

# A QUASI-NEWTON ADAPTIVE ALGORITHM FOR ESTIMATING GENERALIZED EIGENVECTORS

G. Mathew<sup>1</sup>      V.U. Reddy<sup>1</sup>      A. Paulraj<sup>2</sup>

<sup>1</sup>Dept. of Electrical Communication Engineering  
Indian Institute of Science, Bangalore- 560012, India  
e-mail: vur@ece.iisc.ernet.in

<sup>2</sup>Information Systems Laboratory, Dept. of Electrical Engineering  
Stanford University, Stanford, CA 94305  
Tel: 415-725-8307 Fax: 415-723-8473

## ABSTRACT

We first introduce a constrained minimization formulation for the generalized symmetric eigenvalue problem and then recast it into an unconstrained minimization problem by constructing an appropriate cost function. Minimizer of this cost function corresponds to the eigenvector corresponding to the minimum eigenvalue of the given symmetric matrix pencil and all minimizers are global minimizers. We also present an inflation technique for obtaining multiple generalized eigenvectors of this pencil. Based on this asymptotic formulation, we derive a quasi-Newton adaptive algorithm for estimating these eigenvectors in the data case. This algorithm is highly modular and parallel with a computational complexity of  $\mathcal{O}(N^2)$  multiplications,  $N$  being the problem-size. Simulation results show fast convergence and good quality of the estimated eigenvectors.

## 1. INTRODUCTION

Consider the matrix pencil  $(\mathbf{R}_y, \mathbf{R}_w)$  where  $\mathbf{R}_y$  and  $\mathbf{R}_w$  are  $N \times N$  symmetric positive definite matrices. Then, the task of computing an  $N \times 1$  vector  $\mathbf{x}$  and a scalar  $\lambda$  such that

$$\mathbf{R}_y \mathbf{x} = \lambda \mathbf{R}_w \mathbf{x} \quad (1)$$

is called the generalized symmetric eigenvalue problem. The solution vector  $\mathbf{x}$  and scalar  $\lambda$  are called the generalized eigenvector and eigenvalue, respectively, of the pencil  $(\mathbf{R}_y, \mathbf{R}_w)$ . This pencil has  $N$  positive eigenvalues viz.  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ , and corresponding real  $\mathbf{R}_w$ -orthonormal eigenvectors  $\mathbf{q}_j$ ,  $j = 1, \dots, N$  [9,1]:

$$\mathbf{R}_y \mathbf{q}_i = \lambda_i \mathbf{R}_w \mathbf{q}_i \quad \text{with} \quad \mathbf{q}_i^T \mathbf{R}_w \mathbf{q}_j = \delta_{ij} \\ i, j \in \{1, \dots, N\} \quad (2)$$

where  $\delta_{ij}$  is the Kronecker delta function.

The generalized symmetric eigenvalue problem has extensive applications in the harmonic retrieval / direction-of-arrival estimation problem when the observed data  $\mathbf{y}(n)$  is the sum of a desired signal  $x(n)$  and a coloured noise which is of the form  $\sigma w(n)$ , where  $\sigma$  is a scalar. For example, consider  $x(n)$  to be the sum of  $P$  real sinusoids and assume  $w(n)$  to be uncorrelated with  $x(n)$ . Then, if  $\mathbf{R}_y$  and  $\mathbf{R}_w$  represent asymptotic covariance matrices (of size  $N \times N$  with  $N \geq 2P+1$ ) of  $\mathbf{y}(n)$  and  $w(n)$ , respectively, with  $\mathbf{R}_w$  assumed to be non-singular, the sinusoidal frequencies

can be obtained from the eigenvector corresponding to the minimum eigenvalue of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

Many researchers have addressed the problem (1) for given  $\mathbf{R}_y$  and  $\mathbf{R}_w$  and proposed methods for solving it. Moler and Stewart [2] proposed a QZ-algorithm and Kaufman [3] proposed an LZ-algorithm for iteratively solving (1). But, their methods do not exploit the structure in  $\mathbf{R}_y$  and  $\mathbf{R}_w$ . On the other hand, the solution of (1) in the special case of symmetric  $\mathbf{R}_y$  and symmetric positive definite  $\mathbf{R}_w$  has been a subject of special interest and several efficient approaches have been reported. By using the Cholesky factorization of  $\mathbf{R}_w$ , this problem can be reduced to the standard eigenvalue problem as reported by Martin and Wilkinson [4]. Bunse-Gerstner [5] proposed an approach using congruence transformations for the simultaneous diagonalization of  $\mathbf{R}_y$  and  $\mathbf{R}_w$ . Shougen and Shuqin [6] reported an algorithm which makes use of Cholesky, QR and Singular value decompositions when  $\mathbf{R}_y$  is also positive definite. This algorithm is stable, faster than QZ-algorithm, and much superior to that of [4]. Auchmuty [1] proposed and analyzed certain cost functions which are minimized at the eigenvectors corresponding to some specific eigenvalues.

