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

New Data Reuse Adaptive Filters with Noise Constraint

Young-Seok Choi

Abstract—We present a new framework of the data-reusing (DR) adaptive algorithms by incorporating a constraint on noise, referred to as a *noise constraint*. The motivation behind this work is that the use of the statistical knowledge of the channel noise can contribute toward improving the convergence performance of an adaptive filter in identifying a noisy linear finite impulse response (FIR) channel. By incorporating the noise constraint into the cost function of the DR adaptive algorithms, the noise constrained DR (NC-DR) adaptive algorithms are derived. Experimental results clearly indicate their superior performance over the conventional DR ones.

Keywords—Adaptive filter, data-reusing, least-mean square (LMS), affine projection (AP), noise constraint.

I. Introduction

THE least mean square (LMS)-type adaptive algorithms yield a deteriorated convergence performance for the highly correlated input regressors [1]. Recently, the data-reusing LMS (DR-LMS) and normalized DR-LMS (NDR-LMS) and affine projection (AP) algorithms have drawn attention among researchers desiring to address the poor convergence issue of the LMS-type adaptive filters [2]–[5]. Comparing with the LMS-type adaptive algorithms, the aforementioned algorithms utilize a block error and a block input regressor for updating the filter weight, accomplishing faster convergence. Nevertheless, the DR adaptive algorithms also suffer from a tradeoff between convergence rate and steady-state misalignment.

When identifying a noisy linear finite impulse response (FIR) channel, the use of noise statistics can play a considerable role in improving the convergence behavior of the adaptive filter [6],[7]. Motivated by this, we make use of the knowledge of the channel noise, more specifically the noise power which is either known or estimated, so as to address a conflicting requirement of fast convergence and low misalignment of the DR adaptive algorithms. By imposing a constraint on noise, referred to as a *noise constraint*, on the unified cost function of the DR adaptive algorithms, an augmented Lagrangian is formulated. Solving the augmented Lagrangian leads to the *noise constrained* DR (NC-DR) adaptive algorithms, which yield a variable step-size feature. Through experiments, we demonstrate that the proposed NC-DR adaptive algorithms posses superiority over the conventional DR ones by carrying out the tradeoff between fast convergence and low misalignment.

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. Data-reusing (DR) Adaptive Algorithms

Consider reference 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. For compact manipulation we also define an input signal matrix and a desired signal vector as

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

Then the error vector is compactly written as

$$\mathbf{e}_i = \mathbf{d}_i - U_i \mathbf{w}_{i-1}. \tag{3}$$

In [8], the unified cost function of the DR adaptive algorithms is given by

$$J(i) = E[\mathbf{e}_i^* \Pi \mathbf{e}_i] \tag{4}$$

where Π is a positive definite matrix. Taking derivatives of $J(i)$ with respect to $\mathbf{w}_{i-1}$ and replacing the expected value by its instantaneous value, we obtain the following stochastic gradient based recursion

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu U_i^* \Pi \mathbf{e}_i \tag{5}$$

where μ is the step-size parameter. Choosing a different parameter Π leads to the distinct DR adaptive algorithms, i.e., the DR-LMS, NDR-LMS, and AP algorithms can be obtained as follows [8]:

$$\begin{aligned} \mathrm{DR-LMS}: \mathbf{w}_i &= \mathbf{w}_{i-1} + \mu U_i^* \mathbf{e}_i, & (6)\\ \mathrm{NDR-LMS}: \mathbf{w}_i &= \mathbf{w}_{i-1} + \mu U_i^* D_i \mathbf{e}_i, & (7)\\ \mathrm{AP}: \mathbf{w}_i &= \mathbf{w}_{i-1} + \mu U_i^* (U_i U_i^*)^{-1} \mathbf{e}_i. & (8) \end{aligned}$$

respectively, where $D_i = \mathrm{diag}[1/\|\mathbf{u}_i\|^2, \ldots, /\|\mathbf{u}_{i-K+1}\|^2]$.

III. Noise Constrained DR (NC-DR) Adaptive Algorithms

Under the model (1), minimizing the unified cost function (4) over $\mathbf{w}$ gives the optimal weight $\mathbf{w} = \mathbf{w}^\circ$ and yields

$$J(i)|_{\mathbf{w}=\mathbf{w}^\circ} = E[\mathbf{v}_i^* \Pi \mathbf{v}_i], \tag{9}$$

where $\mathbf{v}_i = [v(i)\ v(i-1) \ldots v(i-K+1)]^T$. In [8], it is evident that the transient behavior of the weight of the DR adaptive filter is dependent on the statistics of the channel noise, suggesting the usefulness of noise in finding the optimal weight. In addition, it is known that the use of the statistical information of the channel noise enables the LMS-type adaptive filter to improve the convergence performance [6].

