ABSTRACT

The reuse of past coefficient vectors of the NLMS for reducing the steady-state MSD in a low signal-to-noise ratio (SNR) was proposed recently. Its convergence analysis has not been studied yet, so we first derive a steady-state analysis for the NLMS with reusing coefficient vectors for a special case. In addition, this approach slows down the convergence speed while decreasing the steady-state MSD in proportion to the number of reusing coefficient vectors. To address this trade-off, we propose a novel NLMS algorithm which can change the reusing order to achieve both fast convergence speed and low steady-state MSD. The reusing order is decreased or increased by comparing the squared output error with a threshold. The experimental results show that the theoretical results match well with simulation results and the proposed algorithm has fast convergence speed and small steady-state MSD compared to the conventional NLMS.

Index Terms—Adaptive filters, normalized least-mean-square (NLMS), coefficient vector reusing, mean-square deviation (MSD), steady-state analysis, variable reusing order

1. INTRODUCTION

The normalized least mean square (NLMS) algorithm is among the most widely used adaptive filters due to its low computational complexity and ease of implementation [1], [2]. To improve convergence performance of the conventional NLMS algorithm, many researchers have proposed various approaches such as variable step-size (VSS) [3], variable regularization (VR) [4], and set-membership filtering (SMF) [5]. The improved NLMS algorithms employing these schemes have achieved fast convergence speed and small steady-state error.

Recently, an approach reusing past coefficient vectors of the NLMS algorithm was proposed in [6], which we refer to as the NLMS with reusing coefficient-vectors (NLMS-RC). This approach minimizes the summation of each squared Euclidean norm of difference between the update coefficient vector and past coefficient vectors. The NLMS-RC algorithm has small misalignment and its convergence speed is almost the same as the conventional NLMS algorithm in cases where high measurement noise exists. However, the steady-state analysis has not been covered yet in the literature. Since the number of reusing coefficient vectors varies according to the status of the adaptive filter, we refer to it as the NLMS with variable reusing coefficient-vectors (NLMS-VRC). The experimental results demonstrate that the proposed NLMS-VRC algorithm has both fast convergence speed and small steady-state MSD compared to the conventional NLMS.

2. PROPOSED NLMS ALGORITHM

Consider a desired signal \(d(i)\) that arise from an unknown linear system

\[
d(i) = u_i w^o + e(i),
\]

where \(w^o\) is an unknown column vector to be identified with an adaptive filter, \(e(i)\) corresponds to measurement noise which is assumed to be white Gaussian noise with zero mean and variance \(\sigma_v^2\), and \(u_i\) denotes a row input vector with length \(M\) as follows:

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

Let us define \(w_i\) as an estimate for \(w^o\) with length \(M\) at iteration \(i\).

2.1. Review of the Conventional NLMS with Coefficient Vector Reusing

Cho et. al. [6] have proposed a minimization problem for the NLMS which reuses past \(n\) coefficient vectors as follows:

\[
\min_w \sum_{k=1}^{n} \rho^{k-1} ||w_i - w_{i-k}||^2 \text{ subject to } d(i) = u_i w_i,
\]

where \(0 < \rho < 1\) and \(n\) is a positive integer. Then, the modified NLMS with reusing coefficient-vectors (NLMS-RC) reaches the reduced steady-state MSD and the update equation when \(\rho\) is one is given by

\[
w_i = \frac{1}{n} \sum_{k=1}^{n} w_{i-k} + \mu \frac{u_i^T}{||u_i||^2} \left( d(i) - u_i \frac{1}{n} \sum_{k=1}^{n} w_{i-k} \right),
\]

This work was supported by the Brain Korea 21 Project in 2010, Korea.
where $\mu$ is a step-size and we refer to $n$ as a reusing order for the NLMS-RC. In the NLMS-RC, the averaging effect of past coefficient vectors can prevent the updated coefficient vector from heavily fluctuating when the measurement noise is high. In that case, the MSD of the NLMS-RC gets lower without degrading convergence speed. However, if the SNR gets higher, the convergence speed gets slower as $n$ increases. Although the behavior of the NLMS-RC algorithm was geometrically interpreted, the theoretical convergence analysis has not been covered yet. We thus provide a steady-state analysis.

2.2. Steady-State Analysis

We consider energy conservation arguments presented in [7]. For tractable analysis, we assume the following statistics of signals:

**Assumption I** The noise $v(i)$ is independent and identically distributed (i.i.d.) and statistically independent of the input vector $\{u_i\}$.