In the case of adaptive signal processing applications, however,  $\mathbf{R}_y$  and  $\mathbf{R}_w$  correspond to asymptotic covariance matrices and they need to be estimated. In this paper, the data covariance matrix of  $\mathbf{y}(n)$ , i.e.,  $\mathbf{R}_y(n)$ , is taken as the estimate of  $\mathbf{R}_y$  and an estimate of  $\mathbf{R}_w$ , say  $\hat{\mathbf{R}}_w$ , is assumed to be available based on some *a priori* measurements. The sample covariance matrix at  $n^{\text{th}}$  data instant is calculated as

$$\mathbf{R}_y(n) = \frac{1}{n} \sum_{i=1}^n \mathbf{y}(i) \mathbf{y}^T(i)$$

where  $\mathbf{y}(n)$  is the data vector containing the most recent  $N$  samples of  $y(n)$ . Thus, the problem we address in this paper is as follows. Given the time-series  $\mathbf{y}(n)$ , develop an adaptive algorithm for estimating the first  $D$  ( $D \leq N$ ) generalized  $\mathbf{R}_w$ -orthogonal eigenvectors of the pencil  $(\mathbf{R}_y, \mathbf{R}_w)$  using  $(\mathbf{R}_y(n), \hat{\mathbf{R}}_w)$  as an estimate of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

The paper is organized as follows. The generalized eigenvalue problem is translated into an unconstrained minimization problem in Section 2. A quasi-Newton adaptive algorithm is derived in Section 3. Computer simulations are presented in Section 4 and finally, Section 5 concludes the paper.

## 2. A MINIMIZATION FORMULATION FOR GENERALIZED SYMMETRIC EIGENVALUE PROBLEM

In this section, we recast the generalized symmetric eigenvalue problem into an unconstrained minimization framework. We first consider the case of the eigenvector corresponding to the minimum eigenvalue (henceforth referred to as the minimum eigenvector) and then extend it to the case of more than one eigenvector using an inflation approach. The principle used here is a generalization of the approach followed in [8].

### 2.1. Single Eigenvector Case

We first introduce a constrained minimization formulation for seeking the minimum eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ . Consider the following problem:

$$\min_{\mathbf{a}} \mathbf{a}^T \mathbf{R}_y \mathbf{a} \quad \text{subject to} \quad \mathbf{a}^T \mathbf{R}_w \mathbf{a} = 1. \quad (3)$$

Let  $\mathbf{R}_w = \mathbf{L}_w \mathbf{L}_w^T$  be the Cholesky factorization of  $\mathbf{R}_w$ . Then, (3) can be re-written as

$$\min_{\mathbf{b}} \mathbf{b}^T \tilde{\mathbf{R}}_y \mathbf{b} \quad \text{subject to} \quad \mathbf{b}^T \mathbf{b} = 1 \quad (4)$$

where  $\tilde{\mathbf{R}}_y = \mathbf{L}_w^{-1} \mathbf{R}_y \mathbf{L}_w^{-T}$  and  $\mathbf{b} = \mathbf{L}_w^T \mathbf{a}$ . Since the eigenvalues of  $(\mathbf{R}_y, \mathbf{R}_w)$  and  $\tilde{\mathbf{R}}_y$  are identical and the eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$  are the eigenvectors of  $\tilde{\mathbf{R}}_y$  premultiplied by  $\mathbf{L}_w^{-T}$ , it follows from (4) that the solution of (3) is the minimum eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

Using the penalty function method [7], the constrained problem (3) can be translated into unconstrained minimization of the cost function

$$J(\mathbf{a}, \mu) = \frac{\mathbf{a}^T \mathbf{R}_y \mathbf{a}}{2} + \frac{\mu(\mathbf{a}^T \mathbf{R}_w \mathbf{a} - 1)^2}{4} \quad (5)$$

where  $\mu$  is a positive scalar. Below, we give results to establish the correspondence between the minimizer of  $J(\mathbf{a}, \mu)$  and a minimum eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

The gradient vector  $\mathbf{g}$  and Hessian matrix  $\mathbf{H}$  of  $J(\mathbf{a}, \mu)$  with respect to  $\mathbf{a}$  are

$$\mathbf{g}(\mathbf{a}, \mu) = \mathbf{R}_y \mathbf{a} + \mu(\mathbf{a}^T \mathbf{R}_w \mathbf{a} - 1) \mathbf{R}_w \mathbf{a}, \quad (6)$$

$$\mathbf{H}(\mathbf{a}, \mu) = \mathbf{R}_y + 2\mu(\mathbf{R}_w \mathbf{a})(\mathbf{R}_w \mathbf{a})^T + \mu(\mathbf{a}^T \mathbf{R}_w \mathbf{a} - 1) \mathbf{R}_w. \quad (7)$$

Using (6) and (7), we get the following results.

