THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA VOL.25, NO.2E 2006. 6 pp. 43-48

A Recursive Data Least Square Algorithm and Its Channel Equalization Application

Jun-seok Lim*, Jae Soo Kim**
*Department of Electroncis Engineering Sejong University
**Department of Ocean Engineering Korea Maritime University
(Received March 14 2006; Revised Jun 2 2006; Accepted Jun 13 2006)

Abstract

Abstract-Using the recursive generalized eigendecomposition method, we develop a recursive form solution to the data least squares (DLS) problem, in which the error is assumed to lie in the data matrix only. Simulations demonstrate that DLS outperforms ordinary least square for certain types of deconvolution problems.

Keywords: *Data least square method, Generalized eigenvalue problem, Equalization.*

I. Introduction

Linear least squares (LS) problems involve finding "good" approximate solutions to a set of independent, but inconsistent, linear equations

$$A\,x = b, \tag{1}$$

where A is an m x n complex data matrix; b is a complex m × 1 observation vector; and x is a complex n × 1 prediction vector, which is optimally chosen to minimize some kind of squared error measure. It is usually assumed that the underlying noiseless data satisfy (1) with equality. Different classes of LS problems can be defined in terms of the type of perturbation necessary to achieve equality in the system of equations described by (1). For example, in the ordinary least squares (OLS) problem, the error (or perturbation) is assumed to lie in b.

$$Ax_{OLS} = (b + r), \tag{2}$$

where r is the residual error vector that corresponds to a perturbation in b. The OLS solution vector x_{OLS} is chosen so that the Euclidean (or Frobenius) norm of r is minimized. It is implicitly assumed in the OLS problem that A is completely errorless, and therefore the columns of A are not perturbed in the solution [1]. On the other hand, the total least squares (TLS) problem assumes error in both A and b.

$$(A + E\,)x_{TLS} = (b + r). \tag{3}$$

The TLS solution vector is chosen so that the Euclidean norm of [E r] is minimal. Another interesting case that is described and solved in this correspondence assumes that errors occur in A but not b. We call this case the data least squares (DLS) problem because the error is assumed to lie in the data matrix A as indicated by

$$(A + E)x_{DLS} = b. \tag{4}$$

DeGroat, et. al. in [2] developed a close form solution to (4) and demonstrated that it outperformed OLS and TLS in case of noisy data matrix. However, the solution was a kind of batch type algorithm.

In this paper, we develop a recursive form of DLS solution based on the recursive generalized eigendecomposition method,

Corresponding author: Jun-seok Lim (jslim@sejong.ac.kr)
Kwangin-ku, Kunja-Dong 98, Sejong University, Department of Electronics Engineering, Seoul, Korea. 143-747

which was proposed by Y. N. Rao and J. C. Principe in [3]. The usefulness of the DLS method is then demonstrated with simulations of channel equalization.

II. Generalized Total Least Square Problem

Given an unknown system with finite impulse response and assuming that both the input and output are corrupted by the Gaussian white noise, the system should be estimated from the noisy observation of the input and output, as Fig.1. The unknown system is described by

$$\mathbf{h} = [h_0, h_1, \cdots, h_{N-1}]^H \in C^{N\times 1}, \tag{5}$$

where h may be time varying or time invariant. The output is given by

$$d(n) = \tilde{\mathbf{x}}^H(n)\mathbf{h} + n_o(n), \tag{6}$$

where the output noise $n_o(n)$ is a Gaussian white noise with variance σ_o^2 and independent of the input signal, and the noise free input vector is represented as

$$\mathbf{x}(n) = [x(n), x(n-1), \cdots x(n-N+1)]^T. \tag{7}$$

The noisy input vector of the system is given by

$$\tilde{\mathbf{x}}(n) = \mathbf{x}(n) + \mathbf{n}_i(n) \in C^{N\times 1}, \tag{8}$$

where $\mathbf{n}_i(n) = [n_i(n), n_i(n-1), \cdots n_i(n-N+1)]^T$ and the input noise ni(n) is the Gaussian white noise with variance σ_i^2.

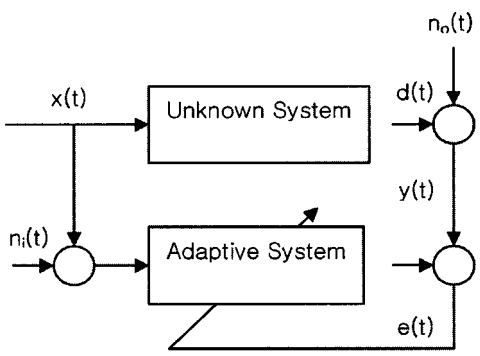

Fig. 1. The model of generalized total least square.