Inspired by these, we formulate a constrained optimization criterion incorporating the knowledge of the channel noise as follows: Minimizing $J(i)$ subject to $J(i) = E[\mathbf{v}_i^* \Pi \mathbf{v}_i]$. An augmented cost function, using a Lagrange multiplier λ, is given by

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

The critical values of (10) are $\mathbf{w} = \mathbf{w}^\circ$ for any λ. This situation may cause convergence problem [6]. To address this, a term $-\gamma\lambda^2$ $(\gamma > 0)$ is incorporated into (10), resulting in a new augmented Lagrangian as follows:

$$J_{\text{NC}}(i) = J(i) + \gamma\lambda \left(J(i) - E[\mathbf{v}_i^* \Pi \mathbf{v}_i] \right) - \gamma\lambda^2. \tag{11}$$

The uses of the γ and $\gamma\lambda^2$ in (11) ensure that the unique critical value of $J_{\text{NC}}(i)$ is $(\mathbf{w}, \lambda) = (\mathbf{w}^\circ, 0)$. In (11), the augmented Lagrangian (11) is minimized with respect to the weight $\mathbf{w}$ and maximized with respect to λ, respectively. By applying the Robbnis-Munro method [9], the weight and λ are updated as follows:

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} - \alpha \nabla_{\mathbf{w}} J_{\text{NC}} &(12a)\\ \lambda_i &= \lambda_{i-1} + \beta \nabla_{\lambda} J_{\text{NC}} &(12b)\end{aligned}$$

where α and β are the positive learning parameters. The gradient terms in (12a)–(12b) are simply derived as

$$\begin{aligned} \nabla_{\mathbf{w}} J_{\text{NC}} &= -(1+\gamma\lambda) E[U_i^* \Pi \mathbf{e}_i] &(13a)\\ \nabla_{\lambda} J_{\text{NC}} &= \gamma \left(E[\mathbf{e}_i^* \Pi \mathbf{e}_i] - E[\mathbf{v}_i^* \Pi \mathbf{v}_i] \right) - 2\gamma\lambda, &(13b)\end{aligned}$$

respectively.

Replacing the expected values of (13a)–(13b) by their instantaneous values, substituting for (12a)–(12b) and replacing $\beta\gamma$ by $\beta/2$, we obtain the following update recursions:

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i U_i^* \Pi \mathbf{e}_i, &(14a)\\ \alpha_i &= \alpha(1+\gamma\lambda_i), &(14b)\\ \lambda_{i+1} &= \lambda_i + \beta \left[\frac{1}{2} (\mathbf{e}_i^* \Pi \mathbf{e}_i - E[\mathbf{v}_i^* \Pi \mathbf{v}_i]) - \lambda_i \right] &(14c)\end{aligned}$$

Finally, (14a)–(14c) are referred to as the *noise constrained* DR (NC-DR) adaptive algorithm. By choosing Π, we can get a new family of the DR adaptive algorithms described in the following subsections.

A. Noise Constrained Data-reusing LMS (NC-DR-LMS) Algorithm

In case of $\Pi = I$, we obtain the NC-DR-LMS algorithm by deriving

$$E[\mathbf{v}_i^* \mathbf{v}_i] = K\sigma_v^2, \tag{15}$$

where $\sigma_v^2 = E[v(i)^2]$ is the power of the channel noise.

Substituting (15) into (14c) leads to the NC-DR-LMS algorithm, being written by

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i U_i^* \mathbf{e}_i &(16a)\\ \alpha_i &= \alpha(1+\gamma\lambda_i) &(16b)\\ \lambda_{i+1} &= \lambda_i + \beta \left[\frac{1}{2} (\mathbf{e}_i^* \mathbf{e}_i - K\sigma_v^2) - \lambda_i \right]. &(16c)\end{aligned}$$

B. Noise Constrained Normalized DR LMS (NC-NDR-LMS) Algorithm

If $\Pi = D_i = \text{diag}[1/\|\mathbf{u}_i\|^2, \ldots, 1/\|\mathbf{u}_{i-K+1}\|^2]$, we arrive at

$$E[\mathbf{v}_i^* D_i \mathbf{v}_i] = \frac{K\sigma_v^2}{M\sigma_u^2(i)}, \tag{17}$$

where $\sigma_u^2 = E[u(i)^2]$ and it can be computed by

$$\sigma_u^2(i) = \theta\sigma_u^2(i-1) + (1-\theta)u^2(i) \tag{18}$$

with $0 \le \theta < 1$.