**Theorem 1**  $\mathbf{a}^*$  is a stationary point of  $J(\mathbf{a}, \mu)$  if and only if  $\mathbf{a}^*$  is an eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$  corresponding to the eigenvalue  $\lambda$  such that  $\lambda = \mu(1 - \mathbf{a}^{*T} \mathbf{R}_w \mathbf{a}^*)$ .

This result immediately follows from (6).

**Theorem 2**  $\mathbf{a}^*$  is a global minimizer of  $J(\mathbf{a}, \mu)$  if and only if  $\mathbf{a}^*$  is a minimum eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$  with eigenvalue  $\lambda_{\min} = \mu(1 - \mathbf{a}^{*T} \mathbf{R}_w \mathbf{a}^*)$ , where  $\lambda_{\min} = \lambda_1$ . Further, all minimizers are global minimizers.

### Proof

If part: Since  $\mathbf{R}_y \mathbf{a}^* = \lambda_{\min} \mathbf{R}_w \mathbf{a}^*$  with  $\lambda_{\min} = \mu(1 - \mathbf{a}^{*T} \mathbf{R}_w \mathbf{a}^*)$ , we get from Theorem 1 that  $\mathbf{a}^*$  is a stationary point of  $J(\mathbf{a}, \mu)$ . Let  $\mathbf{y} = \mathbf{a}^* + \mathbf{p}$ ,  $\mathbf{p} \in \mathbb{R}^N$ . Then, it can be shown that

$$J(\mathbf{y}, \mu) - J(\mathbf{a}^*, \mu) = \frac{\mathbf{p}^T \mathbf{R}_y \mathbf{p}}{2} - \frac{\lambda_{\min}}{2} \mathbf{p}^T \mathbf{R}_w \mathbf{p} + \frac{\mu}{4} [2\mathbf{p}^T \mathbf{R}_w \mathbf{a}^* + \mathbf{p}^T \mathbf{R}_w \mathbf{p}]^2. \quad (8)$$

Expressing  $\mathbf{p}$  as  $\mathbf{p} = \sum_{i=1}^N \alpha_i \mathbf{q}_i$ ,  $\alpha_i \in \mathbb{R}$ , and using (2), we obtain

$$J(\mathbf{y}, \mu) - J(\mathbf{a}^*, \mu) \geq \frac{\mu}{4} [2\mathbf{p}^T \mathbf{R}_w \mathbf{a}^* + \mathbf{p}^T \mathbf{R}_w \mathbf{p}]^2 \geq 0$$

$\forall \mathbf{p} \in \mathbb{R}^N$ . Thus,  $\mathbf{a}^*$  is a global minimizer of  $J(\mathbf{a}, \mu)$ .

Only if part: Since  $\mathbf{a}^*$  is a global minimizer, we have i)  $\mathbf{a}^*$  is a stationary point of  $J(\mathbf{a}, \mu)$  and ii)  $\mathbf{H}(\mathbf{a}^*, \mu)$  is positive semi-definite. Hence, from Theorem 1, we get  $\mathbf{R}_y \mathbf{a}^* = \lambda_m \mathbf{R}_w \mathbf{a}^*$  with  $\lambda_m = \mu(1 - \mathbf{a}^{*T} \mathbf{R}_w \mathbf{a}^*)$ ,  $\mathbf{a}^* = \beta \mathbf{q}_m$  and  $\beta = \sqrt{\mathbf{a}^{*T} \mathbf{R}_w \mathbf{a}^*}$  for some  $m \in \{1, \dots, N\}$ . Then, with  $\mathbf{p} = \sum_{i=1}^N \alpha_i \mathbf{q}_i$ , it can be shown that

$$\mathbf{p}^T \mathbf{H}(\mathbf{a}^*, \mu) \mathbf{p} = \sum_{i=1}^N (\lambda_i - \lambda_m + 2\mu\beta^2 \delta_{mi}) \alpha_i^2. \quad (9)$$

Since  $\mathbf{H}(\mathbf{a}^*, \mu)$  is positive semi-definite, (9) implies that  $\lambda_i \geq \lambda_m$  for all  $i = 1, \dots, N$ . That is,  $\mathbf{a}^*$  is the minimum eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ . Further, since  $\mathbf{H}(\mathbf{a}^*, \mu)$  is positive semi-definite even if  $\mathbf{a}^*$  is a local minimizer, it follows that all minimizers are global minimizers.