Notice that the input noise may originate from the measured error, interference, quantized noise and so on. Hence, we adopt a more general signal model than the least squares based estimation. Moreover, the augmented data vector is defined as

$$\bar{\mathbf{x}}(n) = [\tilde{\mathbf{x}}^T(n), d(n)]^T \in C^{(N+1)\times 1}. \tag{9}$$

The correlation matrix of the augmented data vector has the following structure

$$\bar{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{R}} & \mathbf{p} \\ \mathbf{p}^H & c \end{bmatrix}, \tag{10}$$

where $\mathbf{p} = E\{\tilde{\mathbf{x}}(n)d^*(n)\}$ and $c = E\{d(n)d^*(n)\}$, $\tilde{\mathbf{R}} = E\{\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}^H(n)\} = \mathbf{R} + \sigma_i^2\mathbf{I}$, $\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\}$. We can further establish that $\mathbf{p} = \mathbf{R}^H\mathbf{h}$ and $c = \mathbf{h}^H\tilde{\mathbf{R}}\mathbf{h} + \sigma_o^2$.

Define the constrained Rayleigh quotient as

$$J(\mathbf{w}) = \frac{[\mathbf{w}^T, -1]\bar{\mathbf{R}}[\mathbf{w}^T, -1]^H}{[\mathbf{w}^T, -1]\bar{\mathbf{D}}[\mathbf{w}^T, -1]^H}, \tag{11}$$

where $\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \gamma \end{bmatrix}$ with $\gamma = \frac{\sigma_o^2}{\sigma_i^2}$ [4]. The generalized total least square solution is obtained by solving

$$\min_{\mathbf{w}} J(\mathbf{w}). \tag{12}$$

DLS is a special case in (11) with γ) [4].

III. Generalized Eigendecomposition

From a mathematical perspective, generalized eigendecomposition (GED) involves solving the matrix equation $\mathbf{R}_1\tilde{\mathbf{W}} = \mathbf{R}_2\tilde{\mathbf{W}}\mathbf{\Lambda}$, where R1,R2 are square matrices, $\tilde{\mathbf{w}}$ is the generalized eigenvector matrix and Λ is the diagonal generalized eigenvalue matrix[5]. These are typically the full covariance matrices of zero mean stationary random signals $x_1(n)$ and $x_2(n)$ respectively. For complex symmetric and positive definite matrices, all the generalized eigenvectors are complex and the corresponding generalized eigenvalues are positive. The generalized eigenvectors act as filters in the joint space of the two signals $x_1(n)$ and $x_2(n)$, minimizing the energy of one of the

signals and maximizing the energy of the other at the same time. This property has been successfully applied to deriving a recursive algorithm [3]. The recursive algorithm has been based on the fact that any generalized eigenvector $\tilde{\mathbf{w}}$ that is a column of the matrix $\tilde{\mathbf{w}}$ is a stationary point of the function,

$$J(\tilde{\mathbf{w}}) = \frac{\tilde{\mathbf{w}}^H \mathbf{R}_1 \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \mathbf{R}_2 \tilde{\mathbf{w}}}. \tag{13}$$

This is because

$$\nabla_{\tilde{\mathbf{w}}} J(\tilde{\mathbf{w}}) = \nabla_{\tilde{\mathbf{w}}}\left(\frac{\tilde{\mathbf{w}}^H \mathbf{R}_1 \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \mathbf{R}_2 \tilde{\mathbf{w}}}\right) = 0 \Rightarrow \mathbf{R}_1 \tilde{\mathbf{w}} = \frac{\tilde{\mathbf{w}}^H \mathbf{R}_1 \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \mathbf{R}_2 \tilde{\mathbf{w}}} \mathbf{R}_2 \tilde{\mathbf{w}}. \tag{14}$$

This is nothing but the generalized eigenvalue equation and the generalized eigenvalues are the values of (13) evaluated at the stationary points. Rao, et. al. in[3] stated that the solution of (14) was the generalized eigenvector corresponding to the largest generalized eigenvalue so called the principal generalized eigenvector. The GED equation can be rewritten as,

$$\mathbf{R}_1 \tilde{\mathbf{w}} = \frac{\tilde{\mathbf{w}}^H \mathbf{R}_1 \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \mathbf{R}_2 \tilde{\mathbf{w}}} \mathbf{R}_2 \tilde{\mathbf{w}} \tag{15}$$

