PAPER

A Representation Method of the Convergence Characteristic of the LMS Algorithm Using Tap-Input Vectors

Kiyoshi NISHIKAWA† *and* Hitoshi KIYA†, *Members*

SUMMARY The main purpose of this paper is to give a new representation method of the convergence characteristics of the LMS algorithm using tap-input vectors. The described representation method is an extended version of the interpretation method based on the orthogonal projection. Using this new representation, we can express the convergence characteristics in terms of tap-input vectors instead of the eigenvalues of the input signal. From this representation, we consider a general method for improving the convergence speed.

key words: *adaptive signal processing, system identification, orthogonal projections, oblique expansion*

1. Introduction

This paper provides a new representation method of the convergence characteristic of the least mean square (LMS) algorithm. Ordinary, the convergence characteristic of the LMS algorithm is represented in terms of the eigenvalues of the input signal [1]. In contrast to this tradition, the proposed method uses tap-input vectors instead of the eigenvalues.

The eigenvalue representation is based on the Wiener filter theory, which was originally developed for the statistical optimum system theory based on the least mean square criterion [1], [2]. The most important result derived from this representation might be the fact that the convergence speed of the LMS algorithm is determined by the spread of the eigenvalues of the autocorrelation matrix of the input signal [1], [3]. That is, when the eigenvalues are widely spread then the convergence speed becomes slow. Most of the analyses of the LMS algorithm are based on this representation [4], [5] and many attempts are proposed for improving the slow convergence speed [6].

In practical applications, the adaptive algorithms are applied to a particular realization of stochastic processes. Adaptive algorithms, therefore, are required to converge with one realization in the deterministic sense as well as the mean square sense, i.e., w.p.1 (with probability one) or a.s. (almost sure) convergence [7], [8]. The previous analysis methods based on the stochastic quantities, i.e., eigenvalues, are not useful for describing the deterministic convergence characteristic. Because the convergence in mean square does not assure the w.p.1 convergence. See [7] for more detailed summary of the previous analyses and references.

So far, several papers pointed out this difficulty [7], [8] and thus concerned to construct an alternative analysis methods that enable us to analysis the w.p.1 convergence. One of such analysis methods is the one based on a geometrical interpretation [9], [10]. It uses the concept of the orthogonal projection for analyzing the performance of the algorithm at each time instant n. However, we cannot use this method for considering the global convergence characteristics of the LMS algorithm because it only uses the information at time n. In other words, this method represents only the instantaneous behavior of the adaptive filter at n.

To avoid this difficulty of the conventional method, we propose a new representation method for the LMS algorithm. In order to describe the global characteristics as well as the instantaneous one at n, we introduce a new concept, i.e., the oblique-expansion modeling of the unknown system [11]. By combining the orthogonal projection with this concept, the global characteristics can be represented in terms of the tap-input vectors. Therefore, the proposed method can be regarded as an extended version of the orthogonal projection representation. The reason of the slow convergence of the LMS algorithm can be expressed and a strategy for improving it is considered. For showing the validity of this strategy, we derive a gradient type algorithm that has a fast convergence speed.

The remainder of this paper is organized as follows. In Sect. 2, a review of the LMS and the NLMS algorithms as a preliminary is given. The description of the proposed representation method is presented in Sect. 3. This section includes the introduction of the oblique expansion modeling of the optimum system. In Sects. 4 and 5, derivation of a fast algorithm is considered based on the proposed representation and some computer simulation results are given for showing the efficiency of the derived algorithm. The conclusion is given in Sect. 6

2. Preliminary

This section provides a preliminary for the following description. First, we briefly describe the LMS algo-

Manuscript received March 25, 1995.
†The authors are with the Faculty of Engineering, Tokyo Metropolitan University, Hachioji-shi, 192–03 Japan.

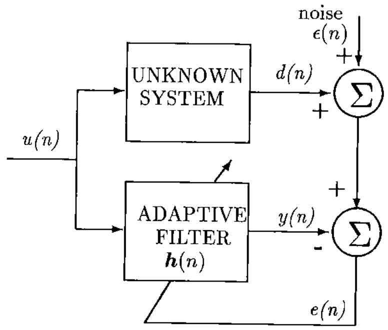


Fig. 1 System identification model.

rithm. Then the interpretation based on the orthogonal projection is reviewed.

