A NEW FAMILY OF EVD TRACKING ALGORITHMS USING GIVENS ROTATIONS

Benoît Champagne and Qing-Guang Liu

INRS-Télécommunications, Université du Québec, 16 Place du Commerce
Verdun, Québec, Canada H3E 1H6, champagne@inrs-telecom.uquebec.ca

ABSTRACT

In this work, we derive new algorithms for tracking the eigenvalue decomposition (EVD) of a time-varying data covariance matrix. These algorithms have parallel structures, low operation counts and good convergence behavior. Their main feature is the use of Givens rotations to update the eigenvector estimates. As a result, orthonormality of the latter can be maintained at all time, which is critical in the application of certain signal-subspace methods. The comparative performance of the new algorithms is illustrated by means of computer experiments.

1. INTRODUCTION

In a recent paper [1], new EVD tracking algorithms were developed using a first-order perturbation approach. These algorithms exhibit attractive computational and convergence properties, but they suffer from a limitation common to many EVD tracking algorithms, namely: they do not produce perfectly orthonormal eigenvectors. In some applications, this is not important; in others, further orthonormalization of the eigenvectors is necessary, which requires additional computations. Clearly, it is desirable to avoid this step by directly producing orthonormal eigenvectors.

In this work, we derive new EVD tracking algorithms which do not suffer from this limitation. A constrained linearization approach is first used to obtain an approximate solution to the updated EVD resulting from a small rank-one modification. It consists of representing the updated eigenvectors, constrained to be orthonormal, in terms of (small) unconstrained parameters and to evaluate the latter by performing a linearization within the parameter space. This parametric representation is then exploited to derive several new, computationally efficient EVD tracking algorithms. Their main feature is the use of Givens rotations to update the eigenvector estimates, so that the constraint of orthonormality can be satisfied at all time. The statistical convergence and numerical stability of the new algorithms are investigated by computer experiments.

2. THE EVD TRACKING PROBLEM

2.1. Formulation

Let $\mathbf{x}(k) \in \mathcal{C}^L$ be an L-dimensional complex data vector observed at discrete-time k. The sequence $\mathbf{x}(k)$ is modeled as a zero-mean, random vector process with covariance matrix

$$R(k) = E[\mathbf{x}(k)\mathbf{x}(k)^H]. \tag{1}$$

This work was supported in part by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

The eigenvalues and corresponding orthonormalized eigenvectors of $R(k)$ are denoted by $\lambda_i(k)$ and $\mathbf{q}_i(k)$, $i = 1, ..., L$, respectively. That is, the matrices

$$\Lambda(k) = \text{diag}(\lambda_1(k), ..., \lambda_L(k)), \tag{2}$$

$$Q(k) = [\mathbf{q}_1(k), ..., \mathbf{q}_L(k)], \tag{3}$$

satisfy

$$R(k) = Q(k)\Lambda(k)Q(k)^H, \tag{4}$$

$$Q(k)^H Q(k) = I_L. \tag{5}$$

Without loss of generality, it is convenient to assume that $\lambda_1(k) \geq \lambda_2(k) \geq ... \geq \lambda_L(k) \geq 0$.

The problem of EVD tracking is to perform on-line estimation of the time-varying EVD parameters of the data covariance matrix $R(k)$ in (1). More specifically, it is desired to develop recursive relationships which can be used to obtain estimates of the EVD at time k, i.e. estimates of $\Lambda(k)$ and $Q(k)$ in (2)-(3), given estimates of $\Lambda(k-1)$ and $Q(k-1)$ and the new data vector $\mathbf{x}(k)$. In certain applications, it is only required to track a subset of the EVD. Due to lack of space, we only consider complete EVD update. However, the new algorithms reported in this paper can be modified appropriately.

Let

$$\Gamma(k) = \text{diag}(\gamma_1(k), ..., \gamma_L(k)), \tag{6}$$

with $\gamma_1(k) \geq \gamma_2(k) \geq ... \geq \gamma_L(k) \geq 0$, and

$$U(k) = [\mathbf{u}_1(k), ..., \mathbf{u}_L(k)], \tag{7}$$

denote the desired estimates of $\Lambda(k)$ and $Q(k)$, respectively. In this work, following [1], we shall seek estimates which approximately satisfy

$$U(k)\Gamma(k)U(k)^H = (1-\epsilon)U(k-1)\Gamma(k-1)U(k-1)^H + \epsilon\,\mathbf{x}(k)\mathbf{x}(k)^H, \tag{8}$$

where $0 < \epsilon < 1$ is a forgetting parameter used to deemphasize the effect of past observations. In addition to this, we shall require that

$$U(k)^H U(k) = I_L, \tag{9}$$

exactly. Enforcing the constraint (9) at all time will ensure that the estimated eigenvectors are orthonormal.

