IEICE TRANS. FUNDAMENTALS, VOL. E80-A, NO. 11 NOVEMBER 1997

PAPER

A Stability Analysis of Predictor-Based Least Squares Algorithm

Kazushi IKEDA†, Youhua WANG†, *and* Kenji NAKAYAMA†, *Members*

SUMMARY The numerical property of the recursive least squares (RLS) algorithm has been extensively studied. However, very few investigations are reported concerning the numerical behavior of the predictor-based least squares (PLS) algorithms which provide the same least squares solutions as the RLS algorithm. In Ref. [9], we gave a comparative study on the numerical performances of the RLS and the backward PLS (BPLS) algorithms. It was shown that the numerical property of the BPLS algorithm is much superior to that of the RLS algorithm under a finite-precision arithmetic because several main instability sources encountered in the RLS algorithm do not appear in the BPLS algorithm. This paper theoretically shows the stability of the BPLS algorithm by error propagation analysis. Since the time-variant nature of the BPLS algorithm, we prove the stability of the BPLS algorithm by using the method as shown in Ref. [6]. The expectation of the transition matrix in the BPLS algorithm is analyzed and its eigenvalues are shown to have values within the unit circle. Therefore we can say that the BPLS algorithm is numerically stable.

key words: *adaptive filter, RLS algorithm, fast RLS algorithm, numerical stability analysis*

1. Introduction

In solving the least squares problem for transversal adaptive filters, the recursive least squares (RLS) algorithm is well known. The principle of the RLS algorithm is based on the matrix inversion lemma in order to get the recursive equations. The RLS algorithm is characterized by a fast convergence rate and a high computational load. Concerning the numerical properties, much research has been done. The results show that divergence phenomenon may occur if the arithmetic precision is not enough or the input signal is ill-conditioned [1], [5], [7].

Another approach for solving the least squares problem is to use the fast least squares (FLS) algorithm. The principle of the algorithm is different from that of the RLS algorithm in that the relation of the forward and backward predictors and the gain vector is exploited, which results in a fast convergence rate with much less computation. However, the numerical instability of the FLS algorithm is so serious that it cannot be continuously used in real applications, especially under finite-precision implementation [2], [4].

The reason for the instability of the FLS algorithm is that the stable structure of the backward predictor has to be destroyed in order to get the recursive equation for computing the gain vector. So if we assume that recursion involves both order- and time-update, the least squares solution can be obtained by using either forward or backward predictor. Therefore, the stable structure of both forward and backward predictors is retained. This leads to the algorithms we called the predictor-based least squares (PLS) algorithms [8].

Although the PLS algorithms can be easily derived from the FLS algorithm, very few investigations concerning their numerical properties are reported in the literature. In Ref. [9], a comparative study on the numerical properties of the PLS and RLS algorithms was presented. Three main instability sources encountered in both the RLS and the FLS algorithms, including the unstable behavior of the conversion factor, the loss of symmetry, and the loss of positive definiteness of the inverse correlation matrix, have been investigated under the finite precision implementations. The results show that these instability sources do not exist in the PLS algorithm. Consequently, the PLS algorithm can provide a more robust and stable numerical performance than those of the RLS and the FLS algorithms. However, even though the effects of three instability sources do not appear in the PLS algorithm, the stability of the PLS algorithm can not be guaranteed without theoretical proof.

This paper gives the theoretical proof of the stability of the backward PLS (BPLS) algorithm. First, we show the approach for the proof, that is, the analysis of error propagation. Next, we apply the method the BPLS algorithm and show its stability.

2. Approach for Error Propagation Analysis

This paper adopts the method used in Ref. [6] for a numerical stability analysis of an adaptive algorithm. The method views an adaptive algorithm as a discrete-time nonlinear dynamical system that can be written in state-space form as

$$\theta(n) = f\left[\theta(n-1), \boldsymbol{u}_m(n)\right] \tag{1}$$

where $\theta(n)$ and $\boldsymbol{u}_m(n)$ denote the state vector and the tap-input vector, respectively. Since a finite precision implementation introduces errors, the infinite precision

Manuscript received February 22, 1997.
Manuscript revised July 8, 1997.

