AN ESTIMATOR FOR THE EIGENVALUES OF THE SYSTEM MATRIX OF A PERIODIC-REFERENCE LMS ALGORITHM

Tuomas Haarnoja and Kari Tammi

VTT Technical Research Centre of Finland
P.O. Box 1000, FI-02044 VTT, Finland
firstname.lastname@vtt.fi

Kai Zenger

Aalto University, School of Electrical Engineering
P.O. Box 15500, FI-00076 Aalto, Finland
kai.zenger@aalto.fi

ABSTRACT

The convergence analysis of the Least Mean Square (LMS) algorithm has been conventionally based on stochastic signals and describes thus only the average behavior of the algorithm. It has been shown previously that a periodic-reference LMS system can be regarded as a linear time-periodic system whose stability can be determined from the monodromy matrix. Generally, the monodromy matrix can only be solved numerically and does not thus reveal the actual factors behind the dynamics of the system. This paper derives an estimator for the eigenvalues of the monodromy matrix. The estimator is easy to calculate, and it also reveals the underlying reason for the bad convergence of the LMS algorithm in some special cases. The estimator is confirmed by comparing it to the precise eigenvalues of the monodromy matrix. The estimator is found to be accurate for the eigenvalues close to unity.

***Index Terms*—** LMS algorithm, monodromy matrix, estimator for the eigenvalues, convergence rate

1. INTRODUCTION

Although the basis of the Least Mean Square (LMS) algorithm was discovered already in 1960's [1, 2], the convergence analysis of it is still not completely understood. Most of the approaches rely on the independence assumption proposed in [3], which states that consecutive reference vectors do not correlate. This assumption has been successfully used in a variety of theoretical studies to derive stability and convergence conditions in the presence of stochastic signals (*e.g* [4, 5]). However, when the reference signal is periodic, the independence assumption is not justified, and it ultimately results in unstable operation of the controller. On the other hand, the periodic reference enables the usage of precise, deterministic analysis tools. For example, Glover [6] found that when the reference signal consists of a single sinusoid, the LMS system can be written as a sum of a Linear Time-Invariant (LTI) and Linear Time-Varying (LTV) transfer function. Similar results were obtained in [7], and in [8] the method was extended to cover a multi-frequency case.

Later, Tammi *et al.* [9] derived an exact time-periodic state-space representation for the algorithm in the case of periodic reference. The work was continued in [10], where a state-space representation for the Filtered-x LMS (FxLMS) algorithm with a full linear plant model (also known as the secondary path) was introduced. The drawback of the method is, however, that the obtained state-space representation is over-large because of the existence of numerous non-controllable and non-observable states. Although these methods are able to predict the possible problem in the convergence of the filter, they still cannot point out the actual reasons for it.

This paper uses a more sophisticated time-periodic state-space representation for the LMS algorithm with a periodic reference signal. The new representation is fully observable and controllable. By assuming a slowly converging system, an approximation for the eigenvalues can be derived. The approximation does not require heavy computing effort and it reveals the reason for the possible problems in convergence as well as lets us approximately calculate the required convergence coefficient when the magnitude of the eigenvalues are specified. The precise LMS model is then used to calculate the exact eigenvalues in several cases, and the values are compared to the ones given by the estimator.

2. STATE-SPACE REPRESENTATION OF AN LMS SYSTEM

Fig. 1 shows a block diagram of the LMS algorithm. With the definitions

$$\mathbf{w}(n) = \begin{bmatrix} w_0(n) & \dots & w_N(n) \end{bmatrix}^{\mathrm{T}} \tag{1a}$$

$$\mathbf{r}(n) = \begin{bmatrix} r_0(n) & \dots & r_N(n) \end{bmatrix}^{\mathrm{T}}, \tag{1b}$$

The work reported in this paper was supported by the Academy of Finland.

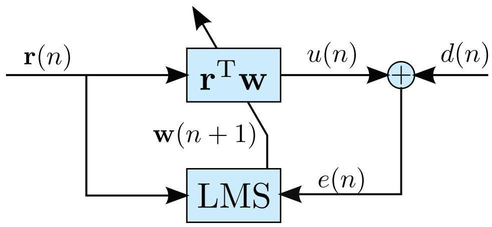

Fig. 1. A block diagram for the LMS filter. The reference signal is a vector whose entries are independent from each other.

where $N+1$ is the number of filter taps, the dynamics of the system can be defined with

$$u(n) = \mathbf{r}^{\mathrm{T}}(n)\mathbf{w}(n) \tag{2a}$$
$$e(n) = u(n) + d(n) \tag{2b}$$
$$\mathbf{w}(n+1) = \mathbf{w}(n) - \alpha\mathbf{r}(n)e(n). \tag{2c}$$