**Corollary 1** The value of  $\mu$  should be greater than  $\lambda_{\min}$ .

**Corollary 2** The eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$  associated with the non-minimum eigenvalues correspond to saddle points of  $J(\mathbf{a}, \mu)$ .

Clearly, minimizer of  $J(\mathbf{a}, \mu)$  will be unique (except for the sign) only if the minimum eigenvalue of  $(\mathbf{R}_y, \mathbf{R}_w)$  is simple. The main result of the above analysis is that all minimizers of  $J(\mathbf{a}, \mu)$  are aligned with the minimum eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$ . Since all minimizers are global minimizers, any simple search technique will suffice to reach the correct solution. Further, since  $\frac{\mathbf{y}^T \mathbf{R}_y \mathbf{y}}{\mathbf{y}^T \mathbf{R}_w \mathbf{y}} \geq \lambda_{\min} \forall \mathbf{y} \in \mathbb{R}^N$ , the bound for  $\mu$  in Corollary 1 can be satisfied by choosing  $\mu > \frac{\mathbf{R}_y(1,1)}{\mathbf{R}_w(1,1)}$ .

### 2.2. Inflation Technique for Seeking Multiple Eigenvectors

Now, we present an inflation technique which, combined with the result of the previous section, is used to seek the first  $D$  ( $D \leq N$ )  $\mathbf{R}_w$ -orthogonal eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$ . The basic approach is to divide the problem into  $D$  sub-problems such that the  $k^{\text{th}}$  sub-problem solves for the  $k^{\text{th}}$  eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

Let  $\mathbf{a}_i^*$ ,  $i = 1, \dots, k-1$ , (with  $2 \leq k \leq D$ ) be the  $\mathbf{R}_w$ -orthogonal eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$  corresponding to the eigenvalues  $\lambda_i = \mu(1 - \mathbf{a}_i^{*T} \mathbf{R}_w \mathbf{a}_i^*)$  for  $i = 1, \dots, k-1$ . To

obtain the next  $\mathbf{R}_w$ -orthogonal eigenvector  $\mathbf{a}_k^*$ , consider the cost function

$$\begin{aligned} J_k(\mathbf{a}_k, \mu, \alpha) &= \frac{\mathbf{a}_k^T \mathbf{R}_y \mathbf{a}_k}{2} + \frac{\mu}{4} (\mathbf{a}_k^T \mathbf{R}_w \mathbf{a}_k - 1)^2 \\ &\quad + \frac{\alpha}{2} \sum_{i=1}^{k-1} (\mathbf{a}_k^T \mathbf{R}_w \mathbf{a}_i^*)^2 \\ &= \frac{\mathbf{a}_k^T \mathbf{R}_{y_k} \mathbf{a}_k}{2} + \frac{\mu}{4} (\mathbf{a}_k^T \mathbf{R}_w \mathbf{a}_k - 1)^2 \end{aligned} \quad (10)$$

where  $\alpha$  is a positive scalar and for  $k \geq 2$

$$\mathbf{R}_{y_k} = \mathbf{R}_y + \alpha \sum_{i=1}^{k-1} (\mathbf{R}_w \mathbf{a}_i^*) (\mathbf{R}_w \mathbf{a}_i^*)^T \quad (11)$$

$$= \mathbf{R}_{y_{k-1}} + \alpha (\mathbf{R}_w \mathbf{a}_{k-1}^*) (\mathbf{R}_w \mathbf{a}_{k-1}^*)^T \quad (12)$$

with  $\mathbf{R}_{y_1} = \mathbf{R}_y$ . Eqn.(12) represents the inflation step and its implication is discussed below.

Let  $\mathbf{a}_i^* = \beta_i \mathbf{q}_i$ ,  $\beta_i > 0$ ,  $i = 1, \dots, k-1$ . Post-multiplying (11) with  $\mathbf{q}_j$ , we get

$$\begin{aligned} \mathbf{R}_{y_k} \mathbf{q}_j &= (\lambda_j + \alpha \beta_j^2) \mathbf{R}_w \mathbf{q}_j \quad j = 1, \dots, k-1 \\ &= \lambda_j \mathbf{R}_w \mathbf{q}_j \quad j = k, \dots, N. \end{aligned}$$

