International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:10, No:1, 2016

Data-Reusing Adaptive Filtering Algorithms with Adaptive Error Constraint

Young-Seok Choi

Abstract—We present a family of data-reusing and affine projection algorithms. For identification of a noisy linear finite impulse response channel, a partial knowledge of a channel, especially noise, can be used to improve the performance of the adaptive filter. Motivated by this fact, the proposed scheme incorporates an estimate of a knowledge of noise. A constraint, called the *adaptive noise constraint*, estimates an unknown information of noise. By imposing this constraint on a cost function of data-reusing and affine projection algorithms, a cost function based on the adaptive noise constraint and Lagrange multiplier is defined. Minimizing the new cost function leads to the *adaptive noise constrained* (ANC) data-reusing and affine projection algorithms. Experimental results comparing the proposed schemes to standard data-reusing and affine projection algorithms clearly indicate their superior performance.

Keywords—Data-reusing, affine projection algorithm, error constraint, system identification.

I. Introduction

HIGH eigenvalue spread of the input signal correlation matrix tend to deteriorate the convergence performance of the least mean square (LMS)-type adaptive filters [1]. Recently, the data-reusing LMS (DR-LMS), the normalized DR-LMS (NDR-LMS,) and the affine projection (AP) algorithms have spawned great interest among researchers desiring to improve convergence at reduced computational cost, and to trade off convergence rate as a function of the computational complexity [2]–[5]. In contrast to LMS-type adaptive filters, these algorithms use block error and block input vector for updating the filter coefficient.

For an identification of a noisy linear finite impulse response (FIR) channel, a partial knowledge of a channel, especially noise, can be used to improve the performance of the adaptive filter [6][7]. Motivated by this fact, we expect that the performance of DR-LMS, NDR-LMS and AP algorithms can be further improved by incorporating a knowledge of noise. However, a knowledge of noise generally is not available to the filter. To overcome this obstacle, a constraint, called *adaptive noise constraint*, which estimates an unknown information of noise is introduced. By imposing this constraint on a cost function of DR and AP algorithms, we present the *adaptive noise constrained* (ANC) DR-LMS, NDR-LMS and AP algorithms, which are based on the adaptive noise constraint and Lagrange multiplier. Through experiments, we illustrate that the proposed algorithms possess better performance than standard DR-LMS, NDR-LMS, AP algorithms in terms of the convergence rate and the misadjustment.

Y.-S. Choi is with the Department of Electronic Engineering, Gangneung-Wonju National University, 7 Jukheon-gil, Gangneung, 210-702 Republic of Korea (phone: +82-33-640-2429; fax: +82-33-646-0740; e-mail: yschoi@gwnu.ac.kr).

II. Adaptive Noise Constrained DR and AP Algorithms

Consider data $d(i)$ that arise from the system identification model

$$d(i) = \mathbf{u}_i \mathbf{w}^{\circ} + v(i), \tag{1}$$

where $\mathbf{w}^{\circ}$ is a column vector for the impulse response of an unknown system that we wish to estimate, $v(i)$ accounts for measurement noise and $\mathbf{u}_i$ denotes the $1 \times M$ row input vector,

$$\mathbf{u}_i = [u(i) \;\; u(i-1) \;\; \cdots u(i-M+1)], \tag{2}$$

and $\mathbf{u}_i$ and $v(i)$ are uncorrelated.

A. Conventional DR-LMS and NDR-LMS and AP Algorithms

Let $\mathbf{w}_i$ be an estimate for $\mathbf{w}^{\circ}$ at iteration i. The DR-LMS, NDR-LMS, and APA take the forms [2]

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{U}_i^* \mathbf{e}_i \tag{3}$$

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{U}_i^* \mathbf{D}_i \mathbf{e}_i \tag{4}$$

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{U}_i^* (\mathbf{U}_i \mathbf{U}_i^* + \delta \mathbf{I})^{-1} \mathbf{e}_i, \tag{5}$$

respectively, where

$$\mathbf{U}_i = \begin{bmatrix} \mathbf{u}_i \\ \mathbf{u}_{i-1} \\ \vdots \\ \mathbf{u}_{i-K+1} \end{bmatrix}, \quad \mathbf{d}_i = \begin{bmatrix} d(i) \\ d(i-1) \\ \vdots \\ d(i-K+1) \end{bmatrix},$$

