering of the noisy process $d(k)$ by the feedback polynomial $1-\hat{A}(q)$ produces bias. This facts suggests one bias elimination technique: counterbalancing of the polynomial $1-\hat{A}(q)$.

Counterbalancing may be achieved by AR filtering of the output error $e_o(k)$ using a proper polynomial. Clearly, the ideal case is to use $1-\hat{A}(q)$. The resulting algorithm will be very similar to the output-error algorithm by Feintuch [2], which may not converge to either a global minimum or a local minimum [18]. Hence, MA filtering (or smoothing) is desirable.

Define

$$e(k) = [1+C(q)]e_e(k) \quad (18)$$

where the polynomial $1+C(q)$ is chosen as

$$\begin{aligned} \frac{1}{1-\hat{A}(q)} &\approx 1+C(q) \\ &= 1+c_1(k)q^{-1}+c_2(k)q^{-2}+\cdots+c_L(k)q^{-L} \end{aligned} \quad (19)$$

for some L. Note that, for a second order $\hat{A}(q)$,

$$\begin{aligned} \frac{1}{1-\hat{A}(z^{-1})} &= \frac{1}{1-\hat{a}_1(k)z^{-1}-\hat{a}_2(k)z^{-2}} \\ &= 1+\hat{a}_1(k)z^{-1}+[\hat{a}_2(k) \\ &\quad +\hat{a}_1^2(k)]z^{-2}+\cdots \end{aligned} \quad (20)$$

and we can take

$$1+C(q) = 1+\hat{a}_1(k)q^{-1}+[\hat{a}_2(k)+\hat{a}_1^2(k)]q^{-2} \quad (21)$$

using a second order approximation. Similarly, for a third order $\hat{A}(q)$, we may have

$$\begin{aligned} 1+C(q) = &1+\hat{a}_1(k)q^{-1}+[\hat{a}_2(k)+\hat{a}_1^2(k)]q^{-2} \\ &+[\hat{a}_3(k)+2\hat{a}_1(k)\hat{a}_2(k)+\hat{a}_1^3(k)]q^{-3} \end{aligned} \quad (22)$$

using a third order approximation.

From the above idea, the LMS-EE algorithm is modified to have the adaptation rule

$$\hat{\theta}(k+1) = \hat{\theta}(k)+\mu e(k)\phi_e(k). \quad (23)$$

Here $e(k)$ is given by (18) and coefficients of error smoothing polynomial $1+C(q)$ are given by

$$\begin{aligned} c_1(k) &= \hat{a}_1(k) \\ c_i(k) &= \hat{a}_i(k)+\sum_{j=1}^{i-1} c_j(k)\hat{a}_{i-j}(k), \quad \text{for } 2 \le i \le N. \end{aligned} \quad (24)$$

The algorithm (23) with the modified adaptation rule (24) will be referred to as the LMS with smoothed equation error (LMS-SEE) algorithm. Figure 4 describes the adaptive filter based on the smoothed equation error formulation. It is interesting to see that the order of $1+C(q)$ is the same as the order of $1-\hat{A}(q)$. Rationale on the order requirement will be discussed in the following anaysis.

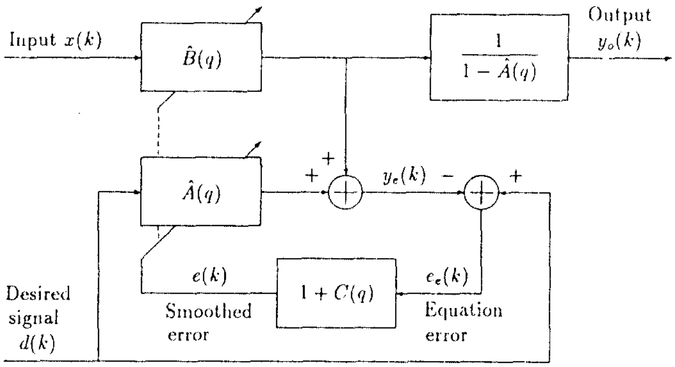

Fig. 4. Adaptive IIR filter based on the smoothed equation error formulation.

Ⅳ. BIAS ANALYSIS OF THE LMS-SEE ALGORITHM

The bias analysis of the LMS-SEE algorithm given in (23) and (24) is very difficult due to the time varying nature of $1+C(q)$ and its correlation with $1-\hat{A}(q)$. To avoid these difficulties, $1+C(q)$ is assumed to be constant in the following analysis. The difference between time-varying and constant $1+C(q)$ in the LMS-SEE algorithm is discussed after the analysis. In addition, the additive noise v(k) is assumed to be white.