Before proceeding with the derivations of new EVD trackers, we need to recast the EVD update problem (8) in a normalized form which will simplify our work.

0-7803-3192-3/96 $5.00©1996 IEEE

2.2. Preprocessing

To simplify the notations, let

$$U \equiv U(k-1), \quad \Gamma \equiv \Gamma(k-1), \quad \mathbf{x} \equiv \mathbf{x}(k), \tag{10}$$

denote the information available at time k and let

$$U' \equiv U(k), \quad \Gamma' \equiv \Gamma(k), \tag{11}$$

denote the updated EVD estimates. Preprocessing consists of four steps. After the ith step ($i = 1, ..., 4$), we have

$$U'\Gamma'U'^H = U_i[(1-\epsilon)\Gamma_i + \epsilon\boldsymbol{\xi}_i\boldsymbol{\xi}_i^H]U_i^H, \tag{12}$$

where U_i, Γ_i and $\boldsymbol{\xi}_i$ are defined so that the problem is gradually simplified [2]. A description of these steps follows.

1. *Diagonalization:* Transform (8) into the rank-one EVD update of a diagonal matrix. To this end, let

$$\boldsymbol{\xi}_1 = U^H\mathbf{x}, \quad U_1 = U, \quad \Gamma_1 = \Gamma. \tag{13}$$

2. *Mapping into real vector space:* Map $\boldsymbol{\xi}_1 \in \mathcal{C}^L$ into $\boldsymbol{\xi}_2 \in \mathcal{R}^L$ so that the updating problem only involves real quantities. To this end, define

$$D = \mathrm{diag}(\xi_{1,i}/|\xi_{1,i}|), \tag{14}$$

where $\xi_{1,i}$ denotes the ith entry of $\boldsymbol{\xi}_1$, and let

$$\boldsymbol{\xi}_2 = D^H\boldsymbol{\xi}_1, \quad U_2 = U_1D, \quad \Gamma_2 = \Gamma_1. \tag{15}$$

3. *Deflation:* Reduce the dimensionality of the problem whenever some of the diagonal elements of Γ_2 are repeated. Specifically, suppose that the number of distinct eigenvalues is $K \leq L$. Then, by using an appropriate block Householder matrix $H = \mathrm{diag}(H_1, ..., H_K)$ [2], it is possible to zero out $L-K$ entries of the vector $\boldsymbol{\xi}_2$ without affecting Γ_2. Thus,

$$\boldsymbol{\xi}_3 = H^T\boldsymbol{\xi}_2, \quad U_3 = U_2H, \quad \Gamma_3 = \Gamma_2. \tag{16}$$

4. *Reordering:* Using a permutation matrix P_1 (see [3]), reorder $\boldsymbol{\xi}_3$, U_3 and Γ_3 via

$$\boldsymbol{\xi}_4 = P_1^T\boldsymbol{\xi}_3, \quad U_4 = U_3P_1, \quad \Gamma_4 = P_1^T\Gamma_3P_1, \tag{17}$$

so that the last $L-K$ entries of $\boldsymbol{\xi}_4$ are zero and the first K diagonal entries of Γ_4 are in decreasing order.

These steps are summarized in Table 1.

Step	Operation
1	$\boldsymbol{\xi} = U^H\mathbf{x}$
2	$D = \mathrm{diag}(\xi_i/\lvert\xi_i\rvert)$ $\boldsymbol{\xi} \leftarrow D^H\boldsymbol{\xi}$ $U \leftarrow UD$
3	$H =$ block Householder matrix $\boldsymbol{\xi} \leftarrow H^T\boldsymbol{\xi}$ $U \leftarrow UH$
4	$P_1 =$ permutation matrix $\boldsymbol{\xi} \leftarrow P_1^T\boldsymbol{\xi}$ $U \leftarrow UP_1$ $\Gamma \leftarrow P_1^T\Gamma P_1$

Table 1: Summary of preprocessing steps.