That is,  $(\mathbf{R}_y, \mathbf{R}_w)$  and  $(\mathbf{R}_{y_k}, \mathbf{R}_w)$  have the same set of eigenvectors but  $k-1$  different eigenvalues. Now, choose  $\alpha$  such that

$$\lambda_k < \lambda_j + \alpha \beta_j^2 \quad j = 1, \dots, k-1. \quad (13)$$

Then, clearly, the minimum eigenvalue of  $(\mathbf{R}_{y_k}, \mathbf{R}_w)$  is  $\lambda_k$  with  $\mathbf{q}_k$  as the corresponding eigenvector. Hence, if  $\mathbf{a}_k^*$  is a minimizer of  $J_k(\mathbf{a}_k, \mu, \alpha)$ , then it follows from Theorem 2 that  $\mathbf{a}_k^*$  is a minimum eigenvector of  $(\mathbf{R}_{y_k}, \mathbf{R}_w)$  with eigenvalue  $\lambda_k = \mu(1 - \mathbf{a}_k^{*T} \mathbf{R}_w \mathbf{a}_k^*)$ . By construction,  $\mathbf{a}_k^*$  is nothing but the  $k^{\text{th}}$   $\mathbf{R}_w$ -orthogonal eigenvector of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

A practical lower bound for  $\alpha$  is  $\mu$ . Thus, by constructing  $D$  cost functions  $J_k(\mathbf{a}_k, \mu, \alpha)$ ,  $k = 1, \dots, D$ , as in (10) and finding their minimizers, we get the first  $D$   $\mathbf{R}_w$ -orthogonal eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$ .

### 3. A QUASI-NEWTON-CUM-INFLATION BASED ADAPTIVE ALGORITHM

In this section, we combine the inflation technique of the previous section with a quasi-Newton method and derive an adaptive algorithm for estimating the first  $D$   $\mathbf{R}_w$ -orthogonal eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$  in the data case.

Let  $\mathbf{a}_k(n)$ ,  $k = 1, \dots, D$ , be the estimates of these vectors at  $n^{\text{th}}$  adaptation instant. Newton algorithm for updating  $\mathbf{a}_k(n-1)$  to  $\mathbf{a}_k(n)$  is of the form

$$\mathbf{a}_k(n) = \mathbf{a}_k(n-1) - \mathbf{H}_k^{-1}(n-1) \mathbf{g}_k(n-1) \quad (14)$$

where  $\mathbf{H}_k(n-1)$  and  $\mathbf{g}_k(n-1)$  are the Hessian and gradient, respectively, of  $J_k(\mathbf{a}_k, \mu, \alpha)$  evaluated at  $\mathbf{a}_k = \mathbf{a}_k(n-1)$ . Since  $\mathbf{R}_y$ ,  $\mathbf{R}_w$  and  $\mathbf{a}_i^*$ ,  $i = 1, \dots, k-1$ , are not available, estimates of  $\mathbf{g}_k(n-1)$  and  $\mathbf{H}_k(n-1)$  can be obtained by

replacing  $\mathbf{R}_y$ ,  $\mathbf{R}_w$  and  $\mathbf{a}_i^*$ ,  $i = 1, \dots, k-1$ , with  $\mathbf{R}_y(n)$ ,  $\hat{\mathbf{R}}_w$  and  $\mathbf{a}_i(n)$ ,  $i = 1, \dots, k-1$ , respectively. Thus, we obtain

$$\hat{\mathbf{g}}_k(n-1) = \mu [\mathbf{a}_k^T(n-1) \mathbf{c}_k(n-1) - 1] \mathbf{c}_k(n-1) + \mathbf{R}_{y_k}(n) \mathbf{a}_k(n-1) \quad (15)$$

$$\hat{\mathbf{H}}_k(n-1) = \mathbf{R}_{y_k}(n) + 2\mu \mathbf{c}_k(n-1) \mathbf{c}_k^T(n-1) + \mu [\mathbf{a}_k^T(n-1) \mathbf{c}_k(n-1) - 1] \hat{\mathbf{R}}_w \quad (16)$$

where  $\mathbf{c}_k(n-1) = \hat{\mathbf{R}}_w \mathbf{a}_k(n-1)$  and

$$\mathbf{R}_{y_k}(n) = \mathbf{R}_{y_{k-1}}(n) + \alpha (\hat{\mathbf{R}}_w \mathbf{a}_{k-1}(n)) (\hat{\mathbf{R}}_w \mathbf{a}_{k-1}(n))^T \quad k = 2, \dots, D \quad (17)$$

with  $\mathbf{R}_{y_1}(n) = \mathbf{R}_y(n)$ .