With constant $1+C(q)$, the smoothed error is given by

$$e(k) = e_e(k) + \sum_{i=1}^{N} c_i e_e(k-i). \tag{25}$$

Now define

$$\phi_y(k) = [y(k-1)\cdots y(k-N)\ x(k)\cdots x(k-M)]^T$$
$$\phi_v(k) = [v(k-1)\cdots v(k-N), 0\cdots 0]^T. \tag{26}$$

Then we can decomposition $\phi_e(k)$ in (10) into

$$\phi_e(k) = \phi_y(k) + \phi_v(k). \tag{27}$$

Now define the optimal coefficient vector as

$$\theta = [a_1\cdots a_N b_0\cdots b_M]^T. \tag{28}$$

The desired signal in (4) is then rewritten as

$$d(k) = \phi_y^T(k)\theta + v(k). \tag{29}$$

Therfore, we can rewrite $e_e(k)$ in (10) into

$$e_e(k) = \phi_y^T(k)\theta + v(k) - [\phi_y(k) + \phi_v(k)]^T\hat{\theta}(k). \tag{30}$$

Substituting (30) into (23) gives

$$\begin{aligned}\hat{\theta}(k+1) = \ & \hat{\theta}(k) + \mu[\phi_y(k)+\phi_v(k)]\{\phi_y^T(k)\theta + v(k) \\ & -[\phi_y(k)+\phi_v(k)]^T\hat{\theta}(k) - \sum_{i=1}^{N} c_i[\phi_y^T(\\ & k-i)\theta + v(k-i) - [\phi_y(k-i) \\ & + \phi_v(k-i)]^T\hat{\theta}(k-i)]\}.\end{aligned} \tag{31}$$

In order to make the analysis as simple as possible, the followings are assumed [14].

(i) $x(k)$ and $v(k)$ are statistically independent.

(ii) $v(k)$ has a zero mean with finite variance.

(iii) $\phi_e(k)$ and $\hat{\theta}(k)$ are independent and E$\{\phi_e(k)\phi_e^T(k)|\hat{\theta}(k)\}$ is finite and positive definite.

Note that assumption (iii) above requires statistical independence of [14]. This requires that $\phi_e(k)$ and $\hat{\theta}(k)$ do not contain any common element, which is clearly not valid. In [14], Mendel used the decorrelation delay to meet this assumption. However, it should be noticed that assumption (iii) is invalid despite the use of decorrelation delay since $y(k\text{-}1)$ is a function of $\hat{\theta}(k\text{-}1)$ and $y(k\text{-}2)$, and $y(k\text{-}2)$ is in turn a function of $\hat{\theta}(k\text{-}2)$ and $y(k\text{-}3)$ and so on. Assumption (iii), however, works quite well for the unit decorrelation delay, which is the case of the LMS-SEE algorithm under investigation. In the analysis of adaptive FIR filtering algorithms, assumption (iii) is accepted as a consequence of so-called "fundamental assumption" [19].

Taking expectation of both sides of (31) and using three assumptions (i)-(iii) with whiteness assumption of $v(k)$ gives

$$\begin{aligned} E[\hat{\theta}(k+1)] = {} & E[\hat{\theta}(k)] + \mu E[\phi_y(k)\ \phi_y^T(k)]\theta \\ & - \mu\{E[\phi_y(k)\ \phi_y^T(k)] + E[\phi_v(k)\ \phi_v^T(k)]\}E[\hat{\theta}(k)] + \mu\sum_{i=1}^{N} c_i E[\phi_y(k)\ \phi_y^T(k-i)]\theta \\ & + \mu\sum_{i=1}^{N} c_i E[\phi_v(k)v(k-i)] \\ & - \mu\sum_{i=1}^{N} c_i\{E[\phi_y(k)\ \phi_y^T(k-i)] + E[\phi_v(k)\ \phi_v^T(k-i)]\}E[\hat{\theta}(k-i)]. \end{aligned} \tag{32}$$

Assume that the algorithm converges. Then, at steady state, it is true

$$\lim_{k\to\infty} E[\hat{\theta}(k)] = \lim_{k\to\infty} E[\hat{\theta}(k-i)], \quad \text{for } 1\le i\le N. \tag{33}$$