**Assumption II** $\bar{w}_{i-k}$ is independent of $u_i^T u_i/\|u_i\|^2$ for $k = 1, \ldots, n$.

By the procedure presented in [7], the weighted variance relation for the NLMS-RC is obtained as

$$E[\|\bar{w}_i\|^2] = E\left[\frac{1}{n} \sum_{k=1}^{n} \|\bar{w}_{i-k}\|^2\right] + \mu^2\sigma_v^2(\gamma^T \gamma)$$

(4)

where $\bar{w}_i = w^o - w_i$,

$$F = I - \mu (E[P_1^T] \otimes I + I \otimes E[P_1]) + \mu^2 E[P_1^T \otimes P_1],$$

$$P_1 = u_i^T u_i/\|u_i\|^2,$$

$$\gamma = \text{vec}(E[u_i^T u_i/\|u_i\|^4]),$$

and $\sigma = \text{vec}(\Sigma)$ for an arbitrary Hermitian positive-definite matrix $\Sigma$. The $\text{vec}\{\cdot\}$ notation replaces an $M \times M$ arbitrary matrix by an $M^2 \times 1$ column vector or an $M \times M$ matrix. The first term on the right-hand side of (4) makes differences from the NLMS. Since calculating the first term for a general $n$ is not easy, we here derive the steady-state analysis for specific cases, i.e., $n = 2$ and 3.

For $n = 2$, the first term is represented by

$$E\left[\frac{1}{2} \sum_{k=1}^{2} \|\bar{w}_{i-k}\|^2\right] = E\left[\frac{1}{4} E]\|\bar{w}_{i-1}\|^2 + E\left[\frac{1}{4} E]\|\bar{w}_{i-2}\|^2\right]$$

$$+ \mu^2 E[\bar{w}_{i-1}^T \Sigma' \bar{w}_{i-2}].$$

(5)

where $\Sigma' = F \sigma$ and $\Sigma'_{\gamma} = \text{vec}(\sigma')$. The last term on the right-hand side of (5) should be newly calculated. To obtain the last term, we consider the following relation.

For $n = 2$, $\bar{w}_i$ can be rewritten as

$$\bar{w}_i = \frac{1}{2} (I - \mu P_1) (\bar{w}_{i-1} + \bar{w}_{i-2}) - \mu u_i^T v(i).$$

(6)

Then, $E[\bar{w}_i^T A \bar{w}_{i-1}]$ for any Hermitian matrix $A$ is given by

$$E[\bar{w}_i^T A \bar{w}_{i-1}] = \frac{1}{2} E\left[\bar{w}_{i-1}^T (I - \mu P_1) A \bar{w}_{i-1}\right]$$

$$+ \frac{1}{2} E\left[\bar{w}_{i-2}^T A (I - \mu P_1) \bar{w}_{i-2}\right].$$

(7)

when neglecting the dependency of the measurement noise $v(i)$ on the past coefficient error vector $\bar{w}_{i-1}$. In steady-state, we can assume that

$$E[\bar{w}_i^T A \bar{w}_{i-1}] \approx E\left[\bar{w}_{i-1}^T A \bar{w}_{i-2}\right],$$

(8)

so that

$$E[\bar{w}_i^T A (I - \mu P_1) \bar{w}_{i-2}] = E\left[\bar{w}_{i-1}^T (I - \mu P_1) A \bar{w}_{i-1}\right].$$

(9)

By the Assumption II, (9) can be rewritten as

$$E[\bar{w}_i^T A (I - \mu E[P_1]) \bar{w}_{i-2}] = E\left[\bar{w}_{i-1}^T (I - \mu E[P_1]) A \bar{w}_{i-1}\right].$$

(10)

To obtain $E[\bar{w}_i^T \Sigma' \bar{w}_{i-2}]$, we select $A$ as a solution to the linear system of equations $A(I + \mu E[P_1]) = \Sigma'$ as follows:

$$A = \Sigma' (I + \mu E[P_1])^{-1}.$$  

(11)

Consequently, we obtain the following relation

$$E[\bar{w}_i^T \Sigma' \bar{w}_{i-2}] = E\left[\bar{w}_{i-1}^T (I - \mu E[P_1]) \Sigma' (I + \mu E[P_1])^{-1} \bar{w}_{i-1}\right]$$

$$= E[\|\bar{w}_{i-1}\|^2],$$

(12)

where $\sigma_1 = \text{vec}\{(I - \mu E[P_1]) \Sigma' (I + \mu E[P_1])^{-1}\}$ and $\Sigma_1 = \text{vec}(\sigma_1)$. To represent $\sigma_1$ as a function of $\sigma$, we use the following property: For any matrices $\{P, Q, R\}$ of compatible dimensions, it holds that

$$\text{vec}(PQR) = (R^T \otimes P) \text{vec}(Q);$$

(13)

where $A \otimes B$ means the Kronecker product of two matrices $A$ and $B$. Thus, $\sigma_1$ can be rewritten by

$$\sigma_1 = \{(I + \mu E[P_1])^{-1} \otimes (I - \mu E[P_1])\} F \sigma = F_2 F \sigma$$