In Fig. 1, a model of the system identification problem that is used in the following consideration is shown. Note that only the system identification problems using the LMS algorithm will be considered and we assume signals and filter coefficients are real. We define the following notations for simplification:

$n \triangleq$ Time index

$u(n) \triangleq n$th input signal

$\boldsymbol{u}(n) \triangleq [u(n), u(n-1), \cdots, u(n-P+1)]^T$: tap-input vector at n

$d(n) \triangleq n$th desired signal

$P \triangleq$ Number of taps of the adaptive FIR filter

$h_l \triangleq l$ th coefficient of the adaptive filter ($l = 1, 2, \cdots, P$)

$\boldsymbol{h}(n) \triangleq [h_1(n), h_2(n), \cdots, h_P(n)]^T$: Adaptive filter coefficients vector

$\epsilon(n) \triangleq$ Noise component

$\boldsymbol{R} \triangleq \mathrm{E}[\boldsymbol{u}(n)\boldsymbol{u}^T(n)]$: Auto correlation matrix of input signal

$\boldsymbol{p} \triangleq \mathrm{E}[d(n)\boldsymbol{u}(n)]$: Cross correlation vector between input signal and desired signal

where the superscript T and $\mathrm{E}[\cdot]$ denote the transpose of a matrix and the expectation operation respectively.

2.1 The LMS and NLMS Algorithms

The LMS algorithm is a gradient-type algorithm for searching the optimum filter [1]. It updates the filter $\boldsymbol{h}(n)$ adaptively using the instantaneous gradient vector [12] which is defined as

$$\tilde{\nabla}(n) = e(n)\boldsymbol{u}(n) \tag{1}$$

where $e(n)$ is the error signal:

$$\begin{aligned} e(n) &= \epsilon(n) + d(n) - y(n) \\ &= \epsilon(n) + d(n) - \boldsymbol{h}^T(n)\boldsymbol{u}(n). \end{aligned} \tag{2}$$

The formula for updating $\boldsymbol{h}(n)$ is given by

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu\tilde{\nabla}(n) \tag{3}$$

where μ is the step size parameter, which perfectly controls the performance of the algorithm. The value of μ must be set for ensuring the w.p.1 convergence of the algorithm. However, the range of the value depends on the ratio of the largest and smallest eigenvalues [1]. In practical applications, these statistical information is rarely obtained, therefore, we must set μ as small value.

To avoid this problem, the normalized LMS (NLMS) algorithm [10] is preferred in practical applications. In the NLMS algorithm, the definition of the step size parameter is modified as

$$\mu = \frac{\alpha}{||\boldsymbol{u}(n)||^2 + \delta} \tag{4}$$

where $||\cdot||^2$ shows the square-norm of a vector. In this equation, α is a parameter for controlling the convergence speed of the algorithm. It is well known that $\alpha = 1$ provides the best convergence performance [13] so that we fix α as $\alpha = 1$ in the following discussion. δ is another parameter that protects the equation from divergence in case of $||\boldsymbol{u}(n)||^2 = 0$. However, we choose $\delta = 0$ for the simplicity of descriptions. Under these simplifications, (3) becomes the next form:

$$\begin{aligned} \boldsymbol{h}(n+1) &= \boldsymbol{h}(n) + \frac{1}{||\boldsymbol{u}(n)||^2}\tilde{\nabla}(n). \\ &= \boldsymbol{h}(n) + \frac{d'(n) - \boldsymbol{u}^T(n)\boldsymbol{h}(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n) \end{aligned} \tag{5}$$

where $d'(n) = \epsilon(n) + d(n)$. Note that if we set

$$\alpha = \mu||\boldsymbol{u}(n)||^2 \tag{6}$$

we can regard the LMS algorithms as a special version of the NLMS, and therefore, we can concentrate on the NLMS algorithm without any loss of generality. Note that it is well known that there are many variations of the LMS algorithm as well as the NLMS algorithm [3], [14].

2.2 The Orthogonal Projection Representation

For later convenience, we describe the orthogonal projection representation of the LMS algorithm [9], [12] here. By rewriting (5), we have

$$\boldsymbol{h}(n+1) = \left[\mathrm{I} - \frac{\boldsymbol{u}(n)\boldsymbol{u}^T(n)}{||\boldsymbol{u}(n)||^2}\right]\boldsymbol{h}(n) + \frac{d'(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n). \tag{7}$$

Let us show that we can interpret this equation in terms of the orthogonal projection.