Left multiplying (15) by $\mathbf{R}_2^{-1}$ and rearranging the terms, we get,

$$\tilde{\mathbf{w}} = \frac{\tilde{\mathbf{w}}^H \mathbf{R}_2 \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \mathbf{R}_1 \tilde{\mathbf{w}}} \mathbf{R}_2^{-1} \mathbf{R}_1 \tilde{\mathbf{w}} \tag{16}$$

Equation (16) is the basis of our iterative algorithm. Let the weight vector $\tilde{\mathbf{w}}(n-1)$ at iteration (n 1) be the estimate of the principal generalized eigenvector. Then, the estimate of the new weight vector at iteration n according to (16) is,

$$\tilde{\mathbf{w}}(n) = \frac{\tilde{\mathbf{w}}^H(n-1)\mathbf{R}_2(n)\tilde{\mathbf{w}}(n-1)}{\tilde{\mathbf{w}}^H(n-1)\mathbf{R}_1(n)\tilde{\mathbf{w}}(n-1)} \mathbf{R}_2^{-1}(n)\mathbf{R}_1(n)\tilde{\mathbf{w}}(n-1) \tag{17}$$

where $\mathbf{R}_1(n) = \mathbf{R}_1(n-1) + \mathbf{x}_1(n)\mathbf{x}_1^H(n)$ and

$\mathbf{R}_2(n) = \mathbf{R}_2(n-1) + \mathbf{x}_2(n)\mathbf{x}_2^H(n)$

If we can observe that (16) and (17) track the GED equation at every time step, it is analogous to the RLS update rule that tracks the Wiener solution with every update[6]. As the update in [6], we need a matrix inversion operation for each update. Application of the Sherman Morrison Woodbury matrix inversion lemma in [5],

$$(\mathbf{A}+\mathbf{BC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I}+\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \tag{18}$$

to the $\mathbf{R}_2^{-1}(n)$ in (17) is straightforward and we obtain

$$\mathbf{R}_2^{-1}(n) = \left[\mathbf{R}_2^{-1}(n-1) - \frac{\mathbf{R}_2^{-1}(n-1)\mathbf{x}_2(n)\mathbf{x}_2^H(n)\mathbf{R}_2^{-1}(n-1)}{1+\mathbf{x}_2^H(n)\mathbf{R}_2^{-1}(n-1)\mathbf{x}_2(n)}\right] \tag{19}$$

Let define $y_1(n) = \tilde{\mathbf{w}}^H(n-1)\mathbf{x}_1(n)$ and $y_2(n) = \tilde{\mathbf{w}}^H(n-1)\mathbf{x}_2(n)$ as the outputs for signals x1(n) and x2(n) respectively, where the vectors x1(n) and x2(n) are built using the delayed versions of the signals x1(n) and x2(n). With this definition, we have

$$\tilde{\mathbf{w}}^H(n-1)\mathbf{R}_1(n)\tilde{\mathbf{w}}(n-1) \approx \sum_{i=1}^{n} |y_1(i)|^2 \tag{20}$$

$$\tilde{\mathbf{w}}^H(n-1)\mathbf{R}_2(n)\tilde{\mathbf{w}}(n-1) \approx \sum_{i=1}^{n} |y_2(i)|^2 \tag{21}$$

This is true in the stationary cases when sample variance estimators can be used instead of the expectation operators. However, for non stationary signals, a simple forgetting factor can be included with a trivial change in the update equation. With these simplifications, we can write the modified update equation for the stationary case as,

$$\begin{aligned} \tilde{\mathbf{w}}(n) &= \frac{\sum_{i=1}^{n}|y_2(i)|^2}{\sum_{i=1}^{n}|y_1(i)|^2} \mathbf{R}_2^{-1}(n)\sum_{i=1}^{n}\mathbf{x}_1(i)y_1(i) \\ y_l(i) &= \tilde{\mathbf{w}}^H(i-1)\mathbf{x}_l(i), \quad l=1,2 \end{aligned} \tag{22}$$

where $\mathbf{R}_2^{-1}(n)$ is estimated by using (19).

