Constant-norm Equation-error 적응 IIR 필터를 위한 가변 Step size 알고리즘

공세진, 신현출, 송우진
포항공과대학교 전자컴퓨터공학부 통신 및 신호처리연구실

A Variable Step-size Algorithm for Constant-norm Equation-error Adaptive IIR Filters

Se-Jin Kong, Hyun-Chool Shin and Woo-Jin Song
Department of Electronic and Electrical Engineering
Pohang University of Science and Technology (POSTECH)
E-mail: wjsong@postech.ac.kr

Abstract Recently a constant-norm constraint equation-error method was proposed to solve the bias problem in adaptive IIR filtering. However, the method adopts a fixed step-size and thus results in slow convergence for a small step-size and significant misadjustment error for a large step-size. In this paper, we propose a variable step-size (VSS) algorithm that greatly improves convergence properties of the *constant-norm constraint equation-error* method. The analysis and the simulation results show that the proposed method indeed achieves both fast convergence and small misadjustment error.

I. Introduction

Adaptive IIR filtering methods are classified into two categories: equation error method and output error method [1]. Equation error method has been preferred to output error method, because it has more attractive characteristics such as the quadratic error surface, global convergence, and guaranteed system stability. However, equation error method has a shortcoming that it may generate biased coefficient estimates in the presence of noise. So many debiasing methods [2]-[3] have tried to solve this problem.

Recently the *constant-norm equation-error* method [2] was proposed. The method has achieved the bias free coefficient estimates with a constant norm constraint. But this method has limits due to a fixed step-size. When the step size is small, the convergence speed becomes slow. Conversely when the step size is large, the misadjustment error increases.

To overcome the limits, a variable step-size (VSS) equation-error method is proposed for the *constant-norm equation-error* method. The proposed VSS method is motivated by the fact that the angle between a gradient vector and a denominator coefficient vector becomes smaller as the filter coefficients reach the optimal solution. The gradient vector and the coefficient vector are projected onto an arbitrary vector. Then as a measure of the closeness between estimated filter coefficients and the optimal solution, we use the norm of the difference vector between two projection vectors. According to the measure, the step size is chosen. Simulation results show that the proposed VSS method indeed achieves both fast convergence and low misadjustment error.

The remainder of this paper is organized as follows. Section II briefly describes the *constant-norm equation-error* algorithm. Then Section III analyzes the stationary condition in the *constant-norm equation-error* algorithm and derives the proposed algorithm. Section IV evaluates the performance of the proposed algorithm in comparison with the *constant-norm equation-error* algorithm by computer simulation. Finally conclusions are presented in Section V.

II. Constant-norm Equation-error Algorithm

An equation error adaptive IIR filter in the system identification configuration is shown in Fig.1. Suppose that the unknown system $H(z)$ is stable and causal represented by a rational system function

$$H(z)=\frac{B(z)}{A(z)}=\frac{\sum_{k=0}^{K}b_k z^{-k}}{\sum_{l=0}^{L}a_l z^{-l}}. \tag{1}$$

The relationship between the input and output signals for the unknown system $H(z)$ can be written as

$$y(n)=\frac{1}{a_0}\left(\sum_{k=0}^{K}b_k x(n-k)-\sum_{l=1}^{L}a_l y(n-l)\right) \tag{2}$$

where $x(n)$ is the input signal. When the output $y(n)$ is corrupted by additive noise, the observed output or the desired signal is given by

$$d(n)=y(n)+v(n) \tag{3}$$

where $v(n)$ is white measurement noise with variance σ_v^2 and is independent of the input. The conventional equation error IIR filtering algorithm has a shortcoming that generate biased coefficient estimations in the presence of noise. In order to remove this bias, the *constant-norm equation-error* algorithm was proposed [1]. The *constant-norm equation-error* IIR filter has two FIR filter outputs

$$w(n)=\mathbf{d}^T(n)\hat{\mathbf{a}} \tag{4}$$

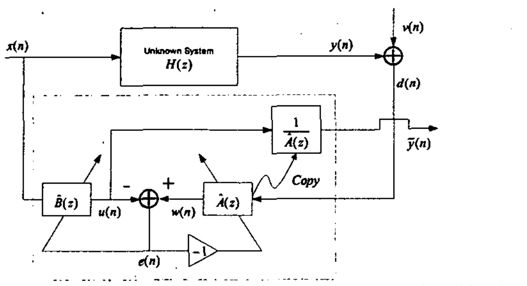


Fig.1 Equation error adaptive IIR filter in the system identification configuration