We start with the description of the first term in the right-hand side using the concept of the orthogonal

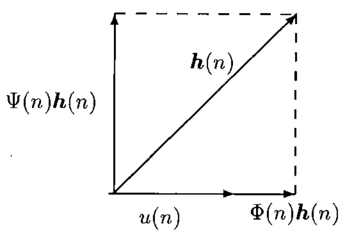


Fig. 2 The orthogonal projection for the two-dimensional case.

projection. In order to explain, we define the next two operators [15]:

$$\Phi(n) \stackrel{\text{def}}{=} \frac{\boldsymbol{u}(n)\boldsymbol{u}^T(n)}{||\boldsymbol{u}(n)||^2} \tag{8}$$

$$\Psi(n) \stackrel{\text{def}}{=} \mathrm{I} - \Phi(n). \tag{9}$$

These two operators are used to extract some particular components from an arbitrary vector, e.g. $\boldsymbol{h}(n)$. In Fig. 2, an example of the two-dimensional case is shown. By applying $\Phi(n)$ to $\boldsymbol{h}(n)$ from the left side, we can extract the component of $\boldsymbol{h}(n)$ which is parallel to $\boldsymbol{u}(n)$. On the other hand, by applying $\Psi(n)$ we can obtain $[\mathrm{I} - \Phi(n)]\boldsymbol{h}(k) = \boldsymbol{h}(k) - \Phi(n)\boldsymbol{h}(n)$.

Using these definitions we can express (7) as:

$$\boldsymbol{h}(n+1) = \Psi(n)\boldsymbol{h}(n) + \frac{d'(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n). \tag{10}$$

From this equation we can interpret the updating process of the LMS algorithm as the next two steps:

1. By applying $\Psi(n)$, the component of $\boldsymbol{h}(n)$ which is parallel to $\boldsymbol{u}(n)$ is subtracted from $\boldsymbol{h}(n)$.
2. Then, $\boldsymbol{u}(n)$ scaled by $\frac{d'(n)}{||\boldsymbol{u}(n)||^2}$ is added to $\boldsymbol{h}(n)$ (the second term in (10)).

This interpretation of the LMS algorithm is useful for considering the w.p.1 convergence [9]. However, obviously from the above description, this representation uses only the instantaneous information at n. So this makes it difficult to describe the global convergence characteristic of the algorithm. Further, this representation does not contain the explanation of the slow convergence under wide spread eigenvalues.

3. Description of a Proposed Representation

In this section, we describe the proposed representation method. We begin describing the definition of the optimum system for making the following consideration clearer. Then we introduce the concept of the oblique-expansion modeling of the optimum system that plays a central roll in the proposed representation. Using this model, we show that the global convergence characteristic can be described in terms of the tap-input vectors.

3.1 Optimum System

As mentioned before, this paper concerns the system identification problems and its purpose is to identify the system that produces $d(n)$ in response to $\boldsymbol{u}(n)$.

When we apply the LMS algorithm, the optimum system that the algorithm searches is said to be the Wiener filter. The Wiener filter is the filter that minimizes the mean squared error (MSE), which is defined as

$$J(n) = \mathrm{E}[e^2(n)]. \tag{11}$$

It is well known that in order to minimize (11), the filter must satisfy the next Wiener-Hopf equation:

$$\boldsymbol{h}_o = \boldsymbol{R}^{-1}\boldsymbol{p} \tag{12}$$

and this optimum filter $\boldsymbol{h}_o$ is called the Wiener filter [1]. However, this definition of the optimum filter uses the stochastic quantities so that it pauses a difficulty when the w.p.1 convergence is mentioned.

To avoid this difficulty, let us consider the deterministic optimum system according the description in [14]. We introduce the next assumptions:

1. $\{\boldsymbol{u}(n)\}$ is a zero mean Gaussian white process with the autocorrelation matrix

$$\boldsymbol{R} = \mathrm{E}[\boldsymbol{u}(n)\boldsymbol{u}^T(n)]. \tag{13}$$

2. There is a true parameter $\hat{\boldsymbol{h}}_o$ such that the error signal

$$\epsilon(n) = d(n) - y(n) = d(n) - \boldsymbol{u}^T(n)\hat{\boldsymbol{h}}_o \tag{14}$$

is a zero mean Gaussian white noise of variance σ_ϵ^2.