$\mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i \mathbf{w}_{i-1}$, $\mathbf{D}_i = \mathrm{diag}[1/\|\mathbf{u}_i\|^2, \ldots, 1/\|\mathbf{u}_{i-K+1}\|^2]$, μ is the step-size, and * denotes the Hermitian transpose. It is known [2] that these algorithms are obtained to minimize the following cost function

$$J(i) = E[\mathbf{e}_i^* \Pi \mathbf{e}_i], \tag{6}$$

where Π is a positive-definite matrix. The gradient vector of $J(i)$ with respect to $\mathbf{w}_{i-1}$ is given by

$$\frac{\partial J(i)}{\partial \mathbf{w}_{i-1}} = -E[\mathbf{e}_i^* \Pi \mathbf{U}_i]. \tag{7}$$

Then, we can obtain the stochastic gradient algorithm

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{U}_i^* \Pi \mathbf{e}_i \tag{8}$$

Note that the choice of Π determines algorithms of (3)–(5). In other words, if we choose $\Pi = I$, i.e., the identity matrix, the DR-LMS algorithm (3) is obtained. And with the choice of data-normalized identity matrix, $\mathbf{D}_i$, we get the NDR-LMS (4). Also the choice $\Pi = (\mathbf{U}_i \mathbf{U}_i^*)^{-1}$ results in the APA (5).

The minimization of (6) over $\mathbf{w}_{i-1}$ yields $\mathbf{w}_{i-1} = \mathbf{w}^\circ$ and $\mathbf{e}_i = \mathbf{v}_i$ where $\mathbf{v}_i = [v(i)\ v(i-1) \dots v(i-K+1)]^T$. This optimal $\mathbf{w}_{i-1}$ results in

$$J(i)|_{\mathbf{w}_{i-1}=\mathbf{w}^\circ} = E[\mathbf{v}_i^* \Pi \mathbf{v}_i]. \quad (9)$$

B. Adaptive Noise Constrained (ANC) Algorithms

Along this line of thought, we organize a constrained optimization problem incorporating the knowledge of $\mathbf{v}_i$. The optimum solution is obtained by minimizing $J(i)$ subject to $J(i) = E[\mathbf{v}_i^* \Pi \mathbf{v}_i]$. The augmented cost function, using a Lagrange multiplier λ, is given by

$$J_1(i) = J(i) + \gamma\lambda \left(J(i) - E[\mathbf{v}_i^* \Pi \mathbf{v}_i]\right) - \gamma\lambda^2, \quad (10)$$

where $\gamma > 0$. To get a unique critical λ, a term $-\gamma\lambda^2$ is used [6]. However, this cost function is based on the knowledge of $\mathbf{v}_i$. To avoid this unpractical obstacle, $E[\mathbf{v}_i^* \Pi \mathbf{v}_i]$ is replaced by an unknown variable ζ, which is adjusted at each iteration. Then the proposed cost function is given by

$$J_{\text{ANC}}(i) = J(i) + \gamma\lambda \left(J(i) - \zeta\right) - \gamma\lambda^2 + \rho\zeta^2, \quad (11)$$

where $\gamma, \rho > 0$. In (11), we know that the proposed cost function is minimized with respect to the weight and ζ, and maximized with respect to λ. Then the update equations are as follows:

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu_{\mathbf{w}} \nabla_{\mathbf{w}} J_{\text{ANC}} \quad (12)$$
$$\lambda_{i+1} = \lambda_i + \mu_\lambda \nabla_\lambda J_{\text{ANC}} \quad (13)$$
$$\zeta_{i+1} = \zeta_i - \mu_\zeta \nabla_\zeta J_{\text{ANC}}, \quad (14)$$

where $\mu_{\mathbf{w}}, \mu_\lambda$ and μ_ζ are positive parameter. The gradients in (12)–(14) are simply derived as

$$\nabla_{\mathbf{w}} J_{\text{ANC}} = -E[\mathbf{e}_i^* \Pi \mathbf{U}_i] \quad (15)$$
$$\nabla_\lambda J_{\text{ANC}} = \gamma \left(E[\mathbf{e}_i^* \Pi \mathbf{e}_i] - \zeta\right) - 2\gamma\lambda \quad (16)$$
$$\nabla_\zeta J_{\text{ANC}} = \gamma\lambda + 2\rho\zeta, \quad (17)$$

respectively. Replacing the expected values of (15)–(17) by its instantaneous values and substituting for (12)–(14), we have the following stochastic gradient based update algorithm:

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu_{\mathbf{w}}(1 + \gamma\lambda_i) \mathbf{U}_i^* \Pi \mathbf{e}_i \quad (18)$$
$$\lambda_{i+1} = \lambda_i + \mu_\lambda \gamma \left[\frac{1}{2}(\mathbf{e}_i^* \Pi \mathbf{e}_i - \zeta_i) - \lambda_i\right] \quad (19)$$
$$\zeta_{i+1} = \zeta_i - \mu_\zeta(\gamma\lambda_i - 2\rho\zeta_i). \quad (20)$$

As mentioned above, a different form of Π, i.e., $\Pi = \mathbf{I}$, diag$[1/\|\mathbf{u}_i\|^2, \dots, 1/\|\mathbf{u}_{i-K+1}\|^2]$, and $(\mathbf{U}_i \mathbf{U}_i^*)^{-1}$, results in the adaptive noise constrained (ANC) DR-LMS, NDR-LMS, and AP algorithms, respectively.

C. Properties of Proposed Algorithms

Let us consider the steady-state mean behavior of λ_i and ζ_i. By taking the expectation of steady-state value of both sides of (20), it leads to

$$E[\zeta_\infty] = E[\zeta_\infty] - \mu_\zeta(\gamma E[\lambda_\infty] - 2\rho E[\zeta_\infty]), \quad (21)$$
$$E[\zeta_\infty] = \frac{\gamma}{2\rho} E[\lambda_\infty], \quad (22)$$

TABLE I
EXPERIMENTAL PARAMETERS OF ADAPTIVE NOISE CONSTRAINED ALGORITHMS

ANC-DR-LMS	ANC-NDR-LMS	ANC-APA
$\mu_{\mathbf{w}} = 0.004$	$\mu_{\mathbf{w}} = 0.04$	$\mu_{\mathbf{w}} = 0.016$
$\mu_\lambda = 10^{-8}$	$\mu_\lambda = 10^{-7}$	$\mu_\lambda = 3 \times 10^{-6}$
$\mu_\zeta = 10^{-5}$	$\mu_\zeta = 10^{-8}$	$\mu_\zeta = 10^{-7}$
$\gamma = 2000$	$\gamma = 5000$	$\gamma = 2000$
$\rho = 100$	$\rho = 100$	$\rho = 100$

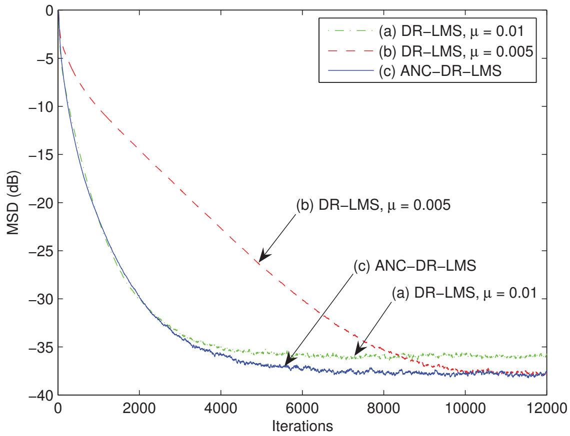

Fig. 1 Plots of MSD for the ANC-DR-LMS and the DR-LMS [K=4, input: ARMA(2,2)]

and

$$E[\lambda_\infty] = \frac{2\rho}{4\rho + \gamma} E[\mathbf{e}_\infty^* \Pi \mathbf{e}_\infty]. \quad (23)$$

In addition, from (20), we find that

$$E[\zeta_\infty] = \frac{\gamma}{2\rho} E[\lambda_\infty]. \quad (24)$$

From (23) and (24), we obtain the following relation as

$$E[\zeta_\infty] = \frac{2\rho}{4\rho + \gamma} E[\mathbf{e}_\infty^* \Pi \mathbf{e}_\infty]. \quad (25)$$

If $\gamma >> 4\rho$, ζ_i converges to $E[\mathbf{e}_\infty^* \Pi \mathbf{e}_\infty]$.

