Steady State Analysis of Two-Dimensional LMS Adaptive Filters Using the Independence Assumption

Maha Shadaydeh Masayuki Kawamata
Graduate School of Engineering, Tohoku University
Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan
Tel: 022-217-7094, Fax: 022-263-9169
E-mail: maha@mk.ecei.tohoku.ac.jp

Abstract

In this paper, we consider the steady state Mean Square Error (MSE) analysis for 2-D LMS algorithm in which the filter’s weights are updated in both vertical and horizontal directions using Fornasini and Marchesini (F-M) state space model. The MSE analysis is conducted using the well-known independence assumption. First we show that computation of the Weight-Error Correlation Matrix (WECM) for F-M model-based 2-D LMS algorithm requires an approximation for the WECMs at large spatial lags. Then we propose a method to solve this problem. Further discussion is carried out for the special case when the input signal is white Gaussian. It is shown that a more strict condition on the upper bounds of the used step size values is required to ensure the convergence of the 2-D LMS in the MSE sense. Simulation experiments are presented to support the obtained analytical results.

keywords: 2-D LMS, steady state analysis, mean square error.

1 Introduction

The main advantage of using 2-D adaptive filters for processing 2-D signals, such as images, is in their ability to change the filtering characteristics based on the statistics of the processed data. Several 2-D adaptive FIR algorithms have been proposed so far [1]. Hadoud and Thomas [2] have proposed a 2-D LMS algorithm by extending the 1-D LMS algorithm [3]. In this algorithm the adaptive filter weight adaptation process is one-dimensional; hence, the convergence analysis is similar to that of the 1-D LMS [4]. Ohki and Hashiguchi [5] have proposed a 2-D LMS algorithm in which the filter’s weight vector is updated in two directions. Thus, the 2-D information of images can be exploited. The condition required for the weight-error vector mean to converge to the optimal solution has been investigated in [5] using stability theory of 2-D Fornasini and Marchesini (F-M) state space model [6]. However, convergence of the mean does not guarantee finite Mean Square Error (MSE).

In this paper, we consider the MSE analysis of 2-D LMS algorithm [5] using the assumption that the successive input vectors are statistically independent, Gaussian-distributed random variables. This assumption, generally referred to as the independence assumption [8], is widely used in the MSE analysis of 1-D LMS for two main reasons. The first is due to the simplification in analysis obtained under such assumption. The second is due to the good agreement between the theory obtained using the independence assumption and experimental results [4], [7]-[10].

Though the 2-D MSE analysis will be significantly simplified when invoking the independence assumption, the truly 2-D nature of the parameter update procedure in the algorithm [5] results in a new problem that is not encountered in the 1-D case. For the 1-D case, MSE analysis reduces mainly to the stability analysis of a set of first order coupled difference equations in the coefficients of the Weight Error Correlation Matrix (WECM). This set of difference equations maintains stability under a general condition imposed on the used step size value [9], [10]. For the 2-D case, as will be shown in the sequel, MSE analysis calls for stability analysis of a set of second-order coupled 2-D difference equations in the coefficients of the WECM, which is very difficult to handle analytically.

In this paper, we show that for the steady state, this set of 2-D difference equations can be reduced to a set of linear simultaneous equations in the coefficients of WECMs at different spatial lags; however, the number of the unknowns in this set exceeds the number of equations. To solve this problem, we propose a method for the approximation of the WECMs at large spatial lags. This approximation is based on extending the direct averaging method [7] to 2-D case. It can also serve as an approximation for the WECM without invoking the independence assumption when the step size values are sufficiently small.