3. $\{\epsilon(n)\}$ and $\{\boldsymbol{u}(n)\}$ are statistically independent.

From these assumptions and the definition of $\boldsymbol{p}$ we have

$$\begin{aligned}\boldsymbol{p} &= \mathrm{E}[u(n)d(n)] \\ &= \mathrm{E}[u(n)\epsilon(n)] + \mathrm{E}[\boldsymbol{u}(n)\boldsymbol{u}^T(n)]\hat{\boldsymbol{h}}_o \\ &= \boldsymbol{R}\hat{\boldsymbol{h}}_o.\end{aligned} \tag{15}$$

Thus we have seen that $\hat{\boldsymbol{h}}_o$ satisfies the Wiener-Hopf equation. In the following consideration, our objective is to search $\hat{\boldsymbol{h}}_o$ using the LMS algorithm. Note that, in the following, we will assume that $\hat{\boldsymbol{h}}_o$ is time-invariant and that $\epsilon(n)$ is zero for simplicity of notations. We also assume that the order of $\boldsymbol{h}(n)$ is high enough to represent $\hat{\boldsymbol{h}}_o$.

3.2 Oblique Expansion Representation of the Optimum System

For describing the proposed representation method, we introduce a new concept here: modeling of $\hat{\boldsymbol{h}}_o$ as a linear combination of the tap-input vectors. For later convenience, we define the set $\boldsymbol{U}$ as

$$\boldsymbol{U} = \{\boldsymbol{u}(n), \boldsymbol{u}(n+1), \cdots, \boldsymbol{u}(n+N-1)\} \tag{16}$$

where N is an integer and its meaning is considered later.

Let us consider identifying $\hat{\boldsymbol{h}}_o$ using the NLMS algorithm. We obtain the next equation by rewriting (5):

$$\begin{aligned}\boldsymbol{h}(n+1) &= \boldsymbol{h}(n) + \frac{e(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n) \\ &= \boldsymbol{h}(n) + g(n)\boldsymbol{u}(n)\end{aligned} \tag{17}$$

where $g(n)$ is set as

$$g(n) = \frac{e(n)}{||\boldsymbol{u}(n)||^2}. \tag{18}$$

Obviously, this procedure has a feedback structure so that we can express $\boldsymbol{h}(n+M)$ using $\boldsymbol{h}(n)$ as

$$\boldsymbol{h}(n+M) = \boldsymbol{h}(n) + \sum_{i=0}^{M-1} g(n+i)\boldsymbol{u}(n+i) \tag{19}$$

where M is an arbitrary integer. We require that $\boldsymbol{h}(n+M)$ approaches $\hat{\boldsymbol{h}}_o$ as M increases [10] with probability one:

$$\begin{aligned}\lim_{M\to\infty}\boldsymbol{h}(n+M) &= \lim_{M\to\infty}\{\boldsymbol{h}(n)+\sum_{i=0}^{M-1} g(n+i)\boldsymbol{u}(n+i)\} \\ &= \hat{\boldsymbol{h}}_o\end{aligned} \tag{20}$$

This suggests that $\hat{\boldsymbol{h}}_o$ can be expressed as a linear combination of the tap-input vectors $\{\boldsymbol{u}(n)\}$.

Motivated from this fact, we propose to model $\hat{\boldsymbol{h}}_o$ as the next form [11]:

$$\hat{\boldsymbol{h}}_o = \boldsymbol{h}(n) + \sum_{i=0}^{N-1} c_i\boldsymbol{u}(n+i) \tag{21}$$

where N is the integer for which the set $\boldsymbol{U}$ has rank P, the number of coefficients of $\boldsymbol{h}(n)$. In this equation c_i is a weighting coefficient corresponding to $\boldsymbol{u}(n+i)$. We can say that under the model (21), the system identification problem turns to be a problem of searching the set $\{c_i\}$ that satisfies (21).