Therefore, taking limit both sides of (32) and using (33) yields

$$\begin{aligned} & \Big\{E[\phi_y(k)\ \phi_y^T(k)] + \sum_{i=1}^{N} c_i E[\phi_y(k)\ \phi_y^T(k-i)]\Big\} \lim_{k\to\infty} E[\hat{\theta}(k)] + \Big\{E[\phi_v(k)\ \phi_v^T(k)] + \sum_{i=1}^{N} c_i E[\phi_v(k)\ \phi_v^T(k-i)]\Big\} \lim_{k\to\infty} E[\hat{\theta}(k)] \\ & = \Big\{E[\phi_y(k)\ \phi_y^T(k)] + \sum_{i=1}^{N} \phi_y(k)\ \phi_y^T(k-i)]\Big\}\theta + \sum_{i=1}^{N} c_i E[\phi_v(k)v(k-i)] \end{aligned} \tag{34}$$

Hence, for $\lim_{k\to\infty} E[\hat{\theta}(k)] = \theta$, it is necessary that

$$\Big\{E[\phi_v(k)\ \phi_v^T(k)] + \sum_{i=1}^{N} \phi_v(k)\ \ \phi_v^T(k-i)]\Big\} \lim_{k\to\infty} E[\hat{\theta}(k)] = \sum_{i=1}^{N} c_i E[\phi_v(k)v(k-i)]. \tag{35}$$

We mat rewrite (35) as

$$\begin{bmatrix} a_1 \\ a_2 + c_1 a_1 \\ a_3 + c_1 a_2 + c_2 a_1 \\ \vdots \\ a_N + c_1 a_{N-1} + \cdots - c_{N-1} a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{36}$$

Therefore, coefficients of $1+C(q)$ for unbiased estimates are given by

$$\begin{aligned} c_1 &= a_1 \\ c_i &= a_i + \sum_{j=1}^{i-1} c_j a_{i-j}, \quad \text{for } 2\le i\le N. \end{aligned} \tag{37}$$

Note that coefficients in (37) are limiting values of (24). If the time varying $1+C(q)$ is used, those $c_i(k)$'s will be computed using current coefficients $\hat{a}_i(k)$, $1\le i\le N$. Hence, for N=2,

$$\begin{aligned} c_1(k) &= \hat{a}_1(k) \\ c_2(k) &= \hat{a}_2(k) + \hat{a}_1^2(k) \end{aligned} \tag{38}$$

and, for N=3

$$\begin{aligned} c_1(k) &= \hat{a}_1(k) \\ c_2(k) &= \hat{a}_2(k) + \hat{a}_1^2(k) \\ c_3(k) &= \hat{a}_3(k) + 2\hat{a}_1(k)\ \hat{a}_2(k) + \hat{a}_1^3(k). \end{aligned} \tag{39}$$

Notice that these error smoothing coefficients are the same as those obtained by inverting $1 - \hat{A}(q)$ and truncating to the order N as in (21) and (22).

As mentioned in the beginning of the analysis, the LMS-SEE algorithm employes the time varying $1+C(q)$. What will happen if we use time varying $1+C(q)$? This can be explained as follows. At time k, coefficient estimates are affected by filtering of the noisy process $d(k)$by feedback coefficients $\hat{a}_i(k)$'s. Thus the equation error $e_e(k)$ contains the undesireable component that is caised by the addtive noise. This component, however, is canceled through error smoothing using $1+C(q)$. Hence, as $\hat{\theta}(k)$ converges to the true value θ, an unbiased estimate $\hat{\theta}(k)$ will result. From this point of view, it is more natural to use time varying $1+C(q)$ than constant $1+C(q)$.

One important thing to mention is the order requirement of $1+C(q)$. Equation (36) shows that the order of $1+C(q)$ should be equal to the order of $1-\hat{A}(q)$. This order requirement on $1+C(q)$ restricts the applicability of the algorithm to those where the order of the adaptive filter is greater than or equal to that of the unknown system. However, in many practical situations of adaptive signal processing applications, adaptive IIR filtering is not superior to its FIR counterpart when it is operating in a reduced order [20]. On the order hand, if the adaptive filter is operating in the overparameterized situation, there will be pole-zero cancellation, and poles and zeros in the adaptive filter correspond to those of the unknown system. Furthermore, the use of extra poles and zeros has the effect of reducing the eigenvalue disparity of the autocorrelation matrix of the input and enhances convergence and stability of adaptive IIR algorithms [13].