Although the 2-D MSE analysis using the independence assumption may be regarded as not completely justified for 2-D adaptive filtering applications as it has been for the 1-D case [11]-[13], there are two motives for considering the analysis under such assumption. The first is to derive an analytical expression for the MSE that gives some insight to the performance of a truly 2-D LMS. The second is to shed a light on the problem that arises in 2-D MSE analysis. This problem is also expected to form a major difficulty in extending the 1-D MSE analysis approaches that do not invoke the independence assumption, such as [12] and [13], to 2-D case.
The organization of this paper is as follows. In Section 2, a brief review of the 2-D LMS algorithm [5] is given. In Section 3, the steady state MSE analysis of the 2-D LMS algorithm using the independence assumption is considered and a method for computing the 2-D WECM is presented. In Section 4, the special case when the input signal is white Gaussian is further discussed, and the condition required to ensure the convergence in the MSE sense is derived. Comparison between experimental and analytical results for the simplified case are presented in Section 5. Finally, Conclusions are drawn in Section 6.

2 Preliminaries

Consider the following $N \times N$ tapped 2-D adaptive FIR filter:

$$y(m, n) = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} h_{m,n}(k,l)x(m-k, n-l)$$

where $x(m, n)$ is a 2-D stationary input signal of size $M_1 \times M_2$. The error between the filter output $y(m, n)$ and the desired signal $d(m, n)$ is given by

$$e(m, n) = d(m, n) - y(m, n),$$

The 2-D LMS algorithm [5] updates the filter weight vector as follows:

$$H_{m+1,n+1} = f_h H_{m,n+1} + f_v H_{m+1,n} + \mu h e(m, n + 1)X_{m,n+1} + \mu v e(m + 1, n)X_{m+1,n};$$

with initial values: $H_{m,0} = 0, m = 0 \cdots M_1$, $H_{0,n} = 0, n = 0 \cdots M_2$;

$$f_h + f_v = 1$$

where $H_{m,n}$ and $X_{m,n}$ are respectively the adaptive filter's weight vector and the input data vector defined as

$$X_{m,n} = [x(m,n), \cdots, x(m-N+1, n-N+1)]^t$$

$$H_{m,n} = [h_{m,n}(0,0), \cdots, h_{m,n}(N-1, N-1)]^t.$$ (4)

In Eq. (3), $\mu_h$ and $\mu_v$ denote the step size values in the horizontal and vertical directions respectively.

From Eqs. (2) and (3), it can be shown that the error vector $C_{m+1,n+1}$ between the adaptive filter’s weight vector $H_{m,n+1}$ and the optimal weight vector $H_{opt}$ is updated as follows [5]:

$$C_{m+1,n+1} = H_{m+1,n+1} - H_{opt}$$

$$= (f_h I - \mu_h X_{m,n+1}X_{m,n+1}^t)C_{m,n+1} + (f_v I - \mu_v X_{m+1,n}X_{m+1,n}^t)C_{m+1,n} + \mu_h (d(m, n + 1)X_{m,n+1} - X_{m,n+1}X_{m,n+1}^tH_{opt})$$

$$+ \mu_v (d(m + 1, n)X_{m+1,n} - X_{m+1,n}X_{m+1,n}^tH_{opt}).$$ (5)

In [5], it has been shown that the adaptive algorithm converges to the optimal solution of Wiener-Hopf equation, i.e. $E\{C_{m,n}\} \to 0$ as $m + n \to \infty$, if the following condition holds:

$$|f_h - \mu_h \lambda_1| + |f_v - \mu_v \lambda_1| < 1$$ (6)

where $\lambda_1, i = 0, \cdots, p - 1$, are the eigenvalues of the input correlation matrix

$$R = E\{X_{m,n}X_{m,n}^t\}$$

with $E$ denoting the statistical expectation operator.

3 Steady State MSE Analysis for the 2-D LMS Algorithm

3.1 MSE Calculation

As an extension of the 1-D MSE analysis using the independence assumption [7] to 2-D case, we will consider the MSE analysis under the following assumptions:

A.1 The input vectors $X_{0,0}, X_{1,0}, \cdots, X_{m,n}$ are zero mean, statistically independent, Gaussian-distributed random variables.

A.2 The error signal $\varepsilon(m, n) = d(m, n) - X_{m,n}^tH_{opt}$

is a zero mean white Gaussian noise with variance $\sigma^2$.

A.3 The error signal $\varepsilon(m, n)$ and the input vector $X_{m,n}$ are statistically independent.