III. EXPERIMENTAL RESULTS

We illustrate the performance of the proposed algorithms by carrying out computer simulations in a channel identification scenario. The unknown channel $H(z)$ has 16 taps and is randomly generated. The adaptive filter and the unknown channel are assumed to have the same number of taps. The input signal is obtained by filtering a white, zero-mean, Gaussian random sequence through a first-order system

$$G(z) = \frac{1 + 0.5z^{-1} + 0.81z^{-2}}{1 - 0.59z^{-1} + 0.4z^{-2}}.$$

This results in a highly correlated Gaussian signal of which the eigenvalue spread is about 105. The signal-to-noise ratio (SNR) is calculated by

$$\text{SNR} = 10\log_{10}\left(E[y^2(i)]/E[v^2(i)]\right), \quad (26)$$

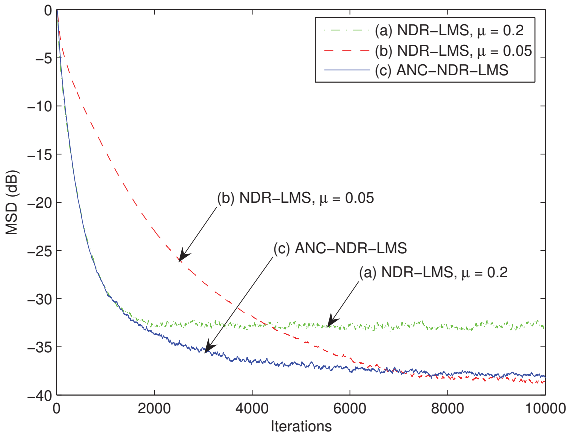

Fig. 2 Plots of MSD for the ANC-NDR-LMS and the NDR-LMS [K=4, input: ARMA(2,2)]

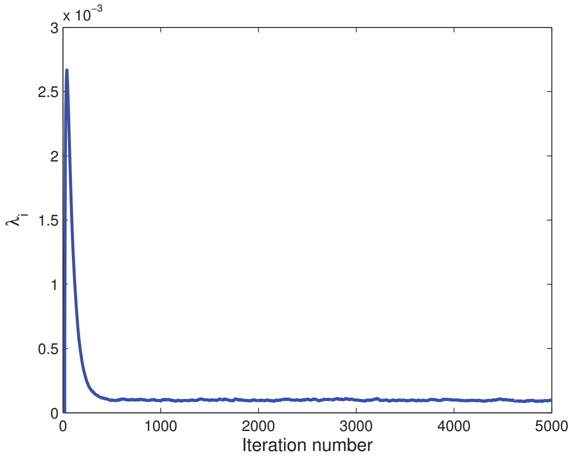

Fig. 4 Time evolution of λ_i of the ANC-APA

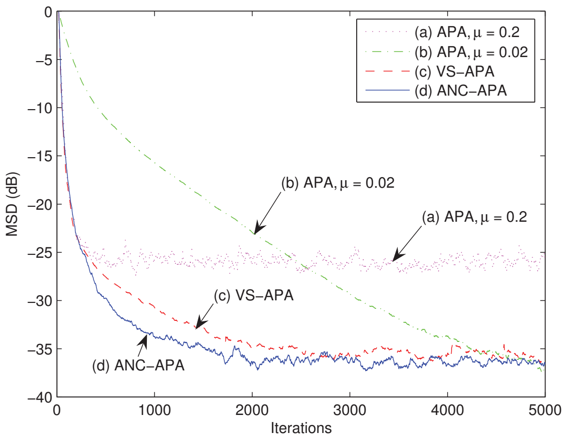

Fig. 3 Plots of MSD for the ANC-APA, the VS-APA [8] and the APA [K=4, input: ARMA(2,2)]