Further, we approximate the Hessian by dropping the last term in (16) so that the approximant is positive definite and a recursion can be obtained directly in terms of its inverse. Substituting this approximated Hessian along with (15) in (14), we obtain the quasi-Newton adaptive algorithm for estimating the eigenvectors of  $(\mathbf{R}_y, \mathbf{R}_w)$  (after some manipulations) as

$$\mathbf{a}_k(n) = l_k(n-1) \mathbf{R}_{y_k}^{-1}(n) \mathbf{c}_k(n-1) \quad k = 1, \dots, D \quad (18)$$

$$l_k(n-1) = \mu \frac{1 + \mathbf{a}_k^T(n-1) \mathbf{c}_k(n-1)}{1 + 2\mu (\mathbf{c}_k^T(n-1) \mathbf{R}_{y_k}^{-1}(n) \mathbf{c}_k(n-1))} \quad k = 1, \dots, D \quad (19)$$

$$\begin{aligned} \mathbf{R}_{y_k}^{-1}(n) &= \mathbf{R}_{y_{k-1}}^{-1}(n) \\ &\quad - \frac{\mathbf{R}_{y_{k-1}}^{-1}(n) \mathbf{c}_{k-1}(n) \mathbf{c}_{k-1}^T(n) \mathbf{R}_{y_{k-1}}^{-1}(n)}{\frac{1}{\alpha} + \mathbf{c}_{k-1}^T(n) \mathbf{R}_{y_{k-1}}^{-1}(n) \mathbf{c}_{k-1}(n)} \end{aligned} \quad k \geq 2 \quad (20)$$

$$\begin{aligned} \mathbf{R}_{y_1}^{-1}(n) &= \mathbf{R}_y^{-1}(n) = \frac{n}{n-1} [\mathbf{R}_y^{-1}(n-1) \\ &\quad - \frac{\mathbf{R}_y^{-1}(n-1) \mathbf{y}(n) \mathbf{y}^T(n) \mathbf{R}_y^{-1}(n-1)}{n-1 + \mathbf{y}^T(n) \mathbf{R}_y^{-1}(n-1) \mathbf{y}(n)}] \end{aligned} \quad n \geq 2 \quad (21)$$

where

$$\mathbf{c}_k(n) = \hat{\mathbf{R}}_w \mathbf{a}_k(n).$$

This algorithm can be implemented using a pipeline architecture as illustrated in Fig. 1. Here,  $k^{\text{th}}$  eigenvector is estimated by the  $k^{\text{th}}$  unit and it goes through the following steps during  $n^{\text{th}}$  sample interval (time indices are suppressed in the figure):

- i) pass on  $\mathbf{R}_{y_k}^{-1}(n-k)$  and  $\mathbf{a}_k(n-k)$  to  $(k+1)^{\text{th}}$  unit,
- ii) accept  $\mathbf{R}_{y_{k-1}}^{-1}(n-k+1)$  and  $\mathbf{a}_{k-1}(n-k+1)$  from  $(k-1)^{\text{th}}$  unit and
- iii) update  $\mathbf{R}_{y_k}^{-1}(n-k)$  and  $\mathbf{a}_k(n-k)$  to  $\mathbf{R}_{y_k}^{-1}(n-k+1)$  and  $\mathbf{a}_k(n-k+1)$ .

Observe from (18)-(21) that the computations required for updating the eigenvector estimates are identical for all the units, thus making the algorithm both modular and parallel. Consequently, the effective computational requirement is equal to that required for updating only one eigenvector estimate and it is about  $5.5N^2$  multiplications per iteration.

We have the following remarks regarding the convergence of this algorithm. Following the steps as in the convergence analysis of a similar-looking algorithm given in [8], we can show that i) the algorithm is locally convergent (asymptotically) and ii) the undesired stationary points (i.e., the undesired eigenvectors) are unstable. However, when  $\mathbf{R}_y(n) = \mathbf{R}_y$  for all  $n$ , the algorithm can get stuck at an undesired eigenvector. Thus, the proposed algorithm is globally convergent with probability one.

#### 4. SIMULATION RESULTS

In this section, we present some computer simulation results to demonstrate the performance of the proposed adaptive algorithm.

The performance measures used for evaluating the quality of the estimated eigenvectors are as follows. To see how close the estimated eigenvectors are to the true subspace, we use the projection error measure,  $E(n)$ , defined as

$$E(n) = \left\| \left[ \mathbf{I} - \mathbf{S}(\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \right] \overline{\mathbf{A}}(n) \right\|_2 \quad (22)$$

where  $\overline{\mathbf{A}}(n) = [\overline{\mathbf{a}}_1(n), \dots, \overline{\mathbf{a}}_D(n)]$ ,  $\overline{\mathbf{a}}_i(n) = \frac{\mathbf{a}_i(n)}{\|\mathbf{a}_i(n)\|}$  and  $\mathbf{S} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_D]$ . In order to know the extent of  $\mathbf{R}_w$ -orthogonality among the estimated eigenvectors, we define an orthogonality measure,  $Orth_{max}(n)$ , as

$$Orth_{max}(n) = \max_{i,j \in \{1, \dots, D\}} \frac{|\mathbf{a}_i^T(n) \mathbf{c}_j(n)|}{\sqrt{\mathbf{a}_i^T(n) \mathbf{c}_i(n) \cdot \mathbf{a}_j^T(n) \mathbf{c}_j(n)}} \quad (23)$$

Thus, the smaller the values of  $E(n)$  and  $Orth_{max}(n)$ , the better is the quality of the estimated eigenvectors, and vice-versa.

