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.

II. ADAPTIVE NOISE CONSTRAINED DR AND AP ALGORITHMS

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

$$d(i) = u_i w^o + v(i),$$

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

$$u_i = [u(i) \ u(i-1) \ \cdots \ u(i-M+1)],$$

and $u_i$ and $v(i)$ are uncorrelated.

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

Let $w_i$ be an estimate for $w^o$ at iteration $i$. The DR-LMS, NDR-LMS, and APA take the forms [2]

$$w_i = w_{i-1} + \mu U_i^* e_i,$$

$$w_i = w_{i-1} + \mu U_i^* D_i e_i,$$

$$w_i = w_{i-1} + \mu U_i^* (U_i U_i^* + \delta I)^{-1} e_i,$$

respectively, where

$$U_i = \begin{bmatrix} u_i \\ u_{i-1} \\ \vdots \\ u_{i-K+1} \end{bmatrix}, \quad d_i = \begin{bmatrix} d(i) \\ d(i-1) \\ \vdots \\ d(i-K+1) \end{bmatrix}$$

$$e_i = d_i - U_i w_{i-1}, \quad D_i = \text{diag}[1/\|u_i\|^2, \ldots, 1/\|u_{i-K+1}\|^2].$$

$\mu$ is the step-size, and $*$ denotes the Hermitian transpose. It is known [2] that these algorithms are obtained to minimize the following cost function

$$J(i) = E[|e_i^*|^2],$$

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

$$\frac{\partial J(i)}{\partial w_{i-1}} = -2E[e_i^* \Pi U_i].$$

Then, we can obtain the stochastic gradient algorithm

$$w_i = w_{i-1} + \mu U_i^* \Pi e_i$$

Note that the choice of $\Pi$ 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, $D_i$, we get the NDR-LMS (4). Also the choice $\Pi = (U_i U_i^*)^{-1}$ results in the APA (5).
The minimization of (6) over \( \mathbf{w}_{i-1} \) yields \( \mathbf{w}_{i-1} = \mathbf{w}^* \) and \( \mathbf{e}_i = \mathbf{v}_i \), where \( \mathbf{v}_i = [v(i) \, v(i-1) \ldots v(i-K+1)]^T \). This optimal \( \mathbf{w}_{i-1} \) results in
\[
J(i)|_{\mathbf{w}_{i-1}=\mathbf{w}^*} = E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i].
\] (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
\[
J(i) = E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i].
\]

The augmented cost function, using a Lagrange multiplier \( \lambda \), is given by
\[
J_1(i) = J(i) + \gamma \lambda (J(i) - E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i]) - \gamma \lambda^2,
\] (10)

where \( \gamma > 0 \). To get a unique critical \( \lambda \), 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, \( E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i] \) is replaced by an unknown variable \( \zeta \), which is adjusted at each iteration. Then the proposed cost function is given by
\[
J_{\text{ANC}}(i) = J(i) + \gamma \lambda (J(i) - \zeta) - \gamma \lambda^2 + \rho \zeta^2,
\] (11)

where \( \gamma, \rho > 0 \). In (11), we know that the proposed cost function is minimized with respect to the weight and \( \zeta \), and maximized with respect to \( \lambda \). Then the update equations are as follows:
\[
\begin{align*}
\mathbf{w}_i &= \mathbf{w}_{i-1} - \mu \mathbf{w} \nabla_w J_{\text{ANC}} \\
\lambda_{i+1} &= \lambda_i + \mu \lambda \nabla_\lambda J_{\text{ANC}} \\
\zeta_{i+1} &= \zeta_i + \mu \zeta \nabla_\zeta J_{\text{ANC}},
\end{align*}
\] (12)-(14)

where \( \mu_w, \mu_\lambda \) and \( \mu_\zeta \) are positive parameter. The gradients in (12)–(14) are simply derived as
\[
\begin{align*}
\nabla_w J_{\text{ANC}} &= -E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i] \\
\nabla_\lambda J_{\text{ANC}} &= \gamma (E[\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i] - \zeta) - 2\gamma \lambda \\
\nabla_\zeta J_{\text{ANC}} &= \gamma \lambda + 2\rho \zeta,
\end{align*}
\] (15)-(17)

respectively. Replacing the expected values of (15)–(17) by their instantaneous values and substituting for (12)–(14), we have the following stochastic gradient based update algorithm:
\[
\begin{align*}
\mathbf{w}_i &= \mathbf{w}_{i-1} - \mu \mathbf{w}(1 + \gamma \lambda_i) \mathbf{U}_i^T \mathbf{P} \mathbf{v}_i \\
\lambda_{i+1} &= \lambda_i + \mu \lambda \left[ \frac{1}{2} (\mathbf{v}_i^T \mathbf{P} \mathbf{v}_i - \zeta_i) - \lambda_i \right] \\
\zeta_{i+1} &= \zeta_i - \mu_\zeta (\gamma \lambda_i - 2\rho \zeta_i).
\end{align*}
\] (18)-(20)

As mentioned above, a different form of \( \mathbf{P} \), i.e., \( \mathbf{P} = \mathbf{I}, \) \( \text{diag}[1/|\mathbf{u}_1|^2, \ldots, 1/|\mathbf{u}_{i-K+1}|^2] \), and \( \mathbf{U}_i \mathbf{U}_i^T \), 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 \( \lambda_i \) and \( \zeta_i \). By taking the expectation of steady-state value of both sides of (20), it leads to
\[
E[\zeta_\infty] = E[\zeta_\infty] - \mu_\zeta (\gamma E[\lambda_\infty] - 2\rho E[\zeta_\infty]),
\] (21)
and
\[
E[\lambda_\infty] = \frac{2\rho}{4\rho + \gamma} E[\mathbf{v}_\infty^T \mathbf{P} \mathbf{v}_\infty].
\] (23)

In addition, from (20), we find that
\[
E[\zeta_\infty] = \gamma E[\lambda_\infty].
\] (24)

From (23) and (24), we obtain the following relation as
\[
E[\zeta_\infty] = \frac{2\rho}{4\rho + \gamma} E[\mathbf{v}_\infty^T \mathbf{P} \mathbf{v}_\infty].
\] (25)

If \( \gamma >> 4\rho \), \( \zeta_i \) converges to \( E[\mathbf{v}_\infty^T \mathbf{P} \mathbf{v}_\infty] \).

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
\[
\text{SNR} = 10 \log_{10} \left( E[y^2(i)]/E[x^2(i)] \right),
\] (26)
where \( y(i) = u_i w^o \). The measurement noise \( v(i) \) is added to \( y(i) \) such that SNR = 30dB. The mean square deviation (MSD), \( E\|w^o - w_i\|^2 \), 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 \( \mu_{\text{max}} = 1.0 \) 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 \( \lambda_i \) of the ANC-APA is depicted. It exhibits the time evolution of \( \lambda_i \) which increases rapidly and then converges. Fig. 5 indicates that \( \zeta_i \) converges to \( E[v_i^* (U_i U_i^*)^{-1} v_i]/2 \) 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