IV. Recursive Data Least Square Algorithm

We can apply the generalized eigendecomposition method in section III to solution of DLS. If we modify (11) and (12), the object function for DLS becomes as follows.

$$\tilde{J}(\mathbf{w}) = \frac{\tilde{\mathbf{w}}^H \bar{\mathbf{D}} \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \bar{\mathbf{R}} \tilde{\mathbf{w}}} = \frac{[\mathbf{w}^H, -1]\bar{\mathbf{D}}[\mathbf{w}^T, -1]^T}{[\mathbf{w}^H, -1]\bar{\mathbf{R}}[\mathbf{w}^T, -1]^T}, \tag{23}$$

The DSL solution can be derive as (24). We apply the recursive algorithm in section III for the maximization of (24).

$$\max_{\tilde{\mathbf{w}}} \tilde{J}(\tilde{\mathbf{w}}), \text{ and then } \mathbf{w} = \tilde{\mathbf{w}}(1:N)/(-\tilde{\mathbf{w}}(N+1)), \tag{24}$$

where $\tilde{\mathbf{w}}(1:N)$ is a vector with the elements from the 1 st to the N th, and $\tilde{\mathbf{w}}(N+1)$ is the (N+1) th element in $\tilde{\mathbf{w}}$. When we apply the generalized eigendecomposition to (23), we have two simplified equations for (20) and (22), respectively. (20) becomes as follows.

$$\tilde{\mathbf{w}}^H(n-1)\mathbf{R}_1(n)\tilde{\mathbf{w}}(n-1) = \tilde{\mathbf{w}}^H(n-1)\overline{\mathbf{D}}(n)\tilde{\mathbf{w}}(n-1) = \mathbf{w}^H(n-1)\mathbf{w}(n-1) \tag{25}$$

And (22) also becomes as follows.

$$\tilde{\mathbf{w}}(n) = \frac{\sum_{i=1}^{n}|y(i)|^2}{\mathbf{w}^H(n-1)\mathbf{w}(n-1)}\mathbf{R}_2^{-1}(n)\overline{\mathbf{D}}\tilde{\mathbf{w}}(n-1), \text{ where } \quad y(i) = \tilde{\mathbf{w}}^H(i-1)\overline{\mathbf{x}}(i) \tag{26}$$

In DLS, the element in (N+1,N+1) becomes null. However, $\overline{\mathbf{D}}$ is not inverted and the null element makes the denominator in (26) the energy of vector, w, so that the null element does not cause any numerical instability.

We summarize the algorithm in table 1.

Table 1. Recursive Data Least Square (RDLS) Algorithm.

1. Initialize $\overline{\mathbf{x}}(0) = [\mathbf{x}^T(0), d(0)]$, $\tilde{\mathbf{w}}(0) = [\mathbf{w}^T(0), -1]$ with the w(0) $\in$ CN$\times$1 to a random vector
2. Fill the matrix Q(0) $\in$ CN$\times$N with small random values
3. Initialize scalar variables C(0) to zero For j 〉 0
4. Compute $y(j) = \tilde{\mathbf{w}}^H(j-1)\overline{\mathbf{x}}(j)$
5. Update Q as
$$\mathbf{Q}(n) = \left[\mathbf{Q}(n-1) - \frac{\mathbf{Q}(n-1)\overline{\mathbf{x}}(n)\overline{\mathbf{x}}^H(n)\mathbf{Q}(n-1)}{1+\overline{\mathbf{x}}^H(n)\mathbf{Q}(n-1)\overline{\mathbf{x}}(n)}\right]$$
6. Update C as $C(j) = C(j-1) + |y(j)|^2$
7. Update the weight vector as
$\tilde{\mathbf{w}}(j) = [C(j)/(\mathbf{w}^H(j-1)\mathbf{w}(j-1))]\mathbf{Q}(j)\overline{\mathbf{D}}\tilde{\mathbf{w}}(j-1)$
8. Normalize the weight vector
9. $\mathbf{w}(j) = \tilde{\mathbf{w}}(1:n-1)/(-\tilde{\mathbf{w}}(n+1))$ loop

V. A Channel Equalization Application

In this section, we demonstrate the usefulness of the DLS method by comparing it with the optimal method and OLS methods when applied to a channel equalization problem. The channel equalization problem is graphically described by the block diagram in Fig. 2. Basically, the solution vector, w = [w1, w2, . . . , wp]T represents an FIR approximant inverse filter to the channel characteristic H(z). The output of the inverse (equalization) filter can be written in matrix form using the output of the channel as input to the finite impulse response (FIR) equalization filter. The output of the equalized channel should be approximately equal to the original input

$$\begin{bmatrix} \tilde{s}_{p-1} \\ \tilde{s}_p \\ \vdots \\ \tilde{s}_{N-1} \end{bmatrix} = \begin{bmatrix} v_{p-1} & \cdots & v_1 & v_0 \\ v_p & \cdots & v_2 & v_1 \\ \vdots & \vdots & \vdots & \vdots \\ v_{N-1} & \cdots & v_{N-p+1} & v_{N-p} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} \approx \begin{bmatrix} s_{p-1} \\ s_p \\ \vdots \\ s_{N-1} \end{bmatrix}, \tag{27}$$

where p is the FIR filter order; and N is the total number of output samples. In this problem, we assume that the left side in (27) is known without error because the input training signal is assumed to be known without error. It is easy to see that (27) has the form of (4).