†The authors are with the Faculty of Engineering, Kanazawa University, Kanazawa-shi, 920 Japan

state vector $\theta(n)$ and the system (1) can be replaced by $\hat{\theta}(n) = \theta(n) + \Delta\theta(n)$ and

$$\hat{\theta}(n) = f\left[\hat{\theta}(n-1), \boldsymbol{u}_m(n)\right] + V(n), \tag{2}$$

where $\Delta\theta(n)$ denotes the state round-off error and $V(n)$ denotes the driving term that is the instantaneous contribution from roundoff noise. Assuming that the errors are small, we consider linearization of the system (2) around $\theta(n-1)$, which leads to

$$\Delta\theta(n) = A(n)\Delta\theta(n-1) + V(n) \tag{3}$$

where

$$A(n) = \nabla_\theta f[\theta, \boldsymbol{u}_m(n)]|_{\theta=\theta(n-1)}. \tag{4}$$

This is a linear time-variant system with a signal-dependent $A(n)$ matrix, whose exact statement about the deterministic stability is difficult to make. In fact, the state vector of the system may diverge if the worst input signals are given. Nevertheless, one can make certain statistical statements when the input signal $u(n)$ is stationary and ergodic, more precisely, it is shown in Ref. [6] that the state transition matrix

$$\Phi(n,0) = A(n)A(n-1)\cdots A(1)$$

has an asymptotic constant eigendecomposition that can be used to decide about the numerical stability of the original system (1) and then the numerical stability of (1) is determined by the eigenvalues of $\lim_{n\to\infty} \mathrm{Ex}[A(n)]$. Consequently, we show that the eigenvalues of $\lim_{n\to\infty} \mathrm{Ex}[A(n)]$ of the BPLS algorithm are within the unit circle in order to guarantee the numerical stability of the BPLS algorithm in the following sections.

3. Backward PLS Algorithm and Its State-Space Model

In the FLS algorithm [3], the gain vector $\boldsymbol{k}_M(n)$ is obtained using the following two order-update equations,

$$\boldsymbol{k}_{M+1}(n) = \begin{bmatrix} 0 \\ \boldsymbol{k}_M(n-1) \end{bmatrix} + \frac{f_{M+1}(n)}{F_{M+1}(n)}\boldsymbol{a}_{M+1}(n), \tag{5}$$

$$\boldsymbol{k}_{M+1}(n) = \begin{bmatrix} \boldsymbol{k}_M(n) \\ 0 \end{bmatrix} + \frac{b_{M+1}(n)}{B_{M+1}(n)}\boldsymbol{c}_{M+1}(n). \tag{6}$$

Notice that to derive $\boldsymbol{k}_M(n)$ from $\boldsymbol{k}_M(n-1)$, $\boldsymbol{k}_{M+1}(n)$ is used therefore reversely order-updated. This is known to be the main reason for the instability of the FLS algorithm [9]. In order to overcome this difficulty, the PLS algorithms use either Eqs. (5) or (6) to get $\boldsymbol{k}_M(n)$ so that the gain vectors $\boldsymbol{k}_m(n)$ and the predictors $\boldsymbol{a}_m(n)$ or $\boldsymbol{c}_m(n)$ for all $m = 1, \ldots, M$ have to be computed. This increases the computational load from $O(M)$ to $O(M^2)$ [8].

Since the forward PLS (FPLS) and BPLS algorithms, which can be derived from Eqs. (5) and (6), exhibit a very similar numerical performance, only the BPLS algorithm is studied in this paper. For convenience of analysis, we write the BPLS algorithm below:

$$\psi_m(n) = \boldsymbol{c}_m(n-1)^T\boldsymbol{u}_m(n), \tag{7}$$

$$B_m(n) = \lambda B_m(n-1) + \gamma_m(n)\psi_m(n)^2, \tag{8}$$

$$\boldsymbol{c}_m(n) = \boldsymbol{c}_m(n-1) - \psi_m(n)\begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix}, \tag{9}$$

$$\gamma_{m+1}(n) = \frac{\lambda B_m(n-1)}{B_m(n)}\gamma_m(n), \tag{10}$$

$$\boldsymbol{k}_m(n) = \begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix} + \frac{\gamma_m(n)\psi_m(n)}{B_m(n)}\boldsymbol{c}_m(n), \tag{11}$$