Then, the NC-NDR-LMS algorithm is given by

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i U_i^* D_i \mathbf{e}_i &(19a)\\ \alpha_i &= \alpha(1+\gamma\lambda_i) &(19b)\\ \lambda_{i+1} &= \lambda_i + \beta \left[\frac{1}{2} \left(\mathbf{e}_i^* D_i \mathbf{e}_i - \frac{K\sigma_v^2}{M\sigma_u^2(i)} \right) - \lambda_i \right]. &(19c)\end{aligned}$$

C. Noise Constrained Affine Projection (NC-AP) Algorithm

The NC-AP algorithm can be obtained from (14a)–(14c) in case of $\Pi = (U_i^* U_i)^{-1}$. Assuming that the diagonal components of $U_i^* U_i$ are much larger than the off-diagonal components [10], we arrive at

$$E[(U_i^* U_i)^{-1}] \approx E\{\text{diag}[1/\|\mathbf{u}_i\|^2, \ldots, 1/\|\mathbf{u}_{i-K+1}\|^2]\} = \frac{1}{M\sigma_u^2(i)} I,$$

since $E[\|\mathbf{u}_i\|^2] = M\sigma_u^2$, where $\sigma_u^2 = E[u^2(i)]$ which can be estimated as (18). By doing this, we get

$$E[\mathbf{v}_i^* (U_i^* U_i)^{-1} \mathbf{v}_i] \approx \frac{1}{M\sigma_u^2(i)} E[\mathbf{v}_i^* \mathbf{v}_i] = \frac{K\sigma_v^2}{M\sigma_u^2(i)}, \tag{20}$$

under the realistic assumption that the noise $v(i)$ is i.i.d. and statistically independent of the input matrix U_i [5].

Then, the resultant NC-AP algorithm is given by

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} + \alpha_i U_i^* (U_i^* U_i)^{-1} \mathbf{e}_i &(21a)\\ \alpha_i &= \alpha(1+\gamma\lambda_i) &(21b)\\ \lambda_{i+1} &= \lambda_i + \beta \left[\frac{1}{2} \left(\mathbf{e}_i^* (U_i^* U_i)^{-1} \mathbf{e}_i - \frac{K\sigma_v^2}{M\sigma_u^2(i)} \right) - \lambda_i \right]. &(21c)\end{aligned}$$

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

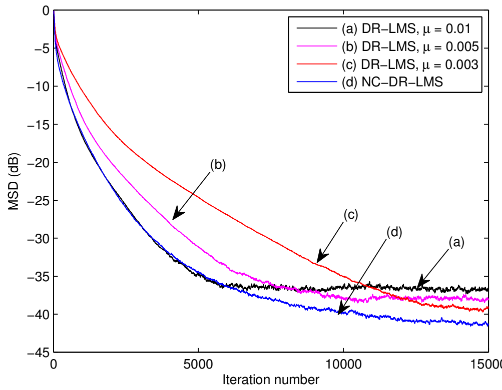


Fig. 1 MSD curves of DR-LMS and NC-DR-LMS [K = 4; M = 16]

IV. Experimental Results

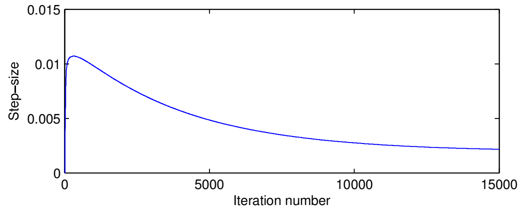
We illustrate the performance of the NC-DR algorithms by carrying out computer simulations in a channel identification scenario. The unknown channel $H(z)$ has 16 taps and is randomly generated. The adaptive filter and the unknown channel are assumed to have the same number of taps. The input signal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

$$G(z) = \frac{1 + 0.5z^{-1} + 0.81z^{-2}}{1 - 0.59z^{-1} + 0.4z^{-2}}. \tag{22}$$

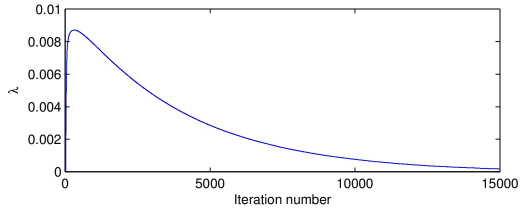
This results in a highly correlated Gaussian signal of which the eigenvalue spread is about 105. The signal-to-noise ratio (SNR) is computed by $10\log_{10}(E[y(i)^2]/E[v(i)^2])$, where $y(i) = \mathbf{u}_i\mathbf{w}^\circ$. The measurement noise $v(i)$ is added to $y(i)$ such that SNR $= 30$dB. In this study, it is assumed that the power of the channel noise, σ_v^2, is known *a priori* since it can be easily estimated in practical applications [11]. The mean square deviation (MSD), $E\|\mathbf{w}^\circ - \mathbf{w}_i\|^2$, is taken and averaged over 100 independent trials. In the NC-NDR-LMS and NC-AP algorithms, $\theta = 0.99$ is used.