(14)

where $F_2 \triangleq (I + \mu E[P_1])^{-1} \otimes (I - \mu E[P_1])$. Substituting (12) into (5), in steady-state, (4) is rewritten as

$$E[\|\bar{w}_{\infty}\|^2] = E\left[\frac{1}{2} E\left[\|\bar{w}_{\infty}\|^2\right] + \frac{1}{2} E\left[\|\bar{w}_{\infty}\|^2\right] + \mu^2 \sigma_v^2(\gamma^T \gamma)\right],$$

(15)

which is equivalent to

$$E\left[\|\bar{w}_{\infty}\|^2 (I - \frac{1}{2} F - \frac{1}{2} F_2 F)\right] = \mu^2 \sigma_v^2(\gamma^T \gamma).$$

(16)

To calculate MSD, we choose $\sigma$ to reduce the weight into the identity matrix. Thus, it needs to be selected to satisfy the following condition such that $(I - \frac{1}{2} F - \frac{1}{2} F_2 F) \sigma = \text{vec}(I)$, then

$$\sigma = \left(I - \frac{1}{2} F - \frac{1}{2} F_2 F\right)^{-1} \text{vec}(I).$$

(17)

Consequently, the left-hand side of (16) becomes the MSD for $n = 2$, and (16) leads to

$$\text{MSD} = \mu^2 \sigma_v^2(\gamma^T \gamma) \left(I - \frac{1}{2} F - \frac{1}{2} F_2 F\right)^{-1} \text{vec}(I), \text{ (for } n = 2).$$

(18)

In a similar way, the MSD for $n = 3$ is given by

$$\text{MSD} = \mu^2 \sigma_v^2(\gamma^T \gamma) \left(I - \frac{1}{3} F - \frac{2}{27} F_3 F\right)^{-1} \text{vec}(I), \text{ (for } n = 3).$$

(19)
where \( F_3 \) is obtained by
\[
F_3 = \left\{ \begin{bmatrix} 2I & I \end{bmatrix} \left\{ \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \right\}^{-1} \begin{bmatrix} I \\ I \end{bmatrix}^T \left( \frac{1}{3} \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 0 \end{bmatrix} \otimes (I - \mu E[P_i]) \right) \right\} \otimes (I - \mu E[P_i]).
\] (20)

If we exchange \( \text{vec}\{I\} \) with \( \text{vec}\{R_u\} \) in (18) and (19), the results become the excess mean square error (EMSE). We are working to find the steady-state analysis for a general \( n \), so it will be our future work.

### 2.3. Proposed NLMS with Variable Reusing Order of Coefficient Vector

As can be seen, the reusing order \( n \) greatly affects the convergence performance. A larger reusing order yields a lower steady-state MSD, but reduces the convergence speed. To address the tradeoff between the convergence rate and the steady-state MSD, we propose an automatic adjusting method of reusing order at every iteration so as to achieve both fast convergence rate and small steady-state MSD. At the initial stage, a small reusing order is used to guarantee fast convergence. Subsequently, at the steady-state, a large reusing order is used, which results in a low steady-state MSD.

We employ an evolutionary approach which automatically adjusts the reusing order according to the status of the adaptive filter proposed by [8]. This can be obtained by
\[
n_i = \begin{cases} \min\{n_{i-1} + 1, n_{\max}\} & \text{if } e^2(i) < \eta \\ \max\{n_{i-1} - 1, 1\} & \text{otherwise} \end{cases}
\] (21)

where \( e(i) = d(i) - u_i w_{i-1} \), \( n_i \) is the reusing order at iteration \( i \), \( \eta \) denotes the threshold, and \( n_{\max} \) is the maximum reusing order. It is to be noted that \( n_i \) is updated so as to satisfy \( 1 \leq n_i \leq n_{\max} \). In determining the threshold, we consider the theoretical steady-state MSE of the NLMS presented in [7], which is given by
\[
\text{MSE} = \frac{\sigma_n^2}{2 - \mu} + \sigma_v^2 = \frac{2\sigma_n^2}{2 - \mu} \triangleq \xi_{\text{NLMS}}.
\] (22)