Ⅴ. SIMULATIONS

Refer back to the system identification configuration in Fig. 1. The additive noise $v(k)$ is assumed to be white and is uncorrelated with the input $x(k)$. Both colored or white noise are used as an the input $x(k)$. As mentioned in section I, a stability monitoring and projection scheme is required to guarantee stability in adaptive IIR filtering. However, if the step size is taken small enough, the adaptive filter would be stable without the scheme. All simulations in this section are performed without a stability monitoring and projection scheme.

The problem is to identify the unknown system

$$H_d(q)=\frac{0.8-1.6q^{-1}}{1-1.5q^{-1}+0.8125q^{-2}} \tag{40}$$

using the adaptive filter

$$H_a(q)=\frac{b_0(k)+b_1(k)q^{-1}}{1-a_1(k)q^{-1}-a_2(k)q^{-2}}. \tag{41}$$

with the white noise input $x(k)$ at SNR = 0 dB. Both the LMS-EE and LMS-SEE algorithms are simulated. Step size values used are μ_F = 0.002 for feedforward coefficients b_j's and μ_B = 0.0002 for feedback coefficients a_i's for all algorithms. It is interesting to note that μ_F is chosen much larger than μ_B. This is because that convergence of the feedfoward coefficients tends to stabilize the convegence process of the feedback coefficients. Figure 5 shows convergence of coefficients for the LMS-EE and LMS-SEE algorithms averaging 10 independent runs. Clearly, all coefficient estimates are biased except b_0. The bias is caused by filtering of noisy process $d(k)$ which is applied to the feedback section of the adaptive filter. Hence, the bias in the feedback coefficients is directly caused by the noise process. This bias in turn affects feedforward coefficients via coupling between $y(k-i)$ and $x(k-j)$for $i < j$. The coefficient estimate b_0 is unbiased since $x(k)$ is uncorrelated with $y(k-i)$ for $i \geq 1$. The LMS-SEE algorithm, however, provides unbiased coefficient estimates for all coefficients as expected.

Last word on this example is speed of convergence. The LMS-SEE algorithm is compared to the LMS-EE and White's algorithms. Figure 6 shows the learning curves for three algorithms for SNR = 30dB. The step size values are set to yield the MSOE of -29.5dB for all algorithms. The LMS-EE converges very slowly due to bias. It may not achieve the target MSOE if the step size values are set too small. On the other hand, White's algorithm and the LMS-SEE algorithm converge about the same time. Now the step size values are set to yield the MSOE of -28.4dB for all algorithms. Relatively large step size values are required for all algorithms for this MSOE. Figure 7 shows the learning curves for three algorithms. The LMS-EE congverges first, the LMS-SEE algorithm the second, White's algorithm the last. However, we can see the frequent instability in White's algorithm due to large step size values. This reveals that the step size values for White's algorithm should be chosen small enough not to cause instability, and fast convergence may not be achieved in general. The LMS-SEE algorithm, however, can achieve fast convergence for all cases. These results reveal the nice feature of the LMS-SEE algorithm over the existing LMS-EE and White's algorithms.

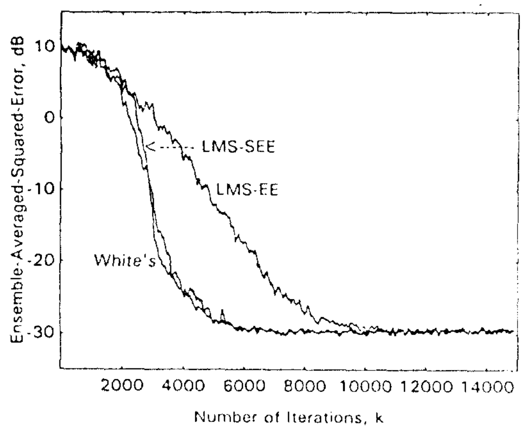

Fig. 6. Convergence of the LMS-EE, LMS-SEE, and White's algorithms for SNR = 30 dB with small step size values.

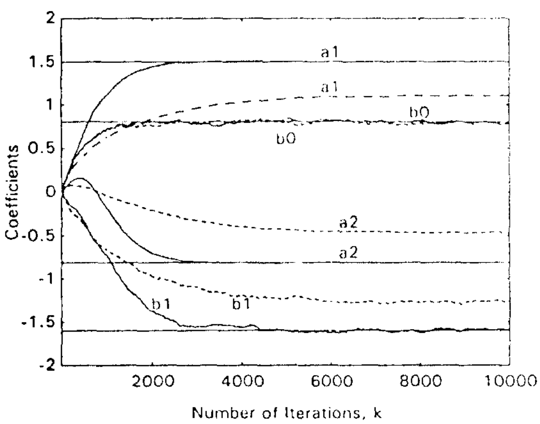

Fig. 5. Coefficient convergence of the LMS-EE(dashed curves) and LMS-SEE(solid curves) algorithms for SNR = 0 dB (optimum coefficients are shown in solid lines).