$$\alpha(n) = d(n) - \boldsymbol{w}_M(n-1)^T\boldsymbol{u}_M(n), \tag{12}$$

$$\boldsymbol{w}_M(n) = \boldsymbol{w}_M(n-1) + \boldsymbol{k}_M(n)\alpha(n), \tag{13}$$

$$m = 1, \ldots, M,$$

where $\psi_m(n)$ is the backward a priori prediction error, $B_m(n)$ is the minimum power of $\psi_m(n)$, $\gamma_m(n)$ is the conversion factor, $\boldsymbol{k}_m(n)$ is the gain vector, $\boldsymbol{c}_m(n)$ is the tap-weight vector of the backward predictor, $\alpha(n)$ is the a priori estimation error, $\boldsymbol{u}_m(n)$ is the tap-input vector, $d(n)$ is the desired signal, and $\boldsymbol{w}_M(n)$ is the tap-weight vector of the adaptive filter. The a posteriori backward prediction error $b_m(n)$ in Eq. (6) is equivalent to $\gamma_m(n)\psi_m(n)$ in the above definition.

To initialize the BPLS algorithm at time $n = 0$, set $\boldsymbol{c}_m(0) = [\mathbf{0}_{m-1}^T, 1]^T$, $B_m(0) = \delta$, $\boldsymbol{k}_m(0) = \mathbf{0}_m$, $\gamma_m(0) = 1$ for $m = 1, 2, \ldots, M$, where δ is a small positive constant and $\mathbf{0}_m$ is the $1 \times m$ vector all of whose elements are zero. And at each iteration $n \geq 1$, generate the first-order variables as $\gamma_1(n) = 1$ and $\boldsymbol{k}_0(n) = [\ ]$. Then, all of the variables are derived with Eqs. (7)–(13) when the input signal $u(i), i = 1, \ldots$ are given.

Next, we derive a state-space model of the BPLS algorithm. The first element of the state-vector is the tap-weight vector of the backward predictor $\boldsymbol{c}_m(n)$. Let $(\boldsymbol{k}_m(n)^T, \mathbf{0}_{j-m}^T)^T$ and $(\boldsymbol{c}_m(n)^T, \mathbf{0}_{j-m}^T)^T$ be denoted by $\boldsymbol{k}_m^j(n)$ and $\boldsymbol{c}_m^j(n)$, respectively. When $j < m$, $\boldsymbol{c}_m^j(n)$ is defined as the vector whose jth element is equal to that of $\boldsymbol{c}_m(n)$. Then, Eq. (9) is rewritten as

$$\boldsymbol{c}_m(n) = \boldsymbol{c}_m(n-1) - \psi_m(n)\boldsymbol{k}_{m-1}^m(n). \tag{14}$$

Substituting Eq. (7) to Eq. (14), we have

$$\boldsymbol{c}_m(n) = (E_m - \boldsymbol{k}_{m-1}^m(n)\boldsymbol{u}_m(n)^T)\boldsymbol{c}_m(n-1) \tag{15}$$

where E_m is an $m \times m$ identity matrix. Because the mth element of $\boldsymbol{c}_m(n)$ is constantly unity, we get the

transition formula of $\boldsymbol{c}_m(n)$ as

$$\begin{aligned}\boldsymbol{c}_m^{m-1}(n) = (&E_{m-1}\\ &-\boldsymbol{k}_{m-1}(n)\boldsymbol{u}_{m-1}(n)^T)\boldsymbol{c}_m^{m-1}(n-1)\\ &-u(n-m+1)\boldsymbol{k}_{m-1}(n). \end{aligned} \tag{16}$$