The notation used here is more general than conventionally because $r_i(n)$'s are considered to be independent reference signals. The original LMS is a special case, in which $r_i(n)$'s are sampled from the same signal: $r_i(n) = r(n-i)$. It is assumed throughout this paper that the reference vector is periodic. Therefore, the reference can be written as a sum of sinusoids as

$$r_i(n) = \sum_{j=1}^{M} A_{ij} \sin\left(\omega_j n + \phi_{ij}\right), \tag{3}$$

where M is the total number of discrete frequencies in the reference, and A_{ij} and ϕ_{ij} are the amplitude and the phase at the normalized frequency of ω_j. The period of the reference is T_{p}, which is the smallest positive integer for which $\mathbf{r}(n+T_{\mathrm{p}}) = \mathbf{r}(n)$.

The reference vector can be written in a matrix form as

$$\mathbf{r}(n) = \mathbf{R}_0\mathbf{s}(n), \tag{4}$$

where

$$\mathbf{R}_0 = \begin{bmatrix} A_{01}\cos(\phi_{01}) & A_{01}\sin(\phi_{01}) & \dots & A_{0M}\sin(\phi_{0M}) \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1}\cos(\phi_{N1}) & A_{N1}\sin(\phi_{N1}) & \dots & A_{NM}\sin(\phi_{NM}) \end{bmatrix} \tag{5}$$

and

$$\mathbf{s}(n) = \begin{bmatrix}\sin(\omega_1 n) & \cos(\omega_1 n) & \dots & \cos(\omega_M n)\end{bmatrix}^{\mathrm{T}}. \tag{6}$$

With this notation, the system (2) can be written as

$$\mathbf{w}(n+1) = \left(\mathbf{I} - \alpha\mathbf{R}_0\mathbf{s}(n)\mathbf{s}^{\mathrm{T}}(n)\mathbf{R}_0^{\mathrm{T}}\right)\mathbf{w}(n), \tag{7}$$

where $d(n)$ is set to zero, because we are only interested in the system matrix, and $d(n)$ would appear as an input of the system. This can be further modified if $\mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0$ is invertible:

$$\mathbf{c}(n+1) = \underbrace{\left(\mathbf{I} - \alpha\mathbf{s}(n)\mathbf{s}^{\mathrm{T}}(n)\mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0\right)}_{\mathbf{A}(n)}\mathbf{c}(n), \tag{8}$$

where $\mathbf{c}(n) = (\mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0)^{-1}\mathbf{R}_0^{\mathrm{T}}\mathbf{w}(n)$. The state vector $\mathbf{c}(n)$ contains the same time-dependent information as $\mathbf{w}(n)$. This will be shown in a later paper and is omitted from here for brevity. The dimension of $\mathbf{A}(n)$ is $2M \times 2M$. Because $\mathbf{A}(n)$ is periodic with the period of T_{p}, the convergence properties of the algorithm are fully described by the time-invariant monodromy matrix

$$\mathcal{A} = \prod_{n=1}^{T_{\mathrm{p}}} \mathbf{A}(n). \tag{9}$$

The calculation of the eigenvalues of the monodromy matrix ($\lambda_{\mathcal{A},i}$) is, however, time consuming, and it cannot be done analytically in a general case. The next section will introduce a simple approximation for the magnitude of these eigenvalues.

3. APPROXIMATION FOR THE EIGENVALUES OF THE MONODROMY MATRIX

The aim is to find an estimate $\tilde{\lambda}_{\mathcal{A},i}$ for the eigenvalues $\lambda_{\mathcal{A},i}$. Let us define $\lambda_{\mathbf{R},i}$ as

$$\mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0\mathbf{v}_i = \lambda_{\mathbf{R},i}\mathbf{v}_i, \tag{10}$$

i.e. $\lambda_{\mathbf{R},i}$ is the i^{th} eigenvalue of $\mathbf{R} = \mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0$, and $\mathbf{v}_i$ is the corresponding eigenvector. All the $\mathbf{v}_i$'s are orthogonal, because $\mathbf{R}$ is symmetric. Now we can write

$$\mathbf{A}(n)\mathbf{v}_i = \underbrace{\left(\mathbf{I} - \alpha\lambda_{\mathbf{R},i}\mathbf{s}(n)\mathbf{s}^{\mathrm{T}}(n)\right)}_{\mathbf{A}_i(n)}\mathbf{v}_i, \tag{11}$$