Now, substituting Eq. (1) in Eq. (2), and using (7) we find

$$e(m, n) = \varepsilon(m, n) - X_{m,n}^tC_{m,n}.$$ (8)

The steady state MSE is then given by

$$\epsilon_\infty = \lim_{m+n \to \infty} E\{e(m, n)^2\}$$

$$= \sigma^2 \epsilon + \lim_{m+n \to \infty} E\{C_{m,n}X_{m,n}X_{m,n}^tC_{m,n}\}$$

$$= \sigma^2 \epsilon + \lim_{m+n \to \infty} \text{tr}(R K_{m,n}^0)$$ (9)

where $K_{m,n}^0$ is the autocorrelation matrix of the weight-error vector defined as

$$K_{m,n}^0 = E\{C_{m,n}C_{m,n}^t\}. (10)$$

In analogy to the 1-D case, the misadjustment of the 2-D LMS is defined as

$$M = \frac{\epsilon_\infty - \epsilon_{\text{min}}}{\epsilon_{\text{min}}}$$

$$= \frac{\epsilon_\infty - \sigma^2 \epsilon}{\sigma^2 \epsilon}$$

$$= \frac{1}{\sigma^2} \lim_{m+n \to \infty} \text{tr}(R K_{m,n}^0).$$ (12)

In the rest of this section we will consider the calculation of the WECM.
3.2 Weight-Error Correlation Matrix

Before proceeding we need to define some necessary notations. Since the input correlation matrix $R$ is symmetric, there exists an orthogonal matrix $Q$ such that

$$QR^iQ^t = \Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{p-1})$$

(14)

$$Q^t = Q^{-1}.$$  

(15)

Thus, we can define the following transformed correlation matrix:

$$\Gamma_{m+1,m+1} = QE\{C_{m+1,m+1}\}Q^t,$$

(16)

where $E\{C_{m+1,m+1}\}$ is the 2-D autocorrelation function.

Now, substituting Eq. (5) in Eq. (11), and after many evaluations based on assumptions A1-A3, we can get

$$\Gamma_{m+1,m+1} = \gamma_{0,i}^{m+1,m+1} + \gamma_{1,i}^{m+1,m+1} + \gamma_{2,i}^{m+1,m+1} + \gamma_{3,i}^{m+1,m+1},$$

(17)

where

1. For the transformed correlation matrix defined in Eq. (16) we have

$$\gamma_{ij}^{m+1,m+1} \leq \gamma_{ij}^{m+1,m+1}.$$  

(18)

That is to say, the boundedness of the diagonal terms of WECM ensures the boundedness of the off-diagonal ones. Hence it is sufficient to consider the set of $N^2$ diagonal scalar equations. Note that, as for the MSE evaluation (see Eq. (10)), we are only interested in the diagonal terms since

$$\lim_{m+n \to \infty} \text{tr}(R\gamma_{0,i}^{m+1,m+1}) = \lim_{m+n \to \infty} \text{tr}(\gamma_{0,i}^{m+1,m+1}) = \lim_{m+n \to \infty} \sum_{j=0}^{p-1} \gamma_{ij}^{m+1,m+1} \lambda_j.$$  

(19)

2. If $\gamma_{ij}^{m+1,m+1}$ has a limit value, say $\gamma_{0,i}^{m+1,m+1}$, then for a finite integer $r$ we have

$$\lim_{m+n \to \infty} \gamma_{ij}^{m+r,n+r} = \gamma_{0,i}^{m+1,m+1}.$$  

(20)

Similarly, if $\gamma_{ij}^{m+r,n+r}$ has a limit value, say $\gamma_{0,i}^{m+1,m+1}$, then

$$\lim_{m+n \to \infty} \gamma_{ij}^{m+r,n+r} = \lim_{m+n \to \infty} \gamma_{0,i}^{m+1,m+1} = \gamma_{0,i}^{m+1,m+1}.$$  

(21)