The second element of the state-vector is the minimum power of the backward prediction error $B_m(n)$. Substituting Eq. (7) to Eq. (8), the transition formula of $B_m(n)$ is written as

$$\begin{aligned}&B_m(n)\\ &= \lambda B_m(n-1)\\ &\quad +\gamma_m(n)\boldsymbol{c}_m(n-1)^T\boldsymbol{u}_m(n)\\ &\quad \cdot\boldsymbol{u}_m(n)^T\boldsymbol{c}_m(n-1)\\ &= \lambda B_m(n-1)\\ &\quad +\gamma_m(n)\boldsymbol{c}_m^{m-1}(n-1)^T\boldsymbol{u}_{m-1}(n)\\ &\quad \cdot\boldsymbol{u}_{m-1}(n)^T\boldsymbol{c}_m^{m-1}(n-1)\\ &\quad +2u(n-m+1)\boldsymbol{u}_{m-1}(n)^T\boldsymbol{c}_m^{m-1}(n-1)\\ &\quad +u(n-m+1)^2. \end{aligned} \tag{17}$$

It is worth noting here that the gain vector $\boldsymbol{k}_m(n)$ and the conversion factor $\gamma_m(n)$ are not state variables because they are given not by time-update but by order-update.

According to Sect. 2, the transition matrix $A(n)$ is given by differentiating the state-space model Eqs. (16) and (17) by the state vector

$$\theta(n) = \begin{pmatrix} \boldsymbol{c}_m^{m-1}(n) \\ B_m(n) \end{pmatrix}$$

which results in

$$A(n) = \begin{pmatrix} E_{m-1} - \boldsymbol{k}_{m-1}(n)\boldsymbol{u}_{m-1}(n)^T & 0 \\ 2\gamma_m(n)\psi_m(n-1)\boldsymbol{u}_m(n)^T & \lambda \end{pmatrix} \tag{18}$$

Since $0 < \lambda < 1$ and since $A(n)$ is block-lower-triangular, it remains to show that all the eigenvalues of $\mathrm{Ex}\left[E_{m-1} - \boldsymbol{k}_{m-1}(n)\boldsymbol{u}_{m-1}(n)^T\right]$ are asymptotically smaller than unity in magnitude.

4. Eigenvalues of the Transition Matrix

In this section, we show that the eigenvalues of $\mathrm{Ex}\left[E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T\right]$ are asymptotically within the unit circle. Actually, they are approximately all equal to λ unless numerical errors exist, since Ref. [3] shows that the gain vector $\boldsymbol{k}_m(n)$ is equal to $\Sigma_m(n)^{-1}\boldsymbol{u}_m(n)$ where

$$\Sigma_m(n) = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{u}_m(i)\boldsymbol{u}_m(i)^T$$

and then

$$\lim_{n\to\infty} \mathrm{Ex}\left[\boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T\right] \approx (1-\lambda)E_m$$

holds when $1-\lambda \ll 1$ because

$$\lim_{n\to\infty} \Sigma_m(n) \approx (1-\lambda)^{-1}R$$

and

$$R = \mathrm{Ex}\left[\boldsymbol{u}_m(n)\boldsymbol{u}_m(n)^T\right].$$

Consequently,

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Ex}\left[E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T\right] &\approx E_m - (1-\lambda)E_m\\ &= \lambda E_m,\end{aligned}$$

which confirms that all the eigenvalues of $\mathrm{Ex}\,[A(n)]$ converge approximately to λ as $n \to \infty$.

The argument above assumes the convergence of the gain vector $\boldsymbol{k}_m(n)$. In the following, we show more strictly that the eigenvalues of $\mathrm{Ex}\,[A(n)]$ exist within the unit circle under the assumption that the input signal is Gaussian. In case of $m = 1$, $\boldsymbol{c}_1(n)$ is constantly equal to 1 and then stable. So, by mathematical induction, we can assume the stability of $\boldsymbol{c}_i(n), i = 1, \ldots, m$ when we show that the eigenvalues of $\mathrm{Ex}\,[A(n)]$ are smaller than unity in magnitude, which leads to the stability of $\boldsymbol{c}_{m+1}(n)$.