where

$$\mathbf{A}_i(n) = \mathbf{I} - \alpha\lambda_{\mathbf{R},i}\mathbf{s}(n)\mathbf{s}^{\mathrm{T}}(n) \tag{12}$$

is the 'directional' system matrix into direction of the eigenvector $\mathbf{v}_i$. Let us assume that the system is in slow convergence, *i.e.*

$$\mathbf{A}(n)\mathbf{x} \approx \mu_{\mathbf{A}(n)\mathbf{x}}\mathbf{x}, \tag{13}$$

for every $\mathbf{x}$, where the value of the scalar $\mu_{\mathbf{A}(n)\mathbf{x}}$ depends on both the matrix $\mathbf{A}(n)$ and the orientation of the vector $\mathbf{x}$. From (9), (11), and (13) we get

$$\mathcal{A}\mathbf{v}_i \approx \prod_{n=1}^{T_{\mathrm{p}}} \mu_{\mathbf{A}_i(n)\mathbf{v}_i}\mathbf{v}_i. \tag{14}$$

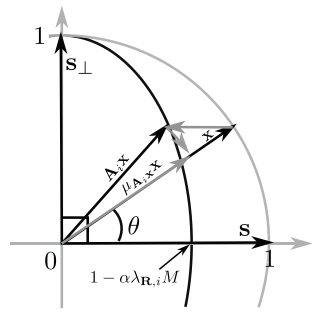

Fig. 2. The figure illustrates the projection of the $2M$ dimensional eigenspace of $\mathbf{A}_i(n)$ onto two dimensional plane spanned by the vectors $\mathbf{s}(n)$ and $\mathbf{x}$. The vector $\mathbf{s}_\perp(n)$ is orthogonal to $\mathbf{s}(n)$. The gray circle is the unit circle, and the black oval illustrates the locus obtained by multiplying the unit circle by $\mathbf{A}_i(n)$. The dependencies on n are omitted for simplicity.

The value of the product

$$\prod_{n=1}^{T_\mathrm{p}} \mu_{\mathbf{A}_i(n)\mathbf{v}_i} = \tilde{\lambda}_{\mathcal{A},i} \tag{15}$$

is thus the approximative eigenvalue we are searching for.

Let us now concentrate only on the directional system matrix $\mathbf{A}_i(n)$ and try to find the estimate (15). One of the eigenvectors of $\mathbf{A}_i(n)$ is $\mathbf{s}(n)$:

$$\begin{aligned}\mathbf{A}_i(n)\mathbf{s}(n) &= \mathbf{s}(n) - \alpha\lambda_{\mathbf{R},i}\mathbf{s}(n)\underbrace{\mathbf{s}^\mathrm{T}(n)\mathbf{s}(n)}_{M} \\ &= (1-\alpha\lambda_{\mathbf{R},i}M)\,\mathbf{s}(n).\end{aligned} \tag{16}$$

Because $\mathbf{A}_i(n)$ is a symmetric matrix, the rest of the eigenvectors are perpendicular to $\mathbf{s}(n)$. Let us denote $\mathbf{s}_\perp(n) \perp \mathbf{s}(n)$ as any linear combination of these perpendicular vectors. For $\mathbf{s}_\perp(n)$ we get

$$\begin{aligned}\mathbf{A}_i(n)\mathbf{s}_\perp(n) &= \mathbf{s}_\perp(n) - \alpha\lambda_{\mathbf{R},i}\mathbf{s}(n)\underbrace{\mathbf{s}^\mathrm{T}(n)\mathbf{s}_\perp(n)}_{=0} \\ &= \mathbf{s}_\perp(n),\end{aligned} \tag{17}$$

i.e. the eigenvalues corresponding to the other eigenvectors are all 1.

This situation can be illustrated as in Fig. 2. The figure shows a two dimensional plane of the eigenspace of $\mathbf{A}_i(n)$. The plane is spanned by $\mathbf{s}(n)$ and an arbitrary vector $\mathbf{x}$. The vector $\mathbf{s}_\perp(n)$ is selected so that it is perpendicular $\mathbf{s}(n)$ and it lies on the same plane. The quarter of the gray circle is the unit circle on the plane, and the part of the black oval shows how the unit circle transforms when multiplied by $\mathbf{A}_i(n)$. The goal is now to calculate the value of $\mu_{\mathbf{A}_i(n)\mathbf{v}_i}$. Let us consider each component of $\mathbf{v}_i(n)$ separately, *i.e.* let us calculate the value of $\mu_{\mathbf{A}_i(n)\mathbf{x}_j}$, where $\mathbf{x}_j = \begin{bmatrix}\delta_{1j} & \delta_{2j} & \delta_{3j} & \dots\end{bmatrix}^\mathrm{T}$. The angle between $\mathbf{s}(n)$ and vector $\mathbf{x}_j$ is