3. CONSTRAINED LINEARIZATION

In the sequel, let $\boldsymbol{\xi}$, U and Γ stand for $\boldsymbol{\xi}_4$, U_4 and Γ_4, respectively. Then, we have

$$\boldsymbol{\xi}^T = [\boldsymbol{\xi}_u^T, \mathbf{0}], \quad \Gamma = \mathrm{diag}(\Gamma_u, \Gamma_l), \tag{18}$$

where $\boldsymbol{\xi}_u = [\xi_1, ..., \xi_K]^T$ with $\xi_i > 0$, $\Gamma_u = \mathrm{diag}(\gamma_1, ..., \gamma_K)$ with $\gamma_1 > \gamma_2 > ... > \gamma_K \geq 0$, $\Gamma_l = \mathrm{diag}(\gamma_{K+1}, ..., \gamma_L)$. Substituting (18) into (12), we obtain

$$U'\Gamma'U'^H = U\begin{bmatrix} (1-\epsilon)\Gamma_u + \epsilon\boldsymbol{\xi}_u\boldsymbol{\xi}_u^T & \mathbf{0} \\ \mathbf{0} & (1-\epsilon)\Gamma_l \end{bmatrix}U^H. \tag{19}$$

Hence, the original EVD update problem (8) over $\mathcal{C}^{L\times L}$ has been reduced to the rank-one EVD update of a diagonal matrix over $\mathcal{R}^{K\times K}$, i.e.:

$$V\Gamma_u'V^T = (1-\epsilon)\Gamma_u + \epsilon\boldsymbol{\xi}_u\boldsymbol{\xi}_u^T, \tag{20}$$

where Γ_u' is diagonal and

$$V^TV = I_K. \tag{21}$$

Once V and Γ_u' are known, Γ' and U' can be obtained from

$$U' = U\begin{bmatrix} V & \mathbf{0} \\ \mathbf{0} & I_{L-K} \end{bmatrix}, \quad \Gamma' = \begin{bmatrix} \Gamma_u' & \mathbf{0} \\ \mathbf{0} & (1-\epsilon)\Gamma_l \end{bmatrix}. \tag{22}$$

Below, a constrained linearization approach is used to derive an approximate solution to (20) which satisfies (21) exactly.

The following observations are at the basis of our derivation: (1) in most applications of EVD tracking, the memory parameter ϵ is small; (2) for ϵ sufficiently small, the modified EVD components Γ_u' and V in (20) can be analytically connected to Γ_u and I_K, respectively, so that $\Gamma_u{}' \to \Gamma_u$ and $V \to I_K$ in the limit $\epsilon \to 0$ (see [1]). Hence, we conclude that in most applications of EVD tracking, the EVD modifications resulting from the update (20) are small, that is, $\|\Gamma_u' - \Gamma_u\|_2 \ll \|\Gamma_u\|_2$ and $\|V - I_K\|_2 \ll 1$.

To emphasize this point, let us write Γ_u' in the form

$$\Gamma_u' = \Gamma_u + \Delta\Gamma_u, \tag{23}$$

$$\Delta\Gamma_u = \mathrm{diag}(\delta_1, ..., \delta_K), \tag{24}$$

where δ_i ($i = 1, ..., K$) represents the modification in the ith eigenvalue. According to the above discussion, $|\delta_i| \ll \gamma_1$ provided ϵ is sufficiently small.

The introduction of a similar representation for V in terms of small parameters requires additional care because of the orthogonality constraint (21). To derive such a representation, we first note that $\det(V) = \pm 1$ as a consequence of (21). Without loss of generality, we shall assume that $\det(V) = +1$. This amounts to multiplying one of the modified eigenvectors by -1. With this additional restriction, V now belongs to the group of $K \times K$ unimodular orthogonal matrices and can thus be expressed in the form [4]

$$V = \exp(\Theta), \tag{25}$$

where $\Theta = [\theta_{ij}]$ is a skew-symmetric matrix in $\mathcal{R}^{K\times K}$ (i.e., $\Theta^T = -\Theta$, or equivalently, $\theta_{ji} = -\theta_{ij}$), and $\exp(.)$ is the matrix exponential function, defined as

$$\exp(\Theta) = \Sigma_{k=0}^{\infty}\Theta^k/k! \tag{26}$$

This parametrization of V in terms of $K(K-1)/2$ real parameters is of particular interest to us. Indeed, provided ϵ is sufficiently small, we have $\|V - I_K\|_2 \ll 1$, which in terms of (25), implies $|\theta_{ij}| \ll 1$.