Consequently, for the steady state, the $N^2$ diagonal coefficients of the WECM (17) should obey the equality

$$\gamma_{0,i}^{m+1,m+1} = \left(f f^t - (\mu_h f + \mu_v f)\lambda_i + (\lambda_f^2 + \lambda_v^2)\lambda_i^2\right),$$

(22)

where

$$\gamma_{0,i}^{m+1,m+1} = (f f^t - (\mu_h f + \mu_v f)\lambda_i + (\lambda_f^2 + \lambda_v^2)\lambda_i^2),$$

(23)

There is a need for another set of equations in the unknowns $\gamma_{0,i}^{m+1,m+1}$ and $\gamma_{1,i}^{m+1,m+1}$. If we apply the same way of analysis to evaluate the matrix

$$\Gamma_{m+1,m+1} = \gamma_{0,i}^{m+1,m+1} + \gamma_{1,i}^{m+1,m+1} + \gamma_{2,i}^{m+1,m+1} + \gamma_{3,i}^{m+1,m+1},$$

(24)

we can find that for the steady state, i.e. $m + n \to \infty$, the diagonal terms of the correlation matrix $\Gamma_{m+1,m+1}$ should obey the equality

$$\gamma_{1,i}^{m+1,m+1} = \left(f f^t - (\mu_h f + \mu_v f)\lambda_i + (\lambda_f^2 + \lambda_v^2)\lambda_i^2\right),$$

(25)

If we continue in similar way evaluating the WECMs for higher diagonal spatial lags, i.e.

$$\lim_{m+n \to \infty} \Gamma_{m+1,m+1} = \gamma_{k,i}^{m+1,m+1}; \quad k = 2, 3, \ldots,$$

at each stage $k, k = 0, 1, \ldots$, we will have a set of $(k + 1) \times N^2$ equations in $(k + 2) \times N^2$ unknowns, namely, $\gamma_{j,i}^{k,i}; 0 \leq j \leq k + 1, 0 \leq i \leq N^2 - 1$. One way to handle this problem is to find an approximation for the correlation coefficients $\gamma_{k,i}^{k+1}$ at specific stage $k \geq 1$. Or under the white Gaussian assumption for the input vector $X_{m,n}$, and the error signal $\varepsilon_{m,n}$, we can assume that for sufficiently large spatial lags, i.e. $k \gg 1$, the correlation coefficients $\gamma_{k,i}^{k+1}$ can be approximated with zero. Thus, the available $(k + 1) \times N^2$ equations can be solved for the $(k + 1) \times N^2$ unknowns to obtain $\gamma_{0,i}^{k,i}, i = 0, \ldots, N^2 - 1$. Then from Eqs. (10) and (19), the steady state MSE can be given by

$$\varepsilon_{\infty} = \sigma_e^2 + \sum_{j=0}^{p-1} \gamma_{0,j}^{i} \lambda_j.$$  

(26)
### 3.3 Approximation of the WECMs

Providing that the step sizes $\mu_h$ and $\mu_v$ are small, the direct averaging method [7] approximates the solution of the stochastic difference Eq. (5) with that of the following equation:

$$
C_{m+1,n+1} = A_h C_{m,n+1} + A_v C_{m+1,n} +
\mu_h \varepsilon_{m,n+1} X_{m,n+1} + \mu_v \varepsilon_{m+1,n} X_{m+1,n},
$$

where

$$A_h = f_h I - \mu_h R$$

$$A_v = f_v I - \mu_v R.$$ 

Eq. (26) is a 2-D F-M state space model with local state space vector $C_{m,n}$ and input vector $\varepsilon_{m,n} X_{m,n}$. This 2-D F-M model is exponentially stable if and only if [6]

$$\det(I - z_1^{-1} A_h - z_2^{-1} A_v) \neq 0$$

in the region

$$U_2 = \left\{ (z_1, z_2) \mid |z_1| \geq 1, |z_2| \geq 1 \right\}.$$ 

Note that the condition (27) is the same condition required for the convergence in the mean which was reduced in [5] to the condition (6).