and

$$u(n) = \mathbf{x}^T(n)\hat{\mathbf{b}} \tag{5}$$

where $\hat{\mathbf{a}}^T = [1 \;\; \hat{a}_1 \;\; \cdots \;\; \hat{a}_L], \hat{\mathbf{b}}^T = [\hat{b}_0 \;\; \hat{b}_1 \;\; \cdots \;\; \hat{b}_K]$, $\mathbf{d}^T(n) = [d(n) \;\; d(n-1) \;\; \cdots \;\; d(n-L)]$, and $\mathbf{x}^T(n) = [x(n) \;\; x(n-1) \;\; \cdots \;\; x(n-K)]$. Then the equation error $e(n)$ and mean square error (MSE) can be represented as

$$e(n) = w(n) - u(n) = \mathbf{d}^T(n)\hat{\mathbf{a}} - \mathbf{x}^T(n)\hat{\mathbf{b}} \tag{6}$$

and

$$E[e^2(n)] = E[(\mathbf{y}^T(n)\hat{\mathbf{a}} - \mathbf{x}^T(n)\hat{\mathbf{b}})^2] + \sigma_v^2\left(1 + \sum_{l=1}^{L}\hat{a}_l^2\right) \tag{7}$$

where $\mathbf{y}^T(n) = [y(n) \;\; y(n-1) \;\; \cdots \;\; y(n-L)]$.

From (7), we can know that $\sigma_v^2\left(1 + \sum_{l=1}^{L}\hat{a}_l^2\right)$ introduces an undesirable bias that depends on the noise power. To solve the problem *the constant- norm equation-error* algorithm uses the constant-norm constraint

$$\left(1 + \sum_{l=1}^{L}\hat{a}_l^2\right) = C \tag{8}$$

where $C > 1$. A bias-free solution is obtained by MSE minimization subject to (8). The augmented cost function, with the *constant-norm constraint*, using Lagrange multiplier λ, is given by

$$J = \frac{E[e^2(n)] + \lambda(C - \hat{\mathbf{a}}^T\hat{\mathbf{a}})}{2}. \tag{9}$$

Finally using LMS adaptive method, the updating equations is as follows:

$$\hat{\mathbf{a}}(n+1) = \hat{\mathbf{a}}(n) - \mu_a e(n)\left(\mathbf{d}(n) - \frac{\mathbf{p}^T\mathbf{d}(n)}{\mathbf{p}^T\hat{\mathbf{a}}(n)}\hat{\mathbf{a}}(n)\right) \tag{10}$$

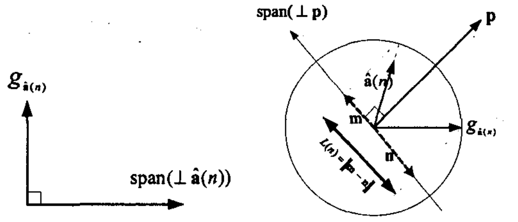
and

$$\hat{\mathbf{b}}(n+1) = \hat{\mathbf{b}}(n) + \mu_b e(n)\mathbf{x}(n) \tag{11}$$

where μ_a and μ_b are the step size for the FIR filter $\hat{A}(z)$ and $\hat{B}(z)$, respectively. The sifting vector $\mathbf{p}$ is an arbitrary vector which satisfies $p^T\hat{a}(n) \neq 0$. The detail description is given in [2].

III. The VSS Constant-norm EE Algorithm

We apply VSS technique to the $\hat{A}(z)$ FIR filter under a constant-norm constraint and $\hat{a}_0 = 1$. On the other hand, a fixed step size method is still applicable for the $\hat{B}(z)$ FIR filter.



(a) (b)

Fig.2 (a) Orthogonality of the gradient of MSE and span (⊥ $\hat{\mathbf{a}}(n)$) (b) Projection length $L(n)$ onto span (⊥ p)

A .VSS algorithm

Firstly we investigate characteristics of stationary point in *constant-norm equation-error* algorithm. Let us define a gradient of MSE with respect to $\hat{\mathbf{a}}(n)$ as

$$g_{\hat{\mathbf{a}}(n)} = \frac{\partial E[e^2(n)]}{\partial \hat{\mathbf{a}}(n)}. \tag{12}$$