The constrained linearization approach that we propose can now be stated as follows: substitute (23) and (25) into (20); perform the necessary expansions and retain only linear terms in $\Delta\Gamma_u$ and Θ; solve for $\Delta\Gamma_u$ and Θ; substitute the solutions back into (23) and (25); The results can be summarized as follows:

$$\Gamma'_u = (1-\epsilon)\Gamma_u + \epsilon\,\mathrm{diag}(\xi_1^2, ..., \xi_K^2), \tag{27}$$

while V is given by (25) with

$$\theta_{ij} = \epsilon\,\xi_i\xi_j/(\gamma_j - \gamma_i), \quad 1 \leq i < j \leq K. \tag{28}$$

The other entries of Θ are obtained from skew-symmetry. The matrices Γ'_u and V as given above are the desired solutions to the simplified EVD update problem (20). Because of the parametrization (25), the orthonormality constraint (21) is automatically satisfied.

4. EVD UPDATE BASED ON GIVENS ROTATION SEQUENCES

In this section, the assumption $\|\Theta\|_2 \ll 1$ is further exploited to derive several (approximate) decompositions of the matrix V (25) as a product of Givens (or plane) rotations. When used in connection with (22), each decomposition leads to a computationally efficient algorithm for EVD update. These decompositions, along with the corresponding algorithms, are derived in the following subsections. We use the acronyms PROTEUS-i to identify these algorithms, where PROTEUS stands for "plane rotation-based EVD update scheme", and the index $i \in \{1, 2, 3\}$.

4.1. PROTEUS-1

Let Ψ_{ij} $(i < j)$ denote the matrix obtained from Θ by setting all its entries to zero, except for the ij and ji-entries (i.e., θ_{ij} and $\theta_{ji} = -\theta_{ij}$), which are left unchanged. It is then possible to express Θ in terms of the matrices Ψ_{ij} as

$$\Theta = \Sigma_{j>i}\Psi_{ij}. \tag{29}$$

Substituting (29) in (25) and using the definition (26) of the exponential matrix function, it can be shown that

$$V = \Pi_{j>i}\exp(\Psi_{ij}) + O(\epsilon^2). \tag{30}$$

Thus, V can be expressed as a finite product of simpler orthogonal matrices, namely $\exp(\Psi_{ij})$, plus a matrix error term of the order of ϵ^2. For small values of ϵ, which is the situation of interest in this work, it is reasonable to neglect this term. Now consider the matrix $\exp(\Psi_{ij})$, which is the basic building block in (30). Using the definition (26) of the matrix exponential function, one can verify that

$$\exp(\Psi_{ij}) = G_{ij}(\theta_{ij}), \tag{31}$$

where $G_{ij}(\theta) \in \mathcal{R}^{K\times K}$ is the well known Givens rotation matrix [3]. Substituting (31) in (30) and neglecting the second degree error term, we obtain a first decomposition of the matrix V (25), namely:

$$V_1 = \Pi_{j>i}G_{ij}(\theta_{ij}), \tag{32}$$

where the rotation parameters θ_{ij} are given by (28). This result simply states that for ϵ small, V can be expressed as the product of $K(K-1)/2$ small Givens rotation matrices with angles θ_{ij} (28).

Step	Operation
1	$\boldsymbol{\xi} \leftarrow \sqrt{\epsilon}\boldsymbol{\xi}$

2	for $i = 1 : K-1$
	  for $j = i+1 : K$
	    $\theta \leftarrow \xi_i\xi_j/(\gamma_j - \gamma_i)$
	    $U \leftarrow UG_{ij}(\theta)$
	  end
	end

3	$\Gamma \leftarrow (1-\epsilon)\Gamma + \mathrm{diag}(\xi_i^2)$

Table 2: PROTEUS-1 algorithm.