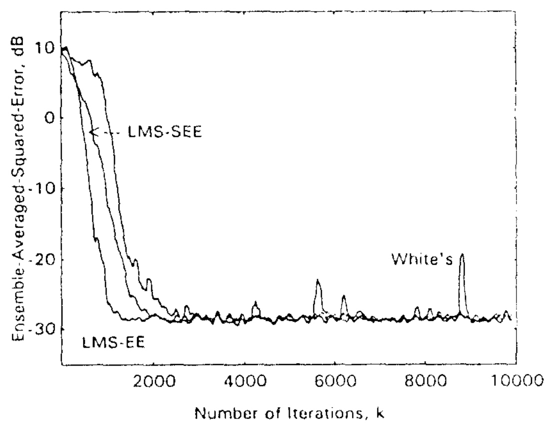

Fig. 7. Convergence of the LMS-EE, LMS-SEE, and White's algorithms for SNR = 30 dB with large step size values.

Next example shows the global convergence of the LMS-SEE algorithm regardless of local minima. The following example is constructed by Soderstrom [21]. The unknown system is given by

$$H_d(q)=\frac{1}{1-1.5q^{-1}+0.49q^{-2}} \tag{42}$$

and the adaptive filter by

$$H_a(q)=\frac{b_0(k)}{1-a_1(k)q^{-1}-a_2(k)q^{-2}}. \tag{43}$$

The input $x(k)$ to the unknown system and the adaptive filter is colored noise such that

$$x(k) = (1-0.7q^{-1})^2(1+0.7q^{-1})^2u(k) \tag{44}$$

where $u(k)$ is white noise. The input $x(k)$ satisfies the persistent excitation condition. Then the MSOE surface has a global minimumat (a_1, a_2, b_0) = (1.4,-0.49,1.0) and a local minimum at (-1.367,0.513,-0.23). Figure 8 shows the contour of constant MSOE and the feedback coefficient trajectory. Filter coefficients are initialized near the local minimum and it converged to the global minimum. Another thing we can observe from Fig. 8 is that the part of the coefficient trajectory is outside the stability triangle, and the adaptive filter remains stable without aid of any stability monitoring and projection scheme. It is due to the self-stabilizing feature comes from the nonrecursive nature of the equation error formulation. This global convergence and nice stability property can be achieved by the LMS-EE algorithm. The coefficient estimates are, however, biased in the presence of additive noise. Therefore, the LMS-SEE algorithm provides unbiased coefficient estimates while retaining the global convergence property of the equation error formulation.

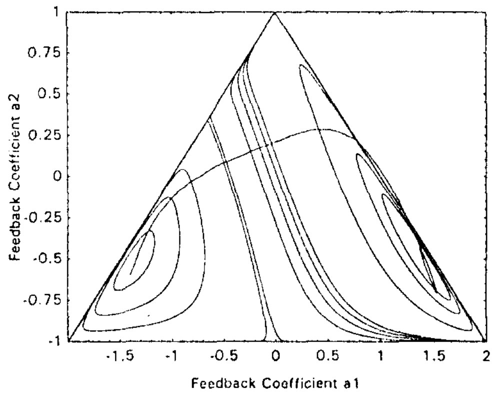

Fig. 8. The coefficient trajectory of the LMS-SEE Algorithm for a multimodal MSOE case.

Simulations in this section show that the LMS-SEE algorithm eliminated the bias in the coefficient estimates if aditive noise is white. Furthermore, it is confirmed that the LMS-SEE algorithm retains advantages of the equation error formulation such as faster convergence speed, and a self stability features. Hence, the LMS-SEE algorithm makes the equation error formulation viable even in the presence of the additive white noise.

Ⅵ. CONCLUSIONS

This paper studied the bias problem in the IIR LMS algorithm. The IIR LMS algorithm produces biased estimates for all coefficients except the zeroth order MA coefficient if there is additive noise in the input. An algorithm is devised to get unbiased coefficient estimates when the additive noise is white. Simulations are performed to demonstrated the performance of the new algorithm. The new algorithm retaines properties such as the self-stabilizing feature, global convergence, and fast convergence speed while provides unbiased coefficient estimates.