In the simulations, the signal  $y(n)$  was generated as  $y(n) = x(n) + \sigma w(n)$  where

$$\begin{aligned} x(n) &= \rho \sin(2\pi(0.2)n + \theta_1) + \rho \sin(2\pi(0.24)n + \theta_2) \\ w(n) &= \sin(2\pi(0.21)n + \theta_3) + \sin(2\pi(0.23)n + \theta_4) \\ &\quad + \sin(2\pi(0.25)n + \theta_5) + v(n). \end{aligned}$$

Here,  $v(n)$  is a zero-mean and unit variance white noise and  $\theta_i$ 's are the initial phases (assumed to be uniform in  $[-\pi, \pi]$ ). Values of  $N$  and  $D$  were fixed at 10 and 6, respectively, and  $\sigma$  at 0.6325.

The algorithm was initialized as  $\mathbf{a}_k(0) = \mathbf{i}_k$ ,  $k = 1, \dots, D$ , and  $\mathbf{R}_y^{-1}(1) = 100\mathbf{I}$  where  $\mathbf{i}_k$  is the  $k^{\text{th}}$  column of  $\mathbf{I}$ . The matrix  $\hat{\mathbf{R}}_w$  was taken to be the asymptotic covariance matrix of  $w(n)$ . 100 Monte Carlo simulations were performed.

Values of  $E(n)$ , averaged over 100 trials, are plotted in Fig. 2 for  $\rho = 5.00$  (high signal to noise ratio (SNR)) and 1.58 (low SNR). Observe that the algorithm converges quite fast especially in the high SNR case. Values of  $Orth_{max}(n)$  (not shown here) are of the order of  $10^{-4}$  and  $10^{-3}$  for  $\rho = 5.00$  and  $\rho = 1.58$ , respectively, implying that the implicit orthogonalization built into the algorithm through the inflation technique is very effective.

In order to see how this subspace quality reflects in frequency estimation, we used the spectral estimator

$$S(f) = \frac{1}{\mathbf{s}^H(f) \mathbf{P} \mathbf{s}(f)} \quad 0 \leq f \leq 0.5 \quad (24)$$

where  $\mathbf{P} = \overline{\mathbf{A}}(L) \left( \overline{\mathbf{A}}^T(L) \overline{\mathbf{A}}(L) \right)^{-1} \overline{\mathbf{A}}^T(L)$ ,  $L$  is the data length,  $\mathbf{s} = [1, \exp(j2\pi f), \exp(j4\pi f), \dots, \exp(j(N-1)2\pi f)]^T$  and  $H$  denotes Hermitian transpose. The peaks of  $S(f)$  are taken as the estimates of the sinusoidal frequencies in the desired signal  $x(n)$ . The averaged results are shown in Fig. 3. Observe that the desired frequencies are well estimated in spite of the fact that the undesired signal frequencies are closely interlaced with the desired signal frequencies.

#### 5. CONCLUSIONS

The problem of seeking the generalized eigenvector corresponding to the minimum eigenvalue of a symmetric matrix positive definite pencil  $(\mathbf{R}_y, \mathbf{R}_w)$  has been translated into an unconstrained minimization problem. This was then extended to the case of more than one eigenvector using an inflation technique. Based on this asymptotic formulation, a quasi-Newton adaptive algorithm was derived for estimating these eigenvectors in the data case. Note that the algorithm requires the knowledge of noise covariance matrix to within a scalar multiple.

#### REFERENCES

- [1] G. Auchmuty, "Globally and Rapidly Convergent Algorithms for Symmetric Eigenproblems," *SIAM J. Matrix Analysis and Applications*, vol.12, no.4, pp.690-706, Oct. 1991.
- [2] C.B. Moler and G.W. Stewart, "An Algorithm for Generalized Matrix Eigenvalue Problems," *SIAM J. Numerical Analysis*, vol.10, pp.241-256, 1973.
- [3] L. Kaufman, "The LZ-Algorithm to solve the Generalized Eigenvalue Problem," *SIAM J. Numerical Analysis*, vol.11, pp.997-1024, 1974.
- [4] R.S. Martin and J.H. Wilkinson, "Reduction of the Symmetric Eigenproblem  $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$  and Related Problems to Standard Form," *Numerische Mathematik*, vol.11, pp.99-110, 1968.
- [5] A. Bunse-Gerstner, "An Algorithm for the Symmetric Generalized Eigenvalue Problem," *Linear Algebra and Applications*, vol.58, pp.43-68, 1984.
- [6] W. Shougen and Z. Shuqin, "An Algorithm for  $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$  with Symmetric and Positive-Definite  $\mathbf{A}$  and  $\mathbf{B}$ ," *SIAM J. Matrix Analysis and Applications*, vol.12, no.4, pp.654-660, Oct. 1991.
- [7] D.G. Luenberger, *Linear and Non-linear Programming*, pp.366-369, Addison-Wesley, 1978.
- [8] G. Mathew, V.U. Reddy and S. Dasgupta, "Adaptive Estimation of Eigensubspace," to appear in *IEEE Trans. Signal Processing*, Feb. 1995.
- [9] B.N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