Based on the above decomposition of V, we can now formulate a complete EVD update algorithm. To this end, we first replace V in (22) by the decomposition V_1 (32). We then rewrite the first equation in (22) in the form

$$U' = U\,\Pi_{j>i}G_{ij}(\theta_{ij}), \tag{33}$$

where $G_{ij}(\theta_{ij})$ now represents a Givens rotation in $\mathcal{R}^{L\times L}$. For the eigenvalue update, we use (27) in connection with (22). The resulting algorithm is presented in Table 2, where the initial values of $\boldsymbol{\xi}$, Γ and U are those obtained after preprocessing (Table 1). For real data, this algorithm requires $3LK^2 + O(K^2)$ flops.

4.2. PROTEUS-2

In this subsection, we derive a simpler algorithm with an $O(LK)$ operation count. The starting point of our derivation is a block representation for the matrix $\Theta = [\theta_{ij}]_{K\times K}$, with entries given by (28). Let $\Theta_k(\xi_1, ..., \xi_k)$ denote the principal submatrix of Θ corresponding to its first k rows and first k columns. Then, we have

$$\Theta = \Theta_K(\xi_1, ..., \xi_K) = \begin{bmatrix} \Theta_{K-2}(\xi_1, ..., \xi_{K-2}) & -\epsilon\,\xi_{K-1}\mathbf{a} & -\epsilon\,\xi_K\mathbf{b} \\ \epsilon\,\xi_{K-1}\mathbf{a}^T & 0 & \theta_{K-1,K} \\ \epsilon\,\xi_K\mathbf{b}^T & -\theta_{K-1,K} & 0 \end{bmatrix} \tag{34}$$

where

$$\mathbf{a} = (a_1, ..., a_{K-2})^T, \quad a_i = \xi_i/(\gamma_i - \gamma_{K-1}), \tag{35}$$

$$\mathbf{b} = (b_1, ..., b_{K-2})^T, \quad b_i = \xi_i/(\gamma_i - \gamma_K). \tag{36}$$

Our second EVD update algorithm is based on the assumption $a_i \approx b_i \approx \xi_i/\gamma_i$. Under this condition, the transformation $\Theta \rightarrow A_{K-1}^T\Theta A_{K-1}$, where $A_{K-1} = G_{K-1,K}(\alpha)$ with α properly defined, can be used to zero out the first $K-2$ entries in the last row and column of Θ. Using additional properties of Givens rotations and of the exponential matrix function, along with mathematical induction, the following decomposition of V (25) can be obtained:

$$V_2 = B_{K-1}...B_2CA_2^T...A_{K-1}^T, \tag{37}$$

$$A_i = G_{i,i+1}(\alpha_i), \quad B_i = G_{i,i+1}(\beta_i), \quad C = G_{1,2}(\theta_1), \tag{38}$$

where α_i, β_i and θ_1 are appropriately defined parameters. The corresponding EVD update algorithm is summarized in Table 3. Its operation count is $12LK + O(K)$ flops. In practice, this algorithm is robust and can be used even when the underlying assumptions do not hold.

Step	Operation
1	$\xi \leftarrow \sqrt{\epsilon}\xi$
2	$\xi'_K \leftarrow \xi_K$ for $i = K-1 : -1 : 2$ $\alpha_i \leftarrow -\arctan(\xi'_{i+1}/\xi_i)$ $\theta \leftarrow \xi_i \xi'_{i+1}/(\gamma_{i+1} - \gamma_i)$ $\xi'_i \leftarrow \sqrt{\xi_i^2 + (\xi'_{i+1})^2}$ $U \leftarrow UG_{i,i+1}(\alpha_i + \theta)$ end $\theta \leftarrow \xi_1 \xi'_2/(\gamma_2 - \gamma_1)$ $U \leftarrow UG_{1,2}(\theta)$ for $i = 2 : K-1$ $U \leftarrow UG^T_{i,i+1}(\alpha_i)$ end
3	$\Gamma \leftarrow (1-\epsilon)\Gamma + \text{diag}(\xi_i^2)$

Table 3: PROTEUS-2 algorithm.

Step	Operation
1	$\xi \leftarrow \sqrt{\epsilon}\xi$
2	for $i = 1 : K-1$ for $j = i+1 : \min(i+l, K)$ $\theta \leftarrow \xi_i \xi_j/(\gamma_j - \gamma_i)$ $U \leftarrow UG_{ij}(\theta)$ end end
3	$\Gamma \leftarrow (1-\epsilon)\Gamma + \text{diag}(\xi_i^2)$

Table 4: PROTEUS-3l algorithm.