$$\begin{aligned}\cos(\theta_j(n)) &= \frac{\mathbf{s}(n)\cdot\mathbf{x}_j}{\|\mathbf{s}(n)\|_2\,\|\mathbf{x}_j\|_2} \\ &= \frac{\sin(n\omega_j+\phi_j)}{\sqrt{M}},\end{aligned} \tag{18}$$

where ϕ_j is either 0 or $\pi/2$ depending on if we are dealing with sine or cosine component of $\mathbf{s}(n)$.

From Fig. 2, one can see that

$$\begin{aligned}\mu_{\mathbf{A}_i(n)\mathbf{x}_j} =& \sqrt{(1-\alpha\lambda_{\mathbf{R},i}M)^2\cos^2(\theta_j(n)) + \sin^2(\theta_j(n))} \\ =& \sqrt{1-(2\alpha\lambda_{\mathbf{R},i}M - (\alpha\lambda_{\mathbf{R},i}M)^2)\cos^2(\theta_j(n))},\end{aligned} \tag{19}$$

and by substituting (18) into (19), we get

$$\mu_{\mathbf{A}_i(n)\mathbf{x}_j} = \sqrt{1-\left(2\alpha\lambda_{\mathbf{R},i} - M\left(\alpha\lambda_{\mathbf{R},i}\right)^2\right)\sin^2(n\omega_j+\phi_j)}. \tag{20}$$

Because $\sin^2(n\omega_j+\phi_j) = \sin^2\left((n+\frac{T_\mathrm{p}}{2})\omega_j+\phi_j\right)$ (for even T_p), we can write

$$\begin{aligned}\prod_{n=1}^{T_\mathrm{p}} \mu_{\mathbf{A}_i(n)\mathbf{x}_j} =& \prod_{n=1}^{T_\mathrm{p}} \sqrt{1-(2\alpha\lambda_{\mathbf{R},i} - M\left(\alpha\lambda_{\mathbf{R},i}\right)^2)\sin^2(n\omega_j+\phi_j)} \\ =& \prod_{n=1}^{T_\mathrm{p}/2} \left(1-(2\alpha\lambda_{\mathbf{R},i} - M\left(\alpha\lambda_{\mathbf{R},i}\right)^2)\sin^2(n\omega_j+\phi_j)\right) \\ =& 1-2\alpha\lambda_{\mathbf{R},i}\underbrace{\sum_{n=1}^{T_\mathrm{p}/2}\sin^2(n\omega_j+\phi_j)}_{T_\mathrm{p}/4} + \mathcal{O}\left\{\left(\alpha\lambda_{\mathbf{R},i}\right)^2\right\} \\ =& 1-\frac{1}{2}\alpha\lambda_{\mathbf{R},i}T_\mathrm{p} + \mathcal{O}\left\{\left(\alpha\lambda_{\mathbf{R},i}\right)^2\right\},\end{aligned} \tag{21}$$

where $\mathcal{O}\left\{\left(\alpha\lambda_{\mathbf{R},i}\right)^2\right\}$ is the error term that diminishes at the rate of $\left(\alpha\lambda_{\mathbf{R},i}\right)^2$ when $\alpha\lambda_{\mathbf{R},i} \to 0$.

Because $\prod_{n=1}^{T_\mathrm{p}} \mu_{\mathbf{A}_i(n)\mathbf{x}_j}$ does not depend on $\mathbf{x}_j$, we have

$$\prod_{n=1}^{T_\mathrm{p}} \mu_{\mathbf{A}_i(n)\mathbf{x}_j} = \prod_{n=1}^{T_\mathrm{p}} \mu_{\mathbf{A}_i(n)\mathbf{v}_i} \approx 1-\frac{1}{2}\alpha\lambda_{\mathbf{R},i}T_\mathrm{p} = \tilde{\lambda}_{\mathcal{A},i}, \tag{22}$$

which is the eigenvalue we are looking for.