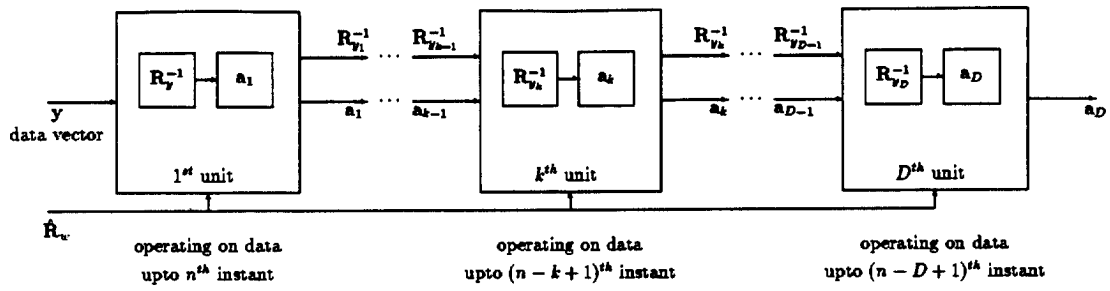


Figure 1. Modular implementation of the proposed method

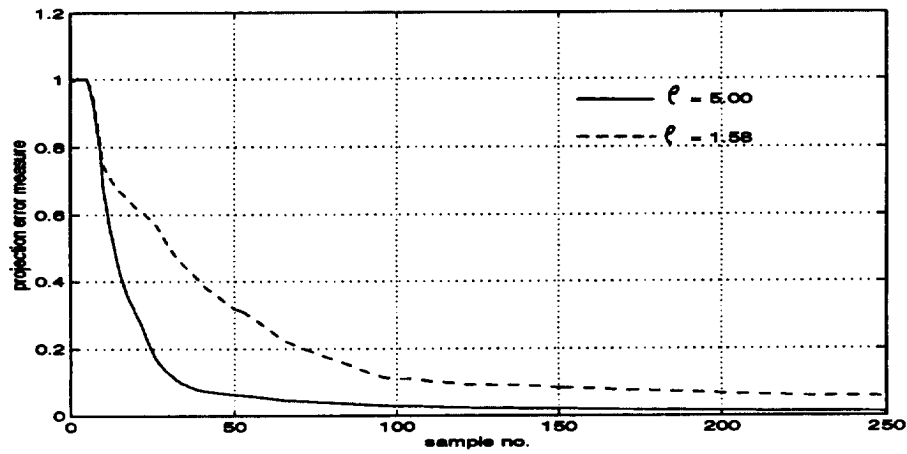


Figure 2. Convergence performance of the proposed method

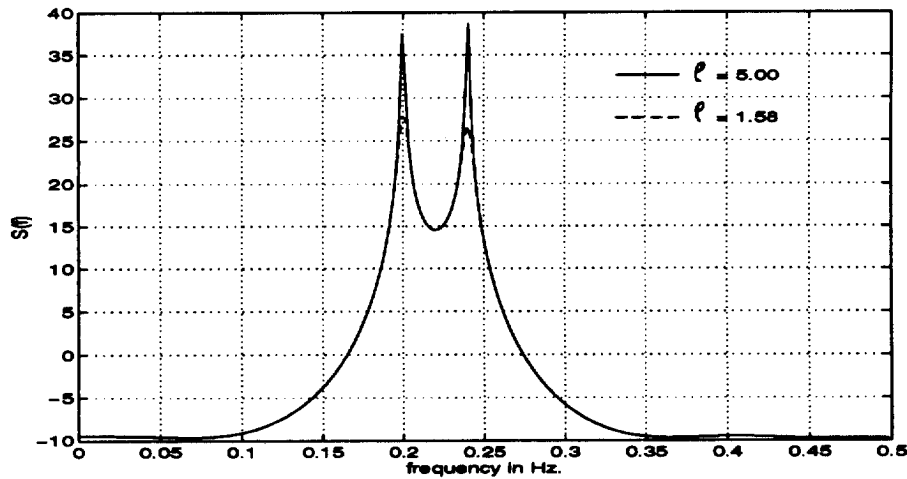


Figure 3. Spectrum estimate using the estimated eigenvectors