Note that if the set $\boldsymbol{U}$ does not have rank P then we cannot uniquely identify the unknown system. In the following, we will assume that $\boldsymbol{U}$ has rank P. Also note that the set $\boldsymbol{U}$ might not be an orthogonal one even if the input signal is a white Gaussian process. Therefore, in practical applications, (21) corresponds to the oblique expansion of $\hat{\boldsymbol{h}}_o$ mostly. In Fig. 3, an example of this expansion is shown when $P=3$.

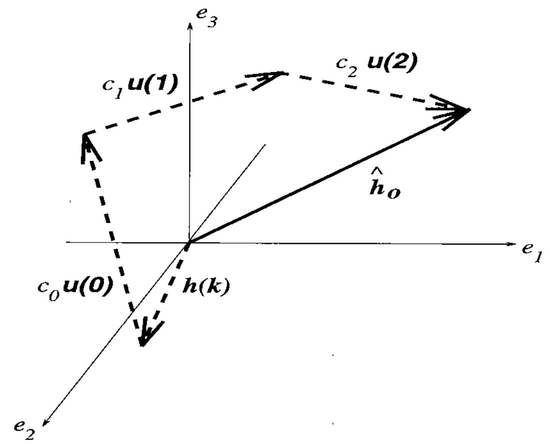


Fig. 3 A representation of the optimum system $\hat{h}_o$ as a linear combination of tap-input vectors.

3.3 The Convergence Characteristic

Let us consider representing the convergence characteristics of the NLMS algorithm using the model (21). For that purpose, let us calculate $\boldsymbol{h}(n+1)$ by updating $\boldsymbol{h}(n)$ using $\boldsymbol{u}(n)$. Using (5) and (21), $\boldsymbol{h}(n+1)$ is calculated as

$$\begin{aligned}\boldsymbol{h}(n+1) &= \boldsymbol{h}(n) + \frac{\hat{\boldsymbol{h}}_o^T\boldsymbol{u}(n) - \boldsymbol{h}^T(n)\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n) \\ &= \boldsymbol{h}(n) + \frac{\sum_{i=0}^{N-1} c_i\boldsymbol{u}^T(n+i)\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}\boldsymbol{u}(n) \\ &= \boldsymbol{h}(n) + \{c_0 + c_0'\}\,\boldsymbol{u}(n)\end{aligned} \tag{22}$$

where c_0' is given as

$$c_0' = \sum_{i=1}^{N-1} c_i\frac{||\boldsymbol{u}(n+i)||}{||\boldsymbol{u}(n)||}\cos\theta_{ni} \tag{23}$$

here $\cos\theta_{ni}$ is defined as

$$\cos\theta_{ni} \stackrel{\mathrm{def}}{=} \boldsymbol{u}^T(n)\boldsymbol{u}(n+i)/\{||\boldsymbol{u}(n)||\,||\boldsymbol{u}(n+i)||\}. \tag{24}$$

Let us consider the result (22). Under the definition (21), it is obvious that the coefficient corresponding to $\boldsymbol{u}(n)$ is c_0. Therefore, c_0' in (22) seems to be an extra term corresponding to $\boldsymbol{u}(n)$. We can conclude that the slow convergence of the LMS algorithm is due to the existence of this extra term c_0'.

Known from (23) and (24), the attribution of the extra term becomes bigger as $\cos\theta_{ni}$ getting bigger so that the convergence speed of the LMS algorithm becomes slower. We can say that this cosine term is a deterministic representation of the eigenvalue in the Wiener theory.

3.4 Discussion

We can say that the combination of this oblique expansion representation with the orthogonal projection provides an alternative image about the convergence characteristics of the algorithm. The LMS algorithm updates the filter coefficients $\boldsymbol{h}(n)$ by using the orthogonal projection at each iteration and only the information of $\boldsymbol{u}(n)$ is used then. This process is very significant if $\boldsymbol{U}$ constructs an orthogonal set and then, the algorithm identifies the optimum system $\hat{\boldsymbol{h}}_o$ with N iterations. However, in practical applications, as mentioned in Sect. 3.2, the set $\boldsymbol{U}$ probably is not an orthogonal one so that the term added to update $\boldsymbol{h}(n)$ becomes greater than the correct value as expressed in (22). We can say that the slow convergence speed of the LMS algorithm is caused by the fact that it uses the orthogonal projection to update the filter $\boldsymbol{h}(n)$ although the set $\boldsymbol{U}$ is not orthogonal. Therefore, the convergence speed can be improved if we decrease the effect of the extra term c_i' in (22).