Fig. 1 shows the MSD curves of the DR-LMS and NC-DR-LMS algorithms for $K = 4$. Here, we choose $\alpha = 0.004$, $\beta = 2 \times 10^{-4}$, and $\gamma = 500$ for the NC-DR-LMS and $\mu = 0.01$, 0.005, and 0.003 for the DR-LMS, respectively. As can be seen, the NC-DR-LMS is superior to the DR-LMS in terms of both the convergence rate and the misalignment. In Fig. 2(a), the transient behavior of the step-size, α_i, of the NC-DR is depicted. The step-size, α_i, has large value at initial period and becomes small, achieving fast convergence as well as low misalignment. Fig. 2(b) exhibits the evolution of λ_i which increases rapidly and then converges to zero.

Fig. 3 exhibits the MSD curves of the NDR-LMS and NC-NDR-LMS for $K = 4$. We use $\alpha = 0.05$, $\beta = 1.5 \times 10^{-3}$, and $\gamma = 5000$ for the NC-NDR-LMS and $\mu = 0.5$, 0.2, and 0.07 for the NDR-LMS, respectively. We can see that the NC-NDR-LMS possesses both rapid and accurate convergency. Next, Figs. 4(a) and 4(b) illustrate the time evolution curves of α_i and λ_i of the NC-NDR-LMS, respectively. Similar results with Figs. 2(a) and 2(b) are observed in Figs. 4(a) and 4(b).



(a) Time evolution of step-size, α_i, for NC-DR-LMS



(b) Time evolution of λ_i for NC-DR-LMS

Fig. 2 Behavior of parameters for NC-DR-LMS

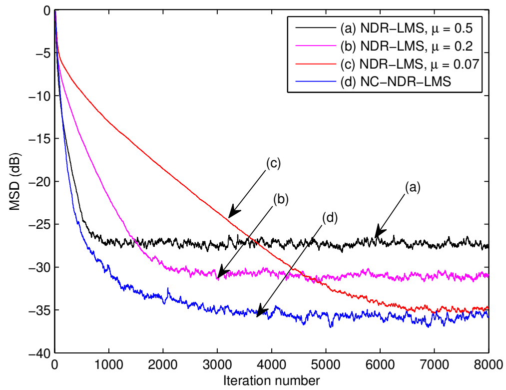
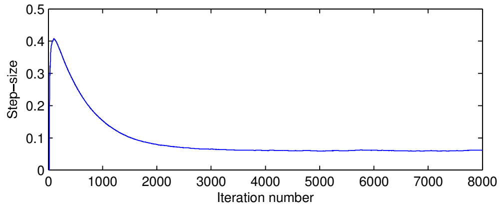


Fig. 3 MSD curves of NDR-LMS and NC-NDR-LMS [K = 4; M = 16]

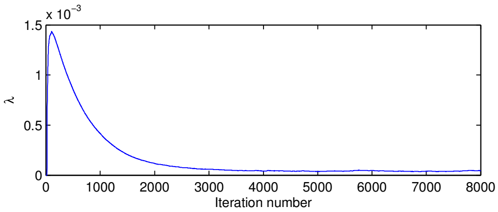
In Fig. 5, the MSD curves of the NC-AP and other conventional AP algorithms for $K = 4$ are depicted. For comparison purpose, we apply a recently developed variable step-size AP (VSS-AP) algorithm [5], which yields $\mu(i) = \mu_{\max} \cdot \frac{\|\hat{\mathbf{p}}_i\|^2}{\|\hat{\mathbf{p}}_i\|^2 + C}$, where $C \approx K/SNR$. The parameters, $C = 5 \times 10^{-3}$ and $\mu_{\max} = 1.0$, are used for the VSS-AP. We choose $\mu = 0.3$, 0.1, and 0.03 for the AP and $\alpha = 0.025$, $\beta = 10^{-3}$, and $\gamma = 5000$ for the NC-AP. In the figure, the NC-AP addresses the trade-off between the convergence rate and the misalignment in the AP and outperforms the VSS-AP. Fig. 6(a) illustrates the time evolutions of step-sizes for the NC-APA and VSS-AP algorithms, revealing the dynamic behavior of the NC-AP. Fig. 6(b) shows the corresponding behavior of λ_i of the NC-AP.