Let $C_m(n)$, $D_m(n)$, and $\Psi_m(n)$ be defined as

$$C_m(n) = (\boldsymbol{c}_1^m(n), \ldots, \boldsymbol{c}_m^m(n)), \tag{19}$$

$$D_m(n) = \mathrm{diag}\,\{d_1(n), \ldots, d_m(n)\}, \tag{20}$$

$$d_i(n) = \frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)}{B_i(n)}, \tag{21}$$

$$\begin{aligned}\Psi_m(n) &= (\psi_1(n), \ldots, \psi_m(n))^T\\ &= C_m(n-1)^T\boldsymbol{u}_m(n),\end{aligned} \tag{22}$$

respectively. By recursive use of

$$\begin{aligned}\boldsymbol{k}_m(n) = &\left(1 - \frac{\gamma_m(n)\psi_m(n)^2}{B_m(n)}\right)\begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix}\\ &+\frac{\gamma_m(n)\psi_m(n)}{B_m(n)}\boldsymbol{c}_m(n-1)\end{aligned} \tag{23}$$

which is given by Eq. (9) and Eq. (11),

$$\begin{aligned}&\boldsymbol{k}_m(n)\\ &= \frac{\lambda B_m(n-1)}{B_m(n)}\boldsymbol{k}_{m-1}^m(n)\\ &\quad +\frac{\gamma_m(n)\psi_m(n)}{B_m(n)}\boldsymbol{c}_m(n-1)\\ &= \frac{\lambda B_m(n-1)}{B_m(n)}\boldsymbol{k}_{m-1}^m(n)\\ &\quad +\frac{\gamma_m(n)}{B_m(n)}\boldsymbol{c}_m(n-1)\boldsymbol{c}_m(n-1)^T\boldsymbol{u}_m(n)\\ &= \cdots\\ &= \sum_{i=1}^{m}\left(\prod_{j=i+1}^{m}\frac{\lambda B_j(n-1)}{B_j(n)}\right)\frac{\gamma_i(n)}{B_i(n)}\\ &\boldsymbol{c}_i(n-1)\boldsymbol{c}_i(n-1)^T\boldsymbol{u}_m(n)\\ &= \sum_{i=1}^{m}\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)}{B_i(n)}\boldsymbol{c}_i(n-1)\boldsymbol{c}_i(n-1)^T\boldsymbol{u}_m(n)\\ &= C_m(n-1)D_m(n)C_m(n-1)^T\boldsymbol{u}_m(n).\end{aligned} \tag{24}$$

is derived. Therefore, the transition matrix $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ is written as

$$\begin{aligned} &E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T \\ &\quad = E_m - C_m(n-1)D_m(n)C_m(n-1)^T \\ &\boldsymbol{u}_m(n)\boldsymbol{u}_m(n)^T \\ &\quad = C_m(n-1)(E_m \\ &\qquad -D_m(n)\Psi_m(n)\Psi_m(n)^T)C_m(n-1)^{-1}. \end{aligned}$$

Consequently, the eigenvalues of the transition matrix coincide with those of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$. From the assumption of the stability of $\boldsymbol{c}_i(n), i = 1,\ldots,m$, and then that of $C_m(n)$, we only need to consider the eigenvalues of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$. Since $\boldsymbol{c}_i(n)$ is the tap-weight vector of the backward predictor and $\psi_i(n)$ means the a priori backward prediction error by their definitions, when $\boldsymbol{c}_i(n)$ converges, $\mathrm{Ex}\left[\psi_i(n)^2\right]$ approaches to its minimal value, that is,

$$\begin{aligned} &\mathrm{Ex}\left[\psi_i(n)^2\right] \\ &\quad = \mathrm{Ex}\left[\boldsymbol{c}_i(n-1)^T\boldsymbol{u}_i(n)\boldsymbol{u}_i(n)^T\boldsymbol{c}_i(n-1)\right] \\ &\quad \rightarrow \min, \end{aligned}$$

and its derivatives by the jth element $(\boldsymbol{c}_i(n-1))_j$ of $\boldsymbol{c}_i(n-1)$ for $j = 1,\ldots,i-1$ become null, that is,

$$\begin{aligned} \frac{\partial \mathrm{Ex}\left[\psi_i(n)^2\right]}{\partial(\boldsymbol{c}_i(n-1))_j} &= \mathrm{Ex}\left[(\boldsymbol{u}_i(n))_j\boldsymbol{u}_i(n)^T\boldsymbol{c}_i(n-1)\right] \\ &= \mathrm{Ex}\left[\psi_i(n)u(n-j)\right] \\ &= 0. \end{aligned}$$