For the simulation, a well known complex nonminimum phase channel model introduced by Cha and Kassam [9] is used to evaluate the proposed recursive data least square (RDLS) equalizer performance for 4 PAM signaling. Although the length of the channel is short, the channel model cannot only simulate the phase change from boundary reflection but also do the nonminimum phase characteristics of the channel in the room acoustics or in the underwater communication. The channel output v(n) (which is also the input of the equalizer) is given by

$$\begin{aligned} v(n) &= (0.34 - j0.27)s(n) + (0.87 + j0.43)s(n-1) \\ &\quad + (0.34 - j0.21)s(n-2) + \eta(n), \quad \eta(n) \sim N(0,0.01) \end{aligned} \tag{28}$$

Where N(0, 0.01) means the white Gaussian noise (of the nonminimum-phase channel) with mean 0 and variance 0.01. 4-PAM symbol sequence $s(n)$ is passed through the channel and the sequence $s(n)$ are valued from the set {±1, ±3}. All the equalizers, the recursive least square (RLS) based equalizer and

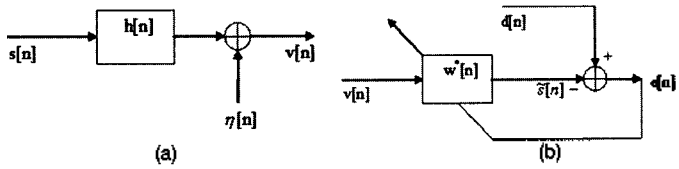

Fig. 2. Transmission and Equalization model: (a) received signal model, (b) equalizer model (s[n]:transmitted signal, h[n]: channel model, η [n]: additive noise, v[n]:received signal, d[n]:training signal).

the RDLS based equalizer, are trained with 1000 data symbols at 16 dB SNR. The RLS is a recursive algorithm for the OLS problem. The order of equalizer is set to 9.

Fig. 3 (a) shows the distribution of the input data of the different equalizers. This figure shows received signals scattered severely due to transmission channel effect. Fig. 3 (b), (c) and (d) show the scatter diagrams of the outputs of the three equalizers, optimal, RLS and RDLS, respectively. As observed from Fig. 3, the equalized signal by the proposed algorithm centres on {±1, ±3} and it is almost the same as the equalized signals by the optimal equalizer which is derived from the Wiener solution. It leads the conclusion that the proposed RDLS outperforms the RLS algorithm. Moreover, it estimates almost the same as optimal equalizer.

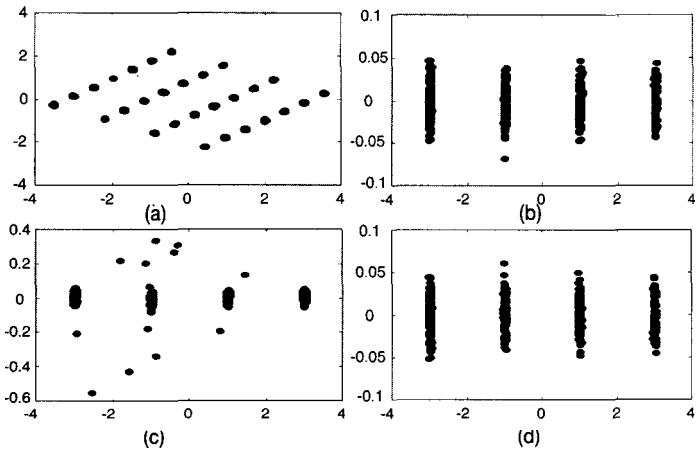

Fig. 3. Performance comparison of three equalizers: (a) scatter diagram of received signals, (b) scatter diagram of optimal equalizer, (c) scatter diagram of RLS equalizer, (d) scatter diagram of the proposed equalizer.