Then the derivative of J with respect to $\hat{\mathbf{a}}(n)$ is given by

$$\begin{aligned}\frac{\partial J}{\partial \hat{\mathbf{a}}(n)} &= \frac{1}{2}\frac{\partial E[e^2(n)]}{\partial \hat{\mathbf{a}}(n)} - \frac{\mathbf{p}^T E[\mathbf{d}(n)e(n)]}{\mathbf{p}^T\hat{\mathbf{a}}(n)}\hat{\mathbf{a}}(n) \\ &= \frac{1}{2}g_{\hat{\mathbf{a}}(n)} - \frac{\mathbf{p}^T E[\mathbf{d}(n)e(n)]}{\mathbf{p}^T\hat{\mathbf{a}}(n)}\hat{\mathbf{a}}(n)\end{aligned} \tag{13}$$

At stationary point, i.e., $\partial J/\partial\hat{\mathbf{a}}(n) = 0$, it follows that

$$\frac{1}{2}g_{\hat{\mathbf{a}}(n)} = \frac{\mathbf{p}^T E[\mathbf{d}(n)e(n)]}{\mathbf{p}^T\hat{\mathbf{a}}(n)}\hat{\mathbf{a}}(n) \tag{14}$$

Let us define $\text{span}(\perp \hat{\mathbf{a}}(n))$ as any vector that is orthogonal to $\hat{\mathbf{a}}(n)$. By multiplying $\text{span}(\perp \hat{\mathbf{a}}(n))$ both sides, we get

$$g^T_{\hat{\mathbf{a}}(n)} \cdot \text{span}(\perp \hat{\mathbf{a}}(n)) = 0, \tag{15}$$

since $\hat{\mathbf{a}}^T(n) \cdot \text{span}(\perp \hat{\mathbf{a}}(n)) = 0$. Therefore, at stationary point a gradient vector is orthogonal to vectors in $\text{span}(\perp \hat{\mathbf{a}}(n))$. From a geometric perspective, the condition in (15) can be illustrated in Fig.2 (a).

When the condition in (15) is satisfied, $g_{\hat{\mathbf{a}}(n)}$ and $\frac{\|g_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}\hat{\mathbf{a}}(n)$ coincides because the two vectors have the same direction and the same length. So the projection vectors of $g_{\hat{\mathbf{a}}(n)}$ and $\frac{\|g_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}\hat{\mathbf{a}}(n)$ onto the space spanned by $\text{span}(\perp \mathbf{p})$ becomes the same one. Therefore the difference vector of two projection vectors can be used as the estimation of the closeness between the estimated parameter $\hat{\mathbf{a}}(n)$ and its optimal solution. In other words, the length of the difference vector is used as a criterion for the step size selection of μ_a. In Fig.2 (b), $\mathbf{m}$ represents $\frac{\|g_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}\left(\hat{\mathbf{a}}(n) - \frac{\hat{\mathbf{a}}^T(n)\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p}\right)$ and $\mathbf{n}$ represents

$\left(g_{\hat{\mathbf{a}}(n)} - \frac{g^T_{\hat{\mathbf{a}}(n)}\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p} \right)$. With $\mathbf{m}$ and $\mathbf{n}$, the length of the difference vector $L(n)$ is given by

$$L(n) = \|\mathbf{m} - \mathbf{n}\| = \left\| \frac{\|g_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}\left(\hat{\mathbf{a}}(n) - \frac{\hat{\mathbf{a}}^T(n)\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p} \right) - \left(g_{\hat{\mathbf{a}}(n)} - \frac{g^T_{\hat{\mathbf{a}}(n)}\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p} \right) \right\| . \tag{16}$$

This fact is illustrated in Fig.2 (b).

Employing the LMS algorithm, the instantaneous gradient vector is given by

$$\hat{g}_{\hat{\mathbf{a}}(n)} = 2e(n)\mathbf{d}(n) \tag{17}$$

where $E[\hat{g}_{\hat{\mathbf{a}}(n)}] = g_{\hat{\mathbf{a}}(n)}$.