This means that $\psi_i(n)$ is statistically orthogonal to $u(n-j), j = 1,\ldots,i-1$, and then $\psi_j(n), j = 1,\ldots,i-1$ because $\psi_j(n)$ is a linear combination of $u(n),\ldots,u(n-j)$. Therefore,

$$\mathrm{Ex}\left[\psi_i(n)\psi_j(n)\right] = 0 \tag{25}$$

is satisfied when $i \neq j$. Equation (25) means that $\psi_i(n), i = 1,\ldots,m$ are statistically independent since $\Psi_m(n)$ obeys a Gaussian distribution by the assumption that $\boldsymbol{u}_m(n)$ is Gaussian. Since the i,j element of $D_m(n)\Psi_m(n)\Psi_m(n)^T$ is

$$\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)\psi_i(n)\psi_j(n)}{\lambda B_i(n-1)+\gamma_i(n)\psi_i(n)^2},$$

its expectation is zero when $i \neq j$ because of their independence, and the diagonal elements

$$\mathrm{Ex}\left[\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)\psi_i(n)^2}{\lambda B_i(n-1)+\gamma_i(n)\psi_i(n)^2}\right]$$

are between 0 and 1 because

$$0 < \frac{\gamma_i(n)\psi_i(n)^2}{\lambda B_i(n-1)+\gamma_i(n)\psi_i(n)^2} < 1$$

and

$$0 < \gamma_{m+1}(n) \leqq \gamma_{i+1}(n)$$

for any $\psi_i(n)$. So, it has been shown that all of the eigenvalues of the expectation of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$ exist between 0 and 1, and so do the eigenvalues of the expectation of

$$\begin{aligned} &C_m(n-1)(E_m \\ &\quad -D_m(n)\Psi_m(n)\Psi_m(n)^T)C_m(n-1)^{-1}. \end{aligned} \tag{26}$$

When $\psi_i(n), i = 1,\ldots,m$ are not independent because of the error of $C_m(n-1)$, the expectation of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$ does not become a diagonal matrix any more. However, if the error is small (for example, order of ε), the eigenvalues of the expectation also move a little (order of ε) and still exist within the unit circle.

Computer simulations are done to confirm that the eigenvalues of the ensemble-averaged transition matrix are asymptotically within the unit circle. A backward predictor with 10 taps is employed for the simulation. Each of the tap-weight vectors of the backward predictors $\boldsymbol{c}_m(n)$ and the gain vectors $\boldsymbol{k}_m(n)$ consists of an 8-bit exponent and a 3-bit mantissa. The input $u(n)$ is made by an AR model $(1, a_1, a_2))$ driven by a white Gaussian noise $N(0,1)$ where

$$a_1 = -2r_x\cos\theta, \tag{27}$$

$$a_2 = r_x^2, \tag{28}$$

$$r_x = 0.82, \tag{29}$$

$$\theta = \pi/4. \tag{30}$$

The initial parameter $\delta = 10$ and the forgetting factor $\lambda = 0.95$ are used. The eigenvalues of the ensemble average (50 samples) of the transition matrices at time $n = 15$ (early stage) and $n = 50$ (convergence stage) are calculated. The simulation results shown in Fig. 1 clearly demonstrate that the eigenvalues are within the unit circle in both cases. More precisely, they are scattered around their theoretical asymptotic value of $(\lambda, 0)$, even though the finite-precision arithmetic is used for computing the predictors and the gain vectors, which support the theoretical results given above.

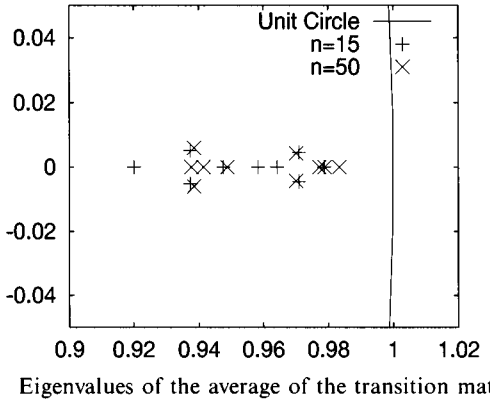


Fig. 1 Eigenvalues of the average of the transition matrices at $n = 15$ and $n = 50$.

