JOURNAL OF COMPUTATIONAL PHYSICS 45, 100–107 (1982)

An Iterative Method for Calculating Several of the Extreme Eigensolutions of Large Real Non-symmetric Matrices

STEN RETTRUP

Department of Physical Chemistry, H. C. Ørsted Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark

Received January 15, 1981; revised August 24, 1981

A new method is described for calculating several of the extreme eigensolutions of large real non-symmetric matrices. The matrices considered are assumed to have real eigenvectors and eigenvalues. An algorithm is presented for both single and multiple vector optimization. The efficiency of the two versions are compared using some test calculations. The method is a generalization of the Davidson–approach.

1. INTRODUCTION

The calculation of the lowest (highest) eigenvalues and the corresponding eigenvectors of large real symmetric matrices has in recent years been extensively studied. Accordingly solution of very large real symmetric eigenvalue problems, which occur frequently in connection with configuration interaction studies of many-body correlations, has become a matter of routine.

Essentially two iterative approaches have been used for computer implementation. One is based on the method of coordinate relaxation (CR) which was proposed by Nesbet [1] and later improved by Shavitt [2] including the root shifting technique by Shavitt *et al.* [3] to be used for calculating several of the extreme eigensolutions. Recently Raffenetti [4] has developed a new modified CR algorithm in which coordinates for several eigenvectors are relaxed simultaneously (SCR). In comparison with previous CR-schemes the algorithm provides a significant improvement in the rate of convergence. This feature is in particular marked when calculating eigensolutions corresponding to degenerate or near degenerate roots.

The expansion method developed by Davidson [5] represents the alternative method. The characteristic feature of this method is, that it automatically keeps different eigenvectors orthogonal by diagonalizing an intermediate matrix of small dimension defined on a small "optimal" set of basis vectors. As a consequence it is found that near degeneracy of the eigenvalues does not lead to significant changes in the convergence rate. In addition the method has the advantage of not requiring the matrix to be stored in a specific mode on external files. This is of particular impor-

0021-9991/82/010100–08$02.00/0

tance for large dimensional matrices for which the elements have not been obtained in a specific order. Which of the two approaches is most appropriate, however, seems to depend strongly on the structure of the actual matrix considered.

Methods for calculating the extreme eigensolutions of large non-Hermitian matrices have been studied much less. In several problems of physical and chemical interest, however, it is necessary to have efficient procedures available for that purpose. One example is in configuration interaction studies using non-orthogonal basis states using a biorthonormal representation (Norbeck and McWeeny [6]). Another is in the study of excited states using the so-called random phase approximation (RPA) or time-dependent Hartree–Fock theory (TDHF) (see, e.g., Oddershede [7]). Usually the problem is solved in practice by transforming the eigenvalue problem into intermediate Hermitian forms (Jørgensen and Linderberg [8]). In this approach to the problem the computational effort involved increases very rapidly with the order of the matrix due to the very time consuming full diagonalization and subsequent transformation required. As a consequence the applicability of the method is practically useful only when the dimensionality of the matrix is relatively small.

In order to overcome these difficulties Bender and Shavitt [9], Nisbet [10] and recently Flament and Gervais [11] have investigated the possibility of solving the eigenvalue problem

$$\mathbf{M}C = \lambda \mathbf{S}C \tag{1}$$

iteratively by applying a generalization of the method of coordinate relaxation. In Eq. (1), **M** is restricted to be a non-symmetric matrix having both real eigenvalues, λ, and eigenvectors **C**. **S** is a real positive definite metric matrix. The algorithms used in [9] and [10] were essentially based on a modification of the original Shavitt and Nesbet procedure [2] and considered only calculation of the lowest root. These authors found that the convergence properties were dependent on the eigenvalues, λ, being well separated from the ratio M_{ii}/S_{ii} of the diagonal elements of **M** and **S**.

In the study of the equation of motion problem Flament and Gervais [11] considered the calculation of several of the lowest eigensolutions of the RPA equations using a modified version of the method of optimal relaxation (MOR [3]). As pointed out by the authors convergence problems may arise in the case of almost degenerate roots. In addition the rate of convergence for roots above the lowest was found to decrease rapidly. In the test example given by the authors 142 iterations were required for calculating the third eigensolution compared with 12 and 42 for the first and second, respectively.

Matrices of very large dimension must be kept on external storage in case they are not extremely sparse. As a consequence a dominant part of the computation time required for diagonalizing such matrices is associated with the process of retrieving the matrix. Therefore the modified MOR procedure does not seem suited for calculating efficiently more than at most the first two or three eigensolutions. Basically the problem appears to be the same as found in the case of real symmetric matrices, namely, to keep higher eigenvectors properly related to lower solutions already calculated.

In the following section we describe an alternative iterative procedure for calculating several of the lowest (or highest) eigensolutions of a large real non-symmetric matrix in an attempt to overcome the drawbacks of the previous methods. As in Eq. (1) the matrix is assumed to have real eigenvalues and eigenvectors. For convenience the metric matrix in Eq. (1) is a unit matrix. The procedure, however, is readily modified to handle the general case. Essentially the method is a generalization of the expansion method developed by Davidson [5]. In Section 2 we outline the procedure and in Section 3 a series of test calculations are described and discussed.