4. A New Fast Gradient Type Algorithm

We have seen that the convergence speed of the LMS algorithm can be improved if we can cancel the extra term that is produced due to its update formula. In this section, a gradient type algorithm is derived which realizes faster convergence speed by attempting to cancel the extra term.

4.1 Derivation

As shown in Sect. 3.2, we can improve the convergence speed of the LMS algorithms if we can cancel the extra term c_i' in (22) which is generated by the result of the form of the update formula of the algorithm. Here, a new LMS-type algorithm with fast convergence speed is derived based on this strategy.

Here, let us consider using the time-averaged (TA) correlations instead of instantaneous ones. However, the TA correlations have only the averaging effect and cannot improve the convergence speed [16]. Then, we propose to use the next modified TA correlations:

$$\tilde{\boldsymbol{\Phi}}(n) = \Psi(n)\tilde{\boldsymbol{\Phi}}(n-1) + \Phi(n) \tag{25}$$

$$\tilde{\varphi}(n) = \Psi(n)\tilde{\varphi}(n-1) + \phi(n) \tag{26}$$

where $\Phi(n)$ and $\phi(n)$ are respectively given by

$$\Phi(n) = \frac{\boldsymbol{u}(n)\boldsymbol{u}^T(n)}{||\boldsymbol{u}(n)||^2}$$

$$\phi(n) = \frac{d(n)\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}. \tag{27}$$

We defined the initial conditions for these recursions as

$$\tilde{\boldsymbol{\Phi}}(-1) = \mathbf{o} \tag{28}$$

$$\tilde{\varphi}(-1) = \mathbf{o} \tag{29}$$

where $\mathbf{o}$ shows a zero matrix.

Using these modified correlations, the filter coefficients are updated according to

$$\begin{aligned}\boldsymbol{h}(n+1) &= \left[\mathrm{I} - \tilde{\boldsymbol{\Phi}}(n)\right]\boldsymbol{h}(n) + \tilde{\varphi}(n)\\ &= \left[\mathrm{I} - \Phi(n)\right]\boldsymbol{h}(n) + \frac{d(n)\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}\\ &\quad -\Psi(n)\left[\tilde{\boldsymbol{\Phi}}(n-1)\boldsymbol{h}(n) - \tilde{\varphi}(n-1)\right]\end{aligned} \tag{30}$$

instead of (7).

Let us show that these formulas ensure that the effects of extra term c_i' are canceled. The right-hand side of (30) can be divided into two parts, the first two terms are identical to the formula of the NLMS and the third term is the unique term in this algorithm. Let us consider the roll of the third.

For that purpose, we first rewrite (26) as the form

$$\begin{aligned}\phi(n) &= \frac{d(n)\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}\\ &= \frac{\boldsymbol{u}^T(n)\hat{\boldsymbol{h}}_o\boldsymbol{u}(n)}{||\boldsymbol{u}(n)||^2}\\ &= \Phi(n)\hat{\boldsymbol{h}}_o.\end{aligned} \tag{31}$$

From this, $\tilde{\varphi}(n)$ can be expressed as

$$\begin{aligned}\tilde{\varphi}(n) &= \{\Psi(n)\tilde{\boldsymbol{\Phi}}(n-1) + \Phi(n)\}\hat{\boldsymbol{h}}_o\\ &= \tilde{\boldsymbol{\Phi}}(n)\hat{\boldsymbol{h}}_o.\end{aligned} \tag{32}$$

Using this equation, the third term in the right-hand side in (30) is rewritten as

$$\begin{aligned}&\Psi(n)\{\tilde{\boldsymbol{\Phi}}(n-1)\boldsymbol{h}(n) - \tilde{\varphi}(n-1)\}\\ &\quad= \Psi(n)\tilde{\boldsymbol{\Phi}}(n-1)\{\boldsymbol{h}(n) - \hat{\boldsymbol{h}}_o\}\end{aligned} \tag{33}$$

where $\Psi(n)\tilde{\boldsymbol{\Phi}}(n-1)$ can be regarded as an orthogonal projection operator. By applying $\Psi(n)$ from the left side, we can extract the component $\boldsymbol{h}_e$ of $\{\boldsymbol{h}(n) - \hat{\boldsymbol{h}}_o\}$ in the subspace spanned by $\{\boldsymbol{u}(0), \boldsymbol{u}(1), \ldots, \boldsymbol{u}(n-1)\}$. Then $\boldsymbol{u}(n)$-direction component is extracted from $\boldsymbol{h}_e$ by applying $\Psi(n)$ from the left side. This extracted components corresponds to the extra components in (22) so that this algorithm will converge faster than the ordinary NLMS algorithm.