5. Conclusion

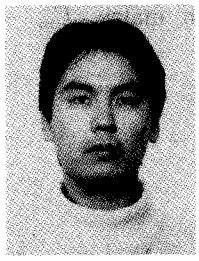
The numerical stability of the BPLS algorithm has been proven in this paper using the method of error-propagation analysis, that is, the expectation of the transition matrix is shown to have eigenvalues within the unit circle. The computer simulation results also show that the transition matrix has eigenvalues within the unit circle in average even under a low-bit word-length implementation. Since the superiority of the BPLS algorithm over the RLS algorithm is supported not only experimentally but also theoretically, it is very promising that the use of the RLS algorithm and its square root versions may be replaced by the BPLS algorithm.

Acknowledgement

The authors would like to express their cordial thanks to the anonymous reviewers whose suggestions have greatly improved the derivation and justification of the results.

References

[1] G.E. Bottomley and S.T. Alexander, "A theoretical basis for the divergence of conventional recursive least squares filters," Proc. ICASSP, pp.908–911, 1989.
[2] J.M. Cioffi, "Limited-precision effects in adaptive filtering," IEEE Trans. CAS, vol.34, pp.821–833, 1987.
[3] S. Haykin, "Adaptive Filter Theory," Prentice-Hall, 2nd edition, 1991.
[4] D.W. Lin, "On digital implementation of the fast Kalman algorithm," IEEE Trans. ASSP, vol.32, pp.998–1005, 1984.
[5] S. Ljung and L. Ljung, "Error propagation properties of recursive least squares adaptation algorithms," Automatica, vol.21, pp.157–167, 1985.
[6] D.T.M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," IEEE Trans. SP, vol.39, pp.92–113, 1991.
[7] M.H. Verhaegen, "Round-off error propagation in four generally-applicable, recursive, least-squares estimation schemes," Automatica, vol.25, pp.437–444, 1989.
[8] Y. Wang and K. Nakayama, "A stable least square algorithm based on predictors and its application to fast Newton transversal filters," IEICE Trans. Fundamentals, vol.E78-A, no.8, pp.999–1003, 1995.
[9] Y. Wang and K. Nakayama, "Numerical Performances of Recursive Least Squares and Predictor Based Least Squares: A Comparative Study," IEICE Trans. Fundamentals, vol.E80-A, no.4, pp.745–752, 1997.



Kazushi Ikeda was born in Shizuoka, Japan, in 1966. He received the B.E., M.E., and Ph.D. degrees in Mathematical Engineering and Information Physics from the University of Tokyo, Tokyo, Japan, in 1989, 1991, and 1994, respectively. In 1995, he was a Research Associate in the Department of Computer Science at the Chinese University of Hong Kong, Hong Kong. He is currently a Research Associate in the Department of Electrical and Computer Engineering at Kanazawa University, Ishikawa, Japan. His research interests include computational learning theory, adaptive systems, and neuro-dynamics.

Youhua Wang received the M.S. degree from Shanghai University of Technology, Shanghai, China, in 1988, and the Doctor of Engineering degree from Kanazawa University, Japan, in 1995. From 1988 to 1990, he held a research and teaching position at the Advanced Technology Research Center, Shenzhen University, Shenzhen, China. From 1995 to 1996, he worked on speech coding and DSP development technology with Matsushita Communication Kanazawa R&D Laboratories. He is currently with the Department of Electrical and Computer Engineering, Kanazawa University as an Assistant Professor. His research interests lie in the areas of digital signal processing, adaptive filtering, and their applications to communications.

Kenji Nakayama received the B.E. and Dr. degrees in electronics engineering from Tokyo Institute of Technology (TIT), Tokyo, Japan, in 1971 and 1983, respectively. From 1971 to 1972 he was engaged in the research on classical network theory in TIT. He was involved in NEC Corp. from 1972 to 1988, where his research subjects were filter design methodology and signal processing algorithms. He joined the Department of Electrical and Computer Engineering at Kanazawa University, in 1988, where he is currently a Professor. His current research interests include neural networks and adaptive filters.