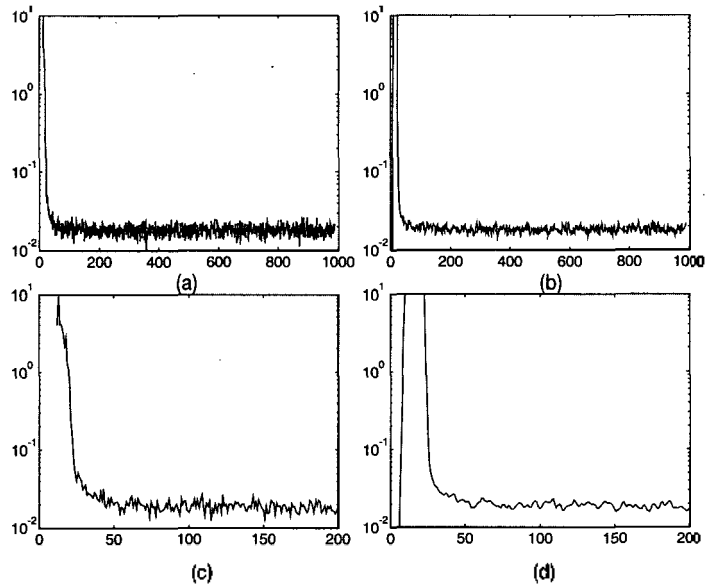

Fig. 4. Convergence comparison: (a) RLS equalizer, (b) the proposed equalizer.

Fig. 4 compares the proposed algorithm with the RLS in respect of the convergence. Fig. 4 shows that the proposed algorithm converges almost the same as RLS. Fig. 4 (c) and (d) provide more better comparison for convergence. Furthermore, the proposed algorithm shows less fluctuation that RLS. That can explain the results in Fig. 3 (c) and in Fig. 3 (d). RLS based equalizer can equalize almost the same as the optimal equalizer in Fig. 3 (b) but some samples are severely scattered far from the symbol centres on {±1, ±3}. It frequently happens in the nonminimum channel. On the contrary, the proposed algorithm concentrates the output symbols on the symbol centres, {±1, ±3}.

VI. Conclusion

In this paper, we proposed a recursive algorithm for data least square (DLS) solution. Channel equalization simulations were performed to compare the proposed algorithm with the algorithms in OLS and we found better performance over OLS methods.

Acknowledgment

This result was supported by ADD in Korea.

References

1. G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," SIAMJ. Numer. Anal., **17**, 883-893, 1980.
2. R.D. DeGroat and E. M. Dowling, "The Data Least Squares and Channel Equalization," IEEE Trans. Signal Processing, **41**, 407-411, 1993.
3. Y. N. Rao and J. C. Principe, "Fast RLS-Like Algorithm for Generalized Eigendecomposition and its Applications," Journal of VLSI Signal Processing, **37**, 333-344, 2004
4. C.E. Davila, "Line Search Algorithm for Adaptive Filtering," IEEE Trans. Signal Processing, **42**, 2490-2494, 1994.
5. G.H. Golub and C.F. Van Loan, Matrix Computations, (The John Hopkins University Press, 1991).
6. S. Haykin, *Adaptive Filter Theory*, (Englewood Cliffs, NJ: Prentice -Hall, c1986).
7. J.C. Principe, N. Euliano, and C. Lefebvre, *Neural systems: Fundamentals through Simulations*, (Wiley, 1999).
8. Y.N. Rao and J.C. Principe, "A Fast On-Line Algorithm for PCA and its Convergence Characteristics," Proc. IEEEWorkshop on Neural Networks for Signal Processing X, 299-308, 2000.
9. Cha, I., Kassam, S.A., "Channel equalization using adaptive complex radial basis function networks," IEEE J. Sel. Area. Comm., 13, 122-131, 1995.

[Profile]

•Jun-seok Lim

1986: B.S. degree in Electronics Engineering, Seoul National University
1988: M.S. degree in Electronics Engineering, Seoul National University
1996: Ph.D. degree in Electronics Engineering, Seoul National University
1988~1993: Research Scientist in Agency of Defense Development
1998~present: Associate Professor, Department of Electronics Engineering, Sejong University

•Jae Soo Kim

1981: B.S. degree in Naval Architecture and Ocean Engineering, Seoul National University
1984: M.S. degree in Ocean Engineering, Univ. of Florida
1989: Ph.D. degree in Underwater Acoustics, MIT
1989~1989: Post-Doctoral Associate in MIT
1990~1991: Senior Research Scientist in Agency of Defense Development
1991~present: Professor, Division of Ocean System Eng., Korea Maritime University