4. CONFIRMATION OF THE ESTIMATOR

Fig. 3 shows a comparison between the estimates and exact eigenvalues. The values are calculated as a function of the

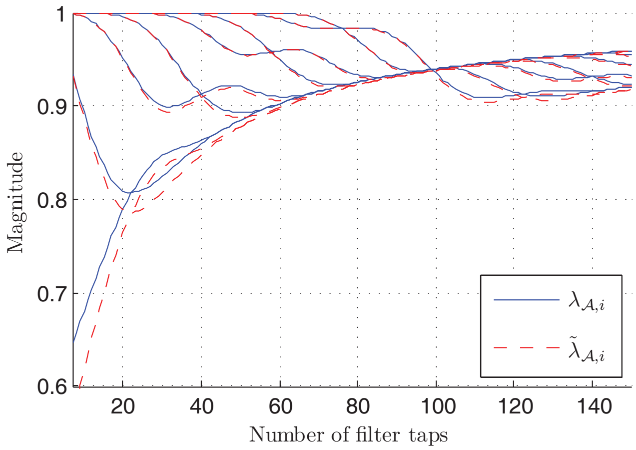

Fig. 3. The precise eigenvalues $\lambda_{\mathcal{A},i}$ (solid line) and the approximations $\tilde{\lambda}_{\mathcal{A},i}$ (dashed line) of the LMS system calculated for several filter lengths.

number of filter taps. The convergence coefficient is chosen as $\alpha = 0.0025/(N+1)$ to prevent the system from becoming unstable on large values of N. The sampling time in each case is $\delta t = 0.001$ $[s]$ and the reference vector is $r_i(n) = \sum_{j=1}^{4} \sin((n-i)j\omega_0)$, where $\omega_0 = 10\delta t 2\pi$. Because now $M = 4$, the order of the system or the number of eigenvalues is 8.

The convergence rate is not adequate for a practical system when the filter length $N < 60$ because the magnitude of at least one of the eigenvalue is close to unity. For filter lengths $N > 100$ the filter time span exceeds the period of the reference, and thus no extra information is obtained from the signal, and the largest eigenvalue does not decrease anymore. At $N = 100$ the filter time span matches the period of the reference, which is the point where all the eigenvalues have the same magnitude. The estimator is very accurate for the eigenvalues greater than approximately 0.9. For smaller eigenvalues the estimation worsens, but is still able to predict the shape of the curve roughly. When the actual eigenvalues are approximately < 0.6 (not shown in the figure), the error becomes large, and the estimates cannot be utilized anymore.

5. CONCLUSION

This work presents a derivation of an estimator for the eigenvalues of an LMS system. The estimator reveals that a single matrix, namely $\mathbf{R}_0$, is the main contributor to the dynamics of the system. Small eigenvalues of the matrix $\mathbf{R}_0^{\mathrm{T}}\mathbf{R}_0$ (or singular values of $\mathbf{R}_0$) indicates the convergence rate of the system will not be sufficient. This happens, for example, if a small number of filter taps is used in a combination with multi-frequency reference. The smallest singular values begin to increase rapidly after the filter time span reaches approximately half of the period of the reference. The estimator cannot, however, predict instability of the system, because possible instabilities are neglected in the derivation of the estimator by assuming a slowly converging system.

6. REFERENCES

[1] B. Widrow and M.E. Hoff, "Adaptive switching circuits," *IRE WESCON CONVENTION RECORD Part 4*, 1960.

[2] C.R. Fuller, S.J. Elliott, and P.A. Nelson, *Active Control of Vibration*, Academic Press, 1997.

[3] W.A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique," *Signal Processing*, vol. 6, no. 2, pp. 113–133, 1984.

[4] E. Bjarnason, "Analysis of the filtered-x LMS algorithm," *Speech and Audio Processing, IEEE Transactions on*, vol. 3, no. 6, pp. 504–514, 1995.

[5] I.T. Ardekani and W.H. Abdulla, "Theoretical convergence analysis of FxLMS algorithm," *Signal Processing*, vol. 90, no. 12, pp. 3046–3055, 2010.

[6] Jr. Glover, J., "Adaptive noise canceling applied to sinusoidal interferences," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 25, no. 6, pp. 484–491, Dec. 1977.

[7] N.J. Bershad and P.L. Feintuch, "Non-wiener solutions for the LMS algorithm – a time domain approach," *Signal Processing, IEEE Transactions on*, vol. 43, no. 5, pp. 1273–1275, 1995.

[8] L. Vicente and E. Masgrau, "Novel FxLMS convergence condition with deterministic reference," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3768 –3774, oct. 2006.

[9] K. Tammi, A. Laiho, E. Lantto, K. Zenger, and A. Arkkio, "On convergence problems of adaptive FIR filter with LMS algorithm," in *The 16th Int. Congress on Sound and Vibration*, Kraków, July 2009.

[10] T. Haarnoja, K. Tammi, and K. Zenger, "Linear time-periodic representation for periodic-reference, discrete-time filtered-x LMS algorithm," in *18th International Congress on Sound & Vibration*, 2011.