Because we cannot know exact expectation of $\hat{g}_{\hat{\mathbf{a}}(n)}$, we approximate it by first-order recursive estimation method. The first-order recursive estimation is given by

$$\bar{g}_{\hat{\mathbf{a}}(n)} = \alpha \bar{g}_{\hat{\mathbf{a}}(n-1)} + (1-\alpha)\hat{g}_{\hat{\mathbf{a}}(n)} \tag{18}$$

where α is the forgetting factor $(0 < \alpha < 1)$. Then $\bar{g}_{\hat{\mathbf{a}}(n)}$ can approximate $E[\hat{g}_{\hat{\mathbf{a}}(n)}]$. As α is chosen to be close to 1, $\bar{g}_{\hat{\mathbf{a}}(n)}$ becomes closer to $E[\hat{g}_{\hat{\mathbf{a}}(n)}]$. By using the estimated gradient vector in (18), the estimated projection length $\hat{L}(n)$ is given by

$$\hat{L}(n) = \left\| \frac{\|\bar{g}_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}\left(\hat{\mathbf{a}}(n) - \frac{\hat{\mathbf{a}}^T(n)\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p} \right) - \left(\bar{g}_{\hat{\mathbf{a}}(n)} - \frac{\bar{g}^T_{\hat{\mathbf{a}}(n)}\mathbf{p}}{\|\mathbf{p}\|^2}\mathbf{p} \right) \right\| . \tag{19}$$

Based on $\hat{L}(n)$, we determine the $\mu_a(n)$ within the range $\mu_{\min} \le \mu_a(n) \le \mu_{\max}$. Let us exploit the linear method in the step size selection. We first determine the upper bound $L_{\max}$ and the lower bound $L_{\min}$ of the projection length. Then, the step size of the VSS *constant-norm equation-error* algorithm is set as

$$\mu_a(n) = \begin{cases} \mu_{\min}, & \text{if } \hat{L}(n) < L_{\min} \\ \mu_{\max}, & \text{if } \hat{L}(n) > L_{\max} \\ \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{L_{\max} - L_{\min}} \cdot (\hat{L}(n) - L_{\min}), & \text{otherwise} \end{cases} \tag{20}$$

where $L_{\max} > L_{\min} > 0$.

B. A family of VSS algorithm

By choosing $\mathbf{p}$ appropriately, various efficient VSS *constant-norm constraint equation-error* can be derived. In this section, we give two examples.

Example 1 - one choice is $\mathbf{p} = \hat{\mathbf{a}}(n)$. As shown in Fig.2 (b), when $\mathbf{p}$ is equal to $\hat{\mathbf{a}}(n)$, $\mathbf{m}$ becomes $\vec{0}$ and thus the projection length $\hat{L}(n)$ becomes

$$\hat{L}(n) = \|\mathbf{n}\| = \sqrt{\|\bar{g}_{\hat{\mathbf{a}}(n)}\|^2 - \left(\frac{\bar{g}^T_{\hat{\mathbf{a}}(n)}\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|} \right)^2} . \tag{21}$$

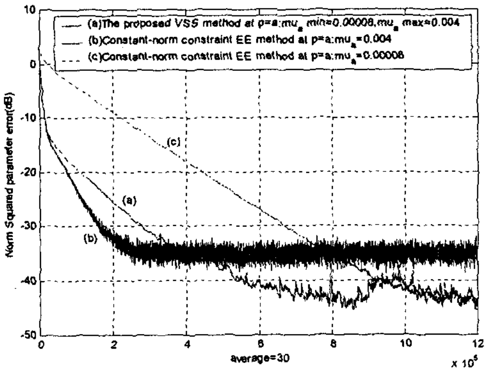


Fig.3 Plots of the norm squared parameter error $(\mathbf{p} = \hat{\mathbf{a}}(n))$

In this case, the proposed VSS method becomes the same as the VSS QCEE method in [4]. The VSS QCEE method corresponds to the one special case of various VSS *constant-norm equation-error* families.

Example 2 - Another possible choice for the sifting vector is $\mathbf{p} = [1 \ 0 \ \cdots \ 0]$. Let us represent $\bar{g}_{\hat{\mathbf{a}}(n)}$ as $\bar{g}^T_{\hat{\mathbf{a}}(n)} \approx [\bar{g}_0 \ \bar{g}_1 \ \cdots \ \bar{g}_L]$. Then the estimated projection length $\hat{L}(n)$ is given by

$$\hat{L}(n) = \left\| \frac{\|\bar{g}_{\hat{\mathbf{a}}(n)}\|}{\|\hat{\mathbf{a}}(n)\|}(\hat{\mathbf{a}}(n) - \mathbf{p}) - (\bar{g}_{\hat{\mathbf{a}}(n)} - \bar{g}_0\mathbf{p}) \right\| . \tag{22}$$

IV. Simulation Results

The simulation results are presented to show that the proposed method indeed achieves both fast convergence and small misadjustment error. The *constant-norm equation-error* algorithm with fixed step size is chosen for comparison of convergence speed and estimation accuracy.