REFERENCES

1. R. P. Gooch, "Adaptive pole-zero filtering : the equation error approach," Ph.D. diss. Stanford University, 1983.
2. P. L. Feintuch, "An adaptive recursive LMS filter," *Proc. IEEE*, vol. 64, no. 11, pp.1622-1624, Nov. 1976.
3. S.Horvath, Jr., "A new adaptive recursive LMS filter," in *Digital Signal Processing*, V.Cappellini and A.G. Constantinides, eds. Academic Press, New York, 1980, pp.21-26.
4. S. A. White, "An adaptive recursive digital filter," in *Proc. 9th Asiloma Conf. Circuits, Systems, Computers*, Pacific Grove, California, Nov. 1975, pp.21-25.
5. J. R. Treichler, M.G. Larimore, and C.R. Johnson, Jr., "Simple adaptive IIR filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, Tulsa, Oklahoma, Apr. 1978, pp.118-122.
6. H. Fan and W.K. Jenkins, "A new adaptive IIR filter," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp.939-947, Oct. 1986.
7. J. R. Treichler, "Adaptive algorithm for infinite impulse response filters," in *Adaptive Filters*, C.F.N. Cowen and P.M. Grant, eds. Prentice-Hall, Englewood Cliffs, N. J., 1985, ch. 4, pp.60-90.
8. John J. Shynk, "Adaptive IIR filtering," *IEEE ASSP Magazine*, pp.4-21, April, 1989.
9. C. R. Johnson, Jr. and A. L. Hamm, "Debiasing equation error frequency detection of a noisy sinusoid without noise variance prespecification," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-27, no. 6, pp.654-656, Dec. 1979.
10. S.D. Stearns, "Error surface of recursive adaptive filters," *IEEE Trans. Circuits Systems*, vol. CAS-28, no. 6, pp.603-606, June, 1981.
11. T. Soderstrom and P. Stoica, "Some properties of the output error method," *Automatica*, vol. 18, pp.93-99, Jan. 1982.
12. H. Fan and M. Nayeri, "On error surface of sufficient order adaptive IIR filters: proofs and counterexamples to a unimodality conjecture," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 37, no. 9, pp.1436-1442, Sept. 1989.
13. G. Long, D. Shewd, and D. D. Falconer, "Study of a pole-zero adaptive echo canceller," *IEEE Trans. Circuits, Ssytems*, vol. CAS-34, no.7, pp.765-769, July, 1987.
14. J. M. Mendel, *Discrete Techniques of Parameter Estimation : The Equation Error Formulation*. Marcel Dekker, New York, 1973.
15. J. R. Treichler, "Transient and convergent behavior of the adaptive line enhancer," in *IEEE Trans. Acoust., Speech, Signal. Proc.*, vol. ASSP-27, pp.53-62, Feb. 1979.
16. J. N. Lin and R. Unbehauen, "Bias-remedy least mean square equation error algorithm for IIR parameter recursive estimation," *IEEE Trans. Signal Processing*, vol. 40, pp.62-69, Jan. 1992.
17. K. C. Ho and Y. T. Chan, "Bias Removal in Equation-Error Adaptive IIR Filters," *IEEE Trans. Signal Processing*, vol. 43, pp.51-62, Jan. 1995.
18. C. R. Johnson, Jr., "Adaptive IIR filtering: Current results and open issues," *IEEE Trans. Inform. Theory*, vol.IT-30, no. 2, pp.237-250, Mar. 1984.
19. S. Haykin, *Adaptive filter theory*, 2nd. ed., Prentice-Hall, N. J., 1986.
20. H. Fan and W. K. Jenkins, "An investigation of an adaptive IIR echo canceller: Advantages and problems," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. 36, no. 12, pp.1819-1833, Dec. 1988.
21. T. Soderstrom, "On the uniqueness of maximum likelihood identification," *Automatica*, vol. 11, pp.193-197, 1975.

南 承 鉉(Seung Hyon Nam) 정회원

1956년 8월 31일생
1980년 2월 : 서강대학교 전자공학과 졸업(공학사)
1987년 8월 : The University of Alabama, Huntsville 전기 및 컴퓨터공학과 졸업(공학석사)
1992년 12월 : Texas A & M University 전기과 졸업(공학박사)
1979년 12월~1985년 6월 : 국방과학연구소 연구원
1993년 3월~현재 : 배재대학교 응용수학과 조교수
※주관심 분야 : 음성 및 오디오 신호처리, 적응신호처리