If \( e^2(i) < \xi_{\text{NLMS}} \), then \( n_i \) should be increased by one from \( n_{i-1} \) for smaller steady-state MSD. Therefore, we set the threshold in (21), \( \eta \), as follows:
\[
\eta = \xi_{\text{NLMS}} = \frac{2\sigma_n^2}{2 - \mu}.
\] (23)

With this proposed variable reusing order, the update equation of the proposed algorithm becomes
\[
w_i = \frac{1}{n_i} \sum_{k=1}^{n_i} w_{i-k} + \mu \frac{u_i^T}{\|u_i\|^2} d(i) - u_i \frac{1}{n_i} \sum_{k=1}^{n_i} w_{i-k}.
\] (24)

When the output error is increased by variation of the environment, the reusing order decreases to track such variation quickly. Thus, the proposed algorithm is expected to have a fast convergence speed. When the proposed algorithm converges on the threshold, that threshold acts as a switching point to increase the reusing order of the proposed algorithm. As a result, the steady-state MSD is expected to decrease further.

### 3. EXPERIMENTAL RESULTS

We demonstrate the performance of the proposed algorithm by carrying out experiments in the system identification configuration. The unknown system to be identified is randomly generated with 16 taps. The adaptive filter is designed to have the same length as the unknown system, \( M = 16 \). The input signal \( u(i) \) is obtained by filtering a white, zero-mean, Gaussian random sequence through a first-order autoregressive (AR) system \( G(z) = 1/(1 - 0.9z^{-1}) \). The signal-to-noise ratio (SNR) is calculated by SNR = 10log_{10}(E[y^2(i)]/E[v^2(i)])], where \( y(i) = u_i w_o \). The measurement noise \( v(i) \) is added to \( y(i) \) such that SNR = 30dB. The MSD, \( E[\|w_o - w_i\|^2] \), is evaluated by ensemble averaging over 200 independent trials. Also, we assumed that the noise variance, \( \sigma_v^2 \), is known \textit{a priori} because it can be easily estimated during silences and on-line in many practical applications [9]–[10].
Fig. 3. MSD curves of the conventional NLMS-RC \((n = 1, 2, 3, 4, 5)\) and the proposed NLMS-VRC \((n_{\text{max}} = 6)\) \([M = 16, \text{Input: Gaussian AR(1), pole at 0.9}]\).

Fig. 1 shows the steady-state MSD according to \(n\) when the SNR varies from 5dB to 30dB and the step-size is set to 1.0. From Fig. 1, we can see that the steady-state MSD gets small regardless of SNR as \(n\) increases. Fig. 2 shows the steady-state MSD curves of the NLMS-RC for colored Gaussian input as a function of the step-size. The step-size varies from 0.04 to 1.0 and the SNR is set to 30dB. The theoretical results are calculated using (18) and (19), and the simulation results are obtained by averaging more than 1000 instantaneous values in the steady-state and then averaging 200 independent trials. The simulation results present good agreement with the theoretical results.

Fig. 3 shows the MSD curves for the proposed NLMS-VRC \((n_{\text{max}} = 6)\) and the conventional NLMS-RC with various reusing orders \((n = 1, 2, 3, 4, 5)\). The SNR and the step-size are set to SNR = 30dB and \(\mu = 1.0\), respectively. Further, to examine the tracking performance, the unknown system is changed abruptly from \(w^0\) to \(-w^0\) at 2500 iterations. As can be seen, the NLMS-RC has a trade-off between fast convergence speed and small steady-state MSD depending on the reusing order \(n\), and the proposed NLMS-VRC results in both fast convergence and small steady-state MSD. Fig 4(a) shows the reusing order over 1 trial for the proposed NLMS-VRC. The reusing order dynamically varies for decreasing steady-state MSD and tracking the change of the unknown system. The average reusing order shows the varying tendency quite well, as represented in Fig. 4(b).

4. CONCLUSIONS

In this paper, we analyzed the steady-state performance of the NLMS-RC using energy conservation arguments and proposed a novel NLMS-VRC algorithm which can vary reusing order to achieve both fast convergence speed and small steady-state MSD. The experimental results confirmed that there is good match between the analysis and practice. The proposed algorithm has been shown to have better convergence performance compared to the conventional NLMS in terms of convergence speed and steady-state MSD. Due to space constraints, the procedure of analysis has not been fully described. The detailed analysis for a general \(n\) will be considered in an upcoming paper.

5. REFERENCES