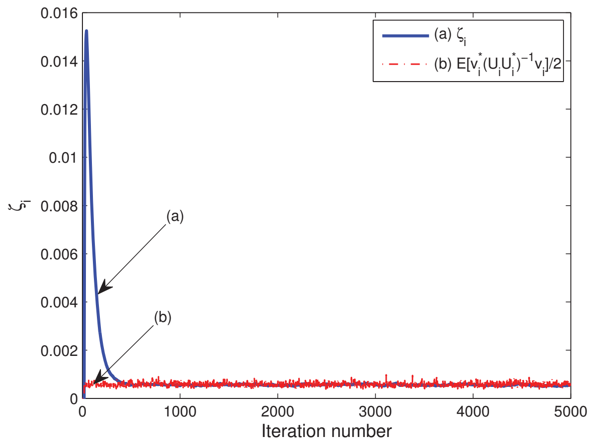

Fig. 5 Time evolution of ζ_i of the ANC-APA

where $y(i) = \mathbf{u}_i\mathbf{w}^{\circ}$. The measurement noise $v(i)$ is added to $y(i)$ such that SNR $=$ 30dB. The mean square deviation (MSD), $E\|\mathbf{w}^{\circ} - \mathbf{w}_i\|^2$, is taken and averaged over 100 independent trials. The parameters used for the proposed algorithms are shown in Table. I.

Fig. 1 indicates the MSD curves of the DR-LMS and the ANC-DR-LMS. Dashed lines indicate the results of the DR-LMS with fixed step-sizes when we choose $\mu = 0.01$ and 0.005. As can be seen, the ANC-DR-LMS has the faster convergence and lower misadjustment than the standard DR-LMS. Fig. 2 shows the MSD curves of the NDR-LMS and the ANC-NDR-LMS. We choose the step-sizes of the NDR-LMS as $\mu = 0.2$ and 0.05. A similar result of Fig. 1 is observed in Fig. 2. Fig. 3 indicates the performance of the APA and the ANC-APA. For a comparison purpose, the variable step-size APA (VS-APA) [8] is presented. we use $C = 0.01$ and $\mu_{\max} = 1.0$ for the VS-APA, which are defined in [8]. We know that the ANC-APA outperforms the APA and is comparable to the VS-APA.

In Fig. 4, the transient behaviour of λ_i of the ANC-APA is depicted. It exhibits the time evolution of λ_i which increases rapidly and then converges. Fig. 5 indicates that ζ_i converges to $E[\mathbf{v}_i^*(\mathbf{U}_i\mathbf{U}_i^*)^{-1}\mathbf{v}_i]/2$ which is the noise related constraint.

IV. Conclusion

In this paper, we present adaptive noise constrained (ANC) DR-LMS, NDR-LMS and AP algorithms, which incorporate a knowledge of noise without a prior information of noise. A cost function based on the adaptive noise constrained optimization using Lagrangian multipliers is introduced to estimate a noise information. As a result, the convergence performance of data-reusing and affine projection algorithms is greatly improved.

References

[1] S. Haykin, *Adaptive Filter Theory*, 4th edition, Upper Saddle River, NJ: Prentice Hall, 2002.

[2] H.-C. Shin, W.-J. Song and A. H. Sayed, "Mean-square performance of data-reusing algorithms," *IEEE Signal Processing Lett.*, vol. 12, pp. 851–854, Dec. 2005.
[3] B. A. Schnaufer and W. K. Jenkins, "New data-reusing LMS algorithms for improved convergence," in *Proc. Asilomar Conf.*, Pacific Groves, CA, May 1993, pp. 1584–1588.
[4] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electro. Commun. Jpn.*, vol. 67-A, no. 5, pp. 19–27, 1984.
[5] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 52, pp. 90–102, Jan. 2004.
[6] Y. Wei, S. B. Gelfand and J. V. krogmeier, "Noise-constrained least mean square algorithm," *IEEE Trans. Signal Processing*, vol. 49, No. 9, pp. 1961–1970, Sep. 2001.
[7] S. Y. Choi, T.-W. Lee and D. S. Hong, "Adaptive error-constrained method for LMS algotihms and applications," *Signal Processing*, vol. 85, pp. 1875–1897, Oct. 2005.
[8] H.-C. Shin, A. H. Sayed and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Processing Lett.*, vol. 11, pp. 132–135, Feb. 2004.