In the next section, we show some computer simulation results to demonstrate both of the validity of the discussion above and the effectiveness of the algorithm.

5. Simulation Results

Here, we show some results of computer simulations to see the convergence property of the new algorithm described in the previous section.

We compared the NLMS algorithm, the RLS (recursive least squares) algorithm, and the proposed algorithm where $\alpha = 1$ for the NLMS algorithm and the exponential weighting factor of the RLS is $\lambda = 0.95$.

The optimum system $\hat{h}_o$ is a lowpass FIR filter with 20th order. The order of the adaptive filter $h(n)$ is the same as that of $\hat{h}_o$. The input signal is an AR(1) process of zero-mean with the AR coefficients a_1. We did simulations with varying the value of a_1. A white Gaussian noise is added to the desired signal and the signal to noise ratio (SNR) is SNR$= -60$ dB. Results are ensemble average of 50 independent processes.

In Fig. 4 and Fig. 5, the results of when $a_1 = 0.0$ and that of $a_1 = 0.9$ are shown respectively. From these figures, we see that the convergence property of the proposed algorithm is not affected by the correlation of the input signal. That is, even if the input signal is a theoretically white process, the new algorithm converges faster than the NLMS algorithm. Under the Wiener filter theory, the LMS algorithm provides the fastest convergence speed under the white input condition. However, as shown in the previous sections, the convergence speed can be improved unless the set U is orthogonal. The result suggests the validity of the proposed analysis method described in this paper.

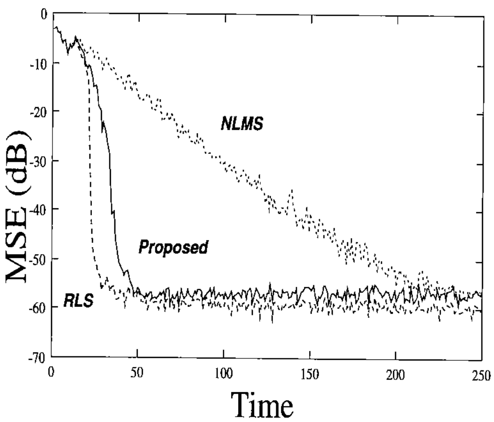


Fig. 4 Results of the computer simulation. The input signal is an AR(1) process and the value of the AR coefficient is 0.0.

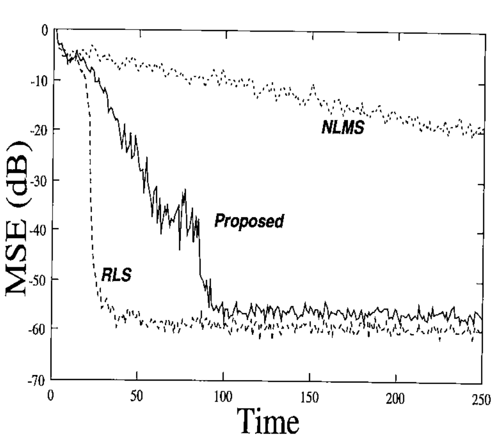


Fig. 5 Results of the computer simulation. The input signal is an AR(1) process and the value of the AR coefficient is 0.9.

6. Conclusion

In this paper, we proposed a new representation method for the convergence characteristics of the LMS algorithm. The proposed method is an extended version of the method based on the orthogonal projection. We used a new concept, i.e., the modeling of the optimum system using oblique expansion, for developing the proposed method. We showed that the reason of the slow convergence speed of the LMS algorithm is expressed in terms of $u(n)$. From this, we considered a strategy for improving the convergence speed. To show the validity of the proposed method, we derived a new gradient type algorithm based on the proposed method and the results of computer simulations were shown. From the results, we showed that this new algorithm improves the convergence speed despite the correlation of the input signal.

References

[1] B. Widrow and S.D. Stearns, "Adaptive Signal Processing," Englewood Cliffs, NJ 07632: Prentice Hall, 1985.