4.3. PROTEUS-3

In this subsection, we derive yet another type of EVD update algorithms with an $O(lLK)$ operation count, where l is a user selectable integer parameter. Consider again the matrix $\Theta = [\theta_{ij}]_{K\times K}$, with entries given by (28). We have observed experimentally that the magnitude of θ_{ij} generally decreases (although not necessarily monotonically) as $|j-i|$ increases, i.e., as we move away from the main diagonal of Θ. This suggests that a simple approximation for the matrix Θ can be obtained by retaining only the first l diagonals of Θ above and below the main diagonal and by setting all the other entries to zero. Here, l is a fixed, small integer (typically 1 or 2). Using this idea in connection with (32), we obtain a third decomposition of V, namely:

$$V_{3,l} = \Pi_{i=1}^{K-1} \Pi_{j=i+1}^{\min(i+l,K)} G_{ij}(\theta_{ij}). \qquad (39)$$

The corresponding EVD update algorithm is summarized in Table 4. Note that for $l = K$, this algorithm is identical to the PROTEUS-1 algorithm. Clearly, significant computational gains are achieved with the algorithm in Table 4 only when $l \ll K$, in which case its operation count is $6lLK + O(K)$ flops.

5. COMPUTER EXPERIMENTS

We consider a uniform linear array of $L = 10$ sensors with half-wavelength spacing. The wavefield consists of $K-1$ Gaussian narrow-band plane wave signals in white noise. The DOAs of the sources are 0^o, 5^o, 20^o and 25^o (w.r.t. broadside) and the corresponding SNRs are 20, 20, 10 and 10dB. The new algorithms, as well as a brute force approach involving exact EVD of a recursive, exponentially weighted

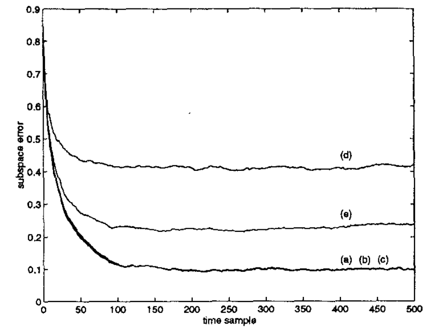

Figure 1: Distance between estimated and true signal-subspace: (a) exact EVD, (b) PROTEUS-1, (c) PROTEUS-2, (d) PROTEUS-3 with $l = 1$, (e) PROTEUS-3 with $l = 2$.

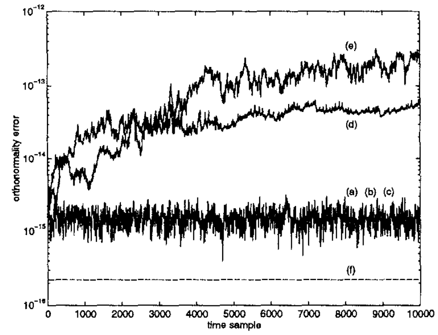

Figure 2: Orthonormality of signal-subspace eigenvectors: (a)-(e) as above, (f) machine accuracy.

covariance matrix estimate, are applied to the data. In all cases, $\epsilon = 0.98$. Fig. 1 shows the distance between the estimated and true signal subspace as a function of k for the various methods (20 run average). Results indicate that the new algorithms PROTEUS-1 and 2 can achieve the same level of performance as the much more costly "exact approach". Fig. 2 shows the quantity $\|U_s(k)^H U_s(k) - I_4\|_2$ versus k for a single run, where $U_s(k)$ contains the estimated signal-subspace eigenvectors. Theses results indicate PROTEUS-1 and 2 are very effective in preserving the orthogonality of the estimated eigenvectors in finite precision arithmetic; the performance of PROTEUS-3 can be improved by using known numerical stabilization mechanisms.

6. REFERENCES

[1] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on first-order perturbations", *IEEE Trans. on Signal Processing*, vol. 42, pp. 2758–2770, Oct. 1994.
[2] J. R. Bunch, C. P. Nielsen and D. C. Sorensen, "Rank-one modification of the symmetric eigenproblem," *Numer. Math.*, vol. 31, pp. 31-48, 1978.
[3] G. H. Golub and C. F. Van Loan, *Matrix Computations, 2nd Ed.*, Baltimore: John Hopkins University Press, 1989.
[4] H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover, 1950.