Now, the transfer function between the input $\varepsilon_{m,n} X_{m,n}$ and state space vector $C_{m,n}$ is given by

$$H(z_1, z_2) = (I - A_h z_1^{-1} - A_v z_2^{-1})^{-1} (\mu_h z_1^{-1} + \mu_v z_2^{-1})$$

$$= (\mu_h z_1^{-1} + \mu_v z_2^{-1}) \sum_{k=0}^{\infty} (A_h z_1^{-1} + A_v z_2^{-1})^k$$

$$= (\mu_h z_1^{-1} + \mu_v z_2^{-1}) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} A^{i,j} z_1^{-i} z_2^{-j}$$

where the series expansion is absolutely convergent in the region $U_2$ [6], and

$$A^{0,0} = I$$

$$A^{i,j} = A_h A^{i-1,j} + A_v A^{i,j-1}, \text{ for } i+j > 0$$

$$A^{i,j} = 0, \text{ for } i < 0 \text{ or } j < 0.$$  

Hence, from Eq. (28), the weight-error vector $C_{m,n}$ can be calculated by

$$C_{m,n} = \sum_{i=0}^{m} \sum_{j=0}^{n} H(i, j) \varepsilon_{m-i,n-j} X_{m-i,n-j}$$

with

$$H(i, j) = \mu_h A^{i-1,j} + \mu_v A^{i,j-1}.$$ 

From Eq. (30), and using assumptions A2 and A3, the WECM $K^{k}_{m,n}$ can be calculated by

$$K^{k}_{m,n} = E\{C_{m,n-k} C^{d}_{m-k,n}\}$$

$$= \sigma^2 \sum_{i=0}^{m} \sum_{j=0}^{n} H(i, j - k) R H(i - k, j).$$  

Stability condition (6) guarantees that the spectral norm of each of the matrices $A_h, A_v, \text{ and } A^{i,j}$ are less than unity. And since these matrices are symmetric, it is straightforward to show that, $\lim_{i,j \to \infty} H(i, j) = 0$. Thus, we can deduce that the error that results from using the approximation (31) decreases as $k$ increases. For sufficiently large $k$, the correlation matrix $K^{k}_{m,n}$ can be approximated with zero as it has been suggested in Subsection 3.2.

Note that, for $k = 0$, Eq. (31) can be used as an approximation of the WECM $K^{0}_{m,n}$.

### 4 Steady State MSE Analysis for a Simplified Case

In this section we will discuss in more details the steady state analysis for the simplified case when the input signal is white Gaussian with variance $\sigma^2_2$; the correlation coefficients $\gamma^{i,j}$, $i = 0, \ldots, N^2 - 1$ are assumed to equal zero; $f_h = f_v$, and $\mu_h = \mu_v = \mu$. We choose to work with this case merely to make the solution of the equations traceable. Similar kind of analysis can be applied to any other case within which A1-A3 hold.

Since for the white Gaussian input case, $\lambda_0 = \cdots, \lambda_{N^2-1} = \sigma^2_2$, then it is clear from Eqs. (22) and (24) that $\gamma_0 = \gamma_0^{i,j} = \gamma_i$, $i = 0, \ldots, p - 1$. Hence, solving Eqs. (22) and (24) for $\gamma_0$ we get

$$\gamma_0 = \left[ \frac{\sigma^2_2}{\sigma^2_2} \left( \frac{\zeta^2}{0.25 - \zeta + (2 + p) \zeta^2} \right) \right] \times$$

$$\left[ \frac{0.375 - \zeta + (0.5 + 1.5p) \zeta^2 + (6 + 2p) \zeta^3 - (4 + 2p) \zeta^4}{0.125 + 3 \zeta - (2.5 + 1.5p) \zeta^2 - (6 + 2p) \zeta^3 + (4 + 2p) \zeta^4} \right]$$

where, for notational convenience, we have defined

$$\zeta = \mu \sigma^2_2.$$  

Now, since the weight-error covariance coefficient $\gamma_0$ should be positive and finite, the range of the step size $\mu$ that ensures the convergence of the 2-D LMS in the MSE sense can be determined by the following condition