2. Present Approach

Let V_n denote the total n-dimensional vector space spanned by the n orthonormal unit vectors e_i, $i = 1, 2,..., n$, consisting of the components $e_i(j)$ satisfying $e_i(j) = \delta_{ij}$ $(1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n)$, where δ_{ij} represents the usual Kronecker delta. Contained in V_n we define an m-dimensional subspace V_m $(1 \leqslant m \ll n)$, spanned by a set of m orthonormal n-dimensional vectors

$$\{Q_i | i = 1, 2,..., m;\ Q_i \cdot Q_j = \delta_{ij}\}.$$

Assume in the following $\mathbf{M}$ is a real square matrix of order n having real eigensolutions. The problem to be studied is now to determine an optimal subspace V_m in such a way that the lowest k eigensolutions of the equation

$$\mathbf{M}C_l = \lambda_l C_l, \qquad l = 1, 2,..., k \tag{2}$$

can be expressed in the form

$$C_l = \sum_{i=1}^{m} Q_i a_{il}^{(m)}, \qquad Q_i \in V_m. \tag{3}$$

This problem can be solved iteratively. In order to initialize the process we follow Ref. [5] and choose a set of m orthonormal basis vectors $\{Q_1, Q_2,..., Q_m\}$, $k \leqslant m$, in such a way that the corresponding space, V_m, contains the dominant parts of the lowest k eigenvectors. Usually it is done by finding the m lowest diagonal elements $M_{l_1 l_1} \leqslant M_{l_2 l_2} \leqslant \cdots \leqslant M_{l_m l_m}$ and setting $Q_i = e_{l_i}$ for $i = 1, 2,..., m$. The first approximate solutions of Eq. (2) are then obtained within the subspace V_m by solving the small eigenvalue problem

$$\mathbf{M}^{(m)}\mathbf{a}^{(m)} = \mathbf{a}^{(m)}\boldsymbol{\lambda}^{(m)}, \tag{4}$$

where

$$M_{ij}^{(m)} = Q_i^{+}\mathbf{M}Q_j \qquad (i, j = 1, 2,..., m), \tag{5}$$

and $\boldsymbol{\lambda}^{(m)}$ is an m-dimensional diagonal matrix containing the approximate eigenvalues of Eq. (2).

It should be noted that the set of eigenvectors $a_i^{(m)}$ in $\mathbf{a}^{(m)}$ does not necessarily constitute an orthonormal set. For general eigenvalue problems as Eq. (4) the right eigenvectors $a_i^{(m)}$ are orthogonal with respect to the left eigenvectors $b_j^{(m)}$ $(i \neq j)$,

$$\mathbf{b}^{(m)}\mathbf{M}^{(m)} = \boldsymbol{\lambda}^{(m)}\mathbf{b}^{(m)}, \tag{6}$$

and are usually normalized according to the biorthogonality relation

$$\mathbf{b}^{(m)} \cdot \mathbf{a}^{(m)} = \mathbf{1}, \tag{7}$$

with $\mathbf{1}$ being a unit matrix. A detailed discussion of orthonormality of eigenvectors for non-Hermitian matrices is found in Ref. [12]. For convenience, however, we choose to normalize separately each of the eigenvectors $a_i^{(m)}$ to unity

$$a_i^{(m)} \cdot a_i^{(m)} = 1, \qquad i = 1, 2, \ldots, m. \tag{8}$$

Due to the fact that m $(k \leqslant m \ll n)$, is a relatively small number Eq. (4) can be solved readily by application of standard procedures.

The next step in the method is to extend the dimensionality of the subspace, V_m, by adding to the set a new "optimal" vector Q_{m+1}. Assuming we are solving Eq. (2) for the lth root $(l = 1, 2, \ldots, k)$, the vector should be chosen as the "best" correction vector for the current estimate of the lth eigenvector given by

$$C_l^{(m)} = \sum_{i=1}^{m} Q_i a_{il}^{(m)}. \tag{9}$$

The rate of convergence depends strongly on the choice of correction vector. Only an approximate vector can be obtained directly, however. Requiring (as in the original work by Nesbet [1]) the following set of equations $(i = 1, 2\ldots, n)$ to be satisfied for the ith component only

$$\mathbf{M}(c_l^{(m)} + w_l^{(m)}(i)\, e_i) = \lambda_{ll}^{(m)}(c_l^{(m)} + w_l^{(m)}(i)\, e_i) \tag{10}$$

we obtain the following n equations for determining the parameter $w_l^{(m)}(i)$

$$w_l^{(m)}(i) = R_l^{(m)}(k)/(\lambda_{ll}^{(m)}\delta_{ki} - M_{ki}), \qquad k = 1, 2, \ldots, n, \tag{11}$$

with $R_l^{(m)}(k)$ being the kth component of the current residual vector

$$R_l^{(m)} = (\mathbf{M} - \lambda_{ll}^