The unknown system is given by

$$H(z) = \eta \frac{1 - 0.5z^{-1}}{1 - 1.0z^{-1} + 0.5z^{-2}} .$$

The gain η is chosen such that $y(n)$ has unit power and it is chosen to $\eta = 0.8465$ in simulations. The initial parameters are set to $\hat{\mathbf{a}}^T(0) = [1 \ 0 \ 0]$ and $\hat{\mathbf{b}}^T(0) = [0 \ 0]$ in the simulation. Also, the initial values of $\mu_a(n)$ and $\bar{g}_{\hat{\mathbf{a}}(n)}$ are set to $\mu_a(0) = \mu_{\min}$ and $\bar{g}_{\hat{\mathbf{a}}(0)} = 0$. The input signal is a white, zero-mean, Gaussian random sequence with unit power, and the measurement noise signal is additive white Gaussian with power 0.1. The squared norm of the parameter estimation error, $\|\mathbf{a} - \hat{\mathbf{a}}(n)\|^2 + \|\mathbf{b} - \hat{\mathbf{b}}(n)\|^2$ is taken and averaged over 30 independent trials. We fix the parameters as $\mu_b = 0.0001$, $\alpha = 0.999$.

The simulation results are shown in Fig.3 and Fig.4. In Fig.3 we use the sifting vector $\mathbf{p} = \hat{\mathbf{a}}(n)$ and we set the parameters in VSS algorithm as $\mu_{\min} = 0.00008$,

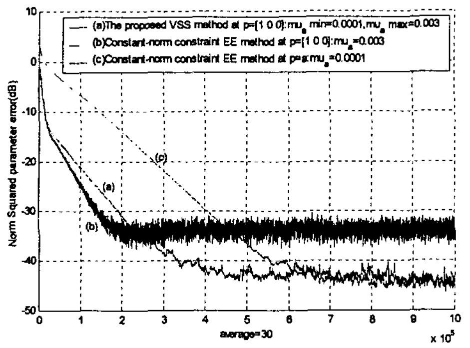


Fig.4 Plots of the norm squared parameter error ($\mathbf{p} = [1 \quad 0 \quad 0]$)

$\mu_{max} = 0.004$, $L_{min} = 0.045$ and $L_{max} = 0.06$. For the comparison purpose the results of the *constant-norm equation-error* algorithm with $\mu_a = 0.004$ and $\mu_a = 0.00008$ are given in the plot (b) and the plot (c), respectively. The result of the proposed method is given in the plot (a). Although there is a trade-off between convergence speed and misadjustment error for the fixed step size, the proposed VSS method has fast convergence and small misadjustment error.

In Fig.4 we use the sifting vector $\mathbf{p} = [1 \quad 0 \quad 0]$ and we set the parameters in VSS algorithm as $\mu_{min} = 0.0001$, $\mu_{max} = 0.003$, $L_{min} = 0.035$ and $L_{max} = 0.07$. For the comparison purpose the results of the *constant-norm equation-error* algorithm with $\mu_a = 0.003$ and $\mu_a = 0.0001$ is given in the plot (b) and the plot (c), respectively. The result of the proposed method is given in the plot (a). As expected, similar results with Fig.3 are obtained.

V. Conclusions

We have proposed the VSS *constant-norm equation-error* adaptive algorithm. The proposed method shows fast convergence speed and small misadjustment error. Also, the proposed algorithm has a great degree of freedom for algorithm selection because of free choice of $\mathbf{p}$. By choosing the sifting vector appropriately, the more efficient and stable VSS *constant-norm constraint equation-error* algorithms can be derived.

Acknowledgments

The authors wish to acknowledge that his work has been partially supported by BK21 program from Ministry of Education, Republic of Korea.

References

[1] J. J. Shynk, "Adaptive IIR filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 6, pp. 4-21, 1989.

[2] H.-C. Shin and W.-J. Song, "Bias-Free Adaptive IIR Filtering," *IEICE Trans. Fundamentals*, vol. E00-A, pp 1273-1279, no. 5, May 2001.

[3] 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.

[4] P. L. Hsu, T. Y. Tsai, and F. C. Lee, "Applications of a Variable Step Size Algorithm to QCEE Adaptive IIR Filters," *IEEE Trans. Signal Processing*, vol. 48, no. 1, pp. 250-254, Jan. 2000.