$$0 \leq \gamma_0 < \infty.$$
For this simplified case, analysis of Eq. (32) reveals that in this equation, the first bracketed term and the numerator of the second bracketed term are always positive for $0 \leq \zeta < 1$. And that for any value of $p \geq 1$, the polynomial in the denominator of the second bracketed term has only one real positive root, say $\zeta_\infty$, in the range $0 \leq \zeta < 1$ where the sign of this polynomial changes from positive to negative. Thus we can deduce that the upper bound of the step size value that ensures finite variance is given by

$$0 \leq \mu < \frac{\zeta_\infty}{\sigma_x^2},$$

(35)

Fig. 1 shows the values of the root $\zeta_\infty$ for different values of $N = \sqrt{p}$. Fig. 2 shows $\gamma_0$ as a function of $\zeta$ with $\sigma_x^2/\sigma_e^2 = 1$ for a $2 \times 2$ tapped 2-D adaptive FIR filter, i.e. $p = 4$. From Fig. 1, it is clear that for any filter order, the condition required for the convergence in the MSE sense, as given in Eq. (35), decreases significantly the convergence region of the 2-D LMS algorithm when comparing to the convergence of the mean condition

$$0 \leq \mu < \frac{1}{\sigma_x^2}$$

(36)

given by Eq. (6).

5 Experiment

To support the analytical results, system identification experiment for the structure shown in Fig. 2 has been conducted for the following 2-D FIR filter:

$$d(m,n) = x(m,n) + 0.5x(m-1,n) + 0.5x(m,n-1) + 0.125x(m-1,n-1) + \varepsilon(m,n),$$

(37)

We used two independent, 2-D white Gaussian sequences with unit variances $\sigma_x^2$ and $\sigma_e^2$ for the input signal $x(m,n)$ and the additive noise $\varepsilon(m,n)$ respectively. For the purpose of comparison, the misadjustment of the 2-D LMS was calculated experimentally using Eqs. (9) and (12) as follows:

$$M = \frac{\sum_{m=1}^{M_1} \sum_{n=1}^{M_0} \sum_{m=1}^{M_2} \sum_{n=1}^{M_0} (C_{m,n}^t X_{m,n} X_{m,n}^t C_{m,n})}{(M_1 - l_1) \times (M_2 - l_2) \sigma_e^2}$$

(38)

where $l_1$ and $l_2$ were chosen large enough to ensure that the adaptive algorithm is in the steady state. Fig. 3 shows a comparison between experimental results and the misadjustment obtained using each of the proposed analytical methods. In the first (the independence assumption method), the coefficients of the WECM in Eq. (13) were calculated using Eq. (32). And in the second (the direct averaging method), the WECM in Eq. (13) was calculated using the approximation given in Eq. (31) with $k = 0$. We can observe that the MSE analysis using both the independent assumption and the direct averaging method gives satisfactory results for small step size value. However, as the step size exceed specific value, the error in the estimated MSE increases significantly. On the other hand, we can notice that the performance of the 2-D LMS was well preserved using the independence assumption based analysis, whereas, the direct averaging method fails completely for large step size value.

6 Conclusions

We have considered the steady state MSE analysis for 2-D LMS algorithm using the independence assumption. We have shown that the evaluation of the WECM for F-M model-based 2-D LMS algorithm requires approximation of the WECMs at larger spatial lags. Then, we have proposed a method to solve this problem. For a simplified case, it has been shown analytically that convergence in the MSE sense occurs for step size range that is significantly smaller than the one necessary for the convergence of the mean. Simulation example was presented to support the analytical results and to show that the analysis using the independence assumption does provide good insight to the performance of the 2-D LMS algorithm.

References

Figure 4: Comparison of the experimental results with the theoretical values for the misadjustment of the 2-D LMS in the simplified setting ($\lambda_0 = \cdots, \lambda_{p-1} = \sigma^2_x, f_h = f_v = 0.5, \text{ and } \mu_h = \mu_v = \mu$).