V. Conclusion

We have presented a noise constrained framework of the DR adaptive algorithms by incorporating the statistical



(a) Time evolution of step-size, α_i, for NC-NDR-LMS



(b) Time evolution of λ_i for NC-NDR-LMS

Fig. 4 Behavior of parameters for NC-NDR-LMS

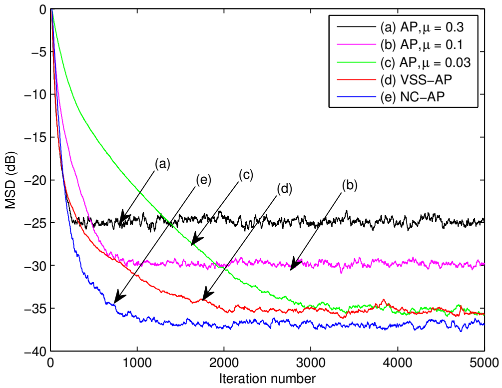
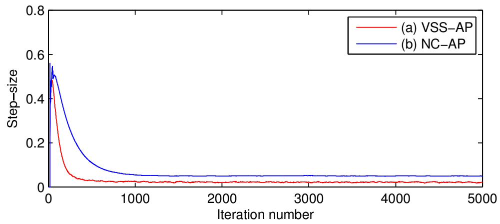
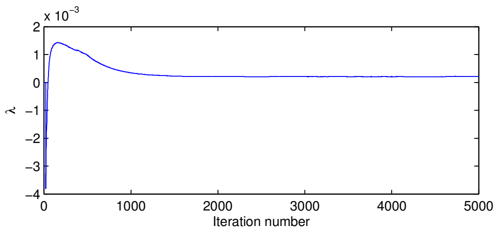


Fig. 5 MSD curves of AP, VSS-AP, and NC-AP [K = 4; M = 16]



(a) Time evolutions of step-sizes for VSS-AP and NC-AP



(b) Time evolution of λ_i for NC-AP

Fig. 6 Behavior of parameters for VSS-AP and NC-AP

information of the channel noise into adaptation. Solving a augmented Lagrangian with a noise constraint leads to a time-varying step-size feature in the DR adaptive algorithms, thus providing a compromise between fast convergence and low misalignment. Experiment results have shown that the proposed DR adaptive algorithms performs better than the conventional DR counterparts.

References

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*, New York: Wiley, 2003.

[2] B. A. Schnaufer and W. K. Jenkins, "New data-reusing LMS algorithms for improved covergence," in *Proc. Asilomar Conf.*, Pacific Groves, CA, pp. 1584–1588, May 1993,

[3] R. A. Soni, K. A. Gallivan and W. K. Jenkins, "Convergence properties of affine projection and normalized data reusing methods,," in *Proc. Asilomar Conf.*, vol. 2, pp. 1166–1170, Nov. 1998,

[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 Process.*, vol. 12, no. 1, pp. 90–102, Jan. 2004.

[6] Y. Wei, S. B. Gelfand and J. V. Krogmeier, "Noise-constrained least mean squares algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1961–1970, Sep. 2001.

[7] A. Zerguine, M. Moinuddin and S. A. A. Imam, "A noise constrained least mean fourth adaptive algorithm," *Signal Process.*, vol. 91, no. 1, pp. 136–149, Jan. 2011.

[8] H.-C. Shin, W.-J. Song and A. H. Sayed, "Mean-square performance of data-reusing adaptive algorithms," *IEEE Trans. Signal Process.*, vol. 12, no. 12, pp. 851–854, Dec. 2005.

[9] M. Dufflo, *Random Iteratie Models*, Springer-Verlag, Berlin, Germany, 1997.

[10] K.-Y. Hwang and W.-J. Song, "An affine projection adaptive filtering algorithm with selective regressors," *IEEE Trans. Circuits Syst. II*, vol. 54, no. 1, pp. 43–46, Jan. 2007.

[11] S.-E. Kim, Y.-S. Choi, M.-K. Song and W.-J. Song, "A subband adaptive filtering algorithm employing dynamic selection of subband filters," *IEEE Trans. Signal Process.*, vol. 17, no. 3, pp. 245–248, Mar. 2010.