[2] J.S. Arora, "Introduction to Optimum Design," New York, NY: McGraw-Hill, 1989.

[3] W.A. Sethares, "The least mean square family," in Adaptive System Identification and Signal Processing Algorithms, ch. 3, Englewood Cliffs, NJ 07632: Prentice Hall, 1993.

[4] N.J. Bershad and O.M. Macci, "Adaptive recovery of a chirped sinusoid in noise, part 2: Performance of the LMS algorithm," IEEE Trans. on Signal Processing, vol.39, pp.595–602, March 1991.

[5] W.A. Gardner, "Nonstationary learning characteristics of the LMS algorithm," IEEE Trans. on Circuits and Systems, vol.34, pp.1199–1207, Oct. 1987.

[6] R.H. Kwong and E.W. Johnston, "A variable step size LMS algorithm," IEEE Trans. on Signal Processing, vol.40, pp.1633–1642, July 1992.

[7] V. Solo, "The limiting behavior of LMS," IEEE Trans. Accoust., Speech, and Signal Processing, vol.37, pp.1909–1922, Dec. 1989.

[8] V. Solo, "The error variance of LMS with time-varying weights," IEEE Trans. on Signal Processing, vol.40, pp.803–813, April 1992.

[9] D.T.M. Slock, "On the convergence speed of the LMS and the normalized LMS algorithms," IEEE Trans. on Signal Processing, vol.41, pp.2811–2825, Sept. 1993.

[10] J. Nagumo and A. Noda, "A learning method for system identification," IEEE Trans. on Automatic Control, vol.12, pp.282–287, June 1967.

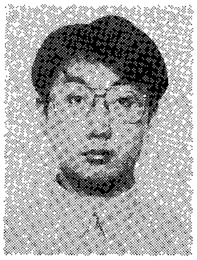
[11] K. Nishikawa and H. Kiya, "A general method for improving the convergence property of the LMS-type algorithms using geometrical interpretation," Proc. 9th DSP Symposium, (Kyoto, Japan), pp.9–14, Nov. 1994.

[12] S. Haykin, "Adaptive Filter Theory," Englewood Cliffs, NJ 07632: Prentice Hall, 1991.

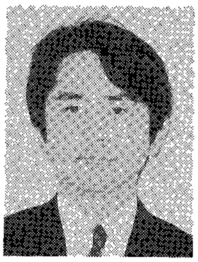
[13] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated gaussian data," IEEE Trans. on Information Theory, vol.34, pp.680–691, July 1988.

[14] V. Solo and X. Kong, "Adaptive Signal Processing Algorithms," Englewood Cliffs, NJ 07632: Prentice Hall, 1994.

[15] E.M. Landesman and M.R. Hestenes, "Linear Algebra for Mathmatics, Science, and Engineering," Englewood Cliffs, NJ 07632: Prentice Hall, 1992.
[16] K. Nishikawa and H. Kiya, "An LS based new gradient type adaptive algorithm – least square gradient," Trans. IEICE, vol.E77-A, pp.1417–1425, Sept. 1994.



Kiyoshi Nishikawa was born in Tokyo, Japan on December 12, 1966. He received the B.E. and the M.E. degrees in electrical engineering from Tokyo Metropolitan University in 1990 and 1992, respectively. From 1992 to 1993, he was at the Computer Systems Laboratory, Nippon Steel Corp. as a researcher. Since 1993, he has been with Tokyo Metropolitan University as an Assistant Professor. His research interest includes the adaptive digital processing. He is a member of IEEE SP, CAS, IT, and Computer Society.



Hitoshi Kiya was born in Yamagata, Japan on November 16, 1957. He received the B.E. and M.E. degrees in electrical engineering from Nagaoka University of Technology, and the D.E. degree in electrical engineering from Tokyo Metropolitan University in 1980, 1982, and 1987, respectively. In 1982, he joined Tokyo Metropolitan University, where he is currently an Associate Professor of Electronics and Information Engineering. His current research interests are in digital signal processing, multirate systems, adaptive filtering, image processing, and efficient algorithms for VLSI implementation. Dr. Kiya is a member of IEEE CAS, ASSP, IT, and COM Society, the institute of image electronics engineers of Japan and the institute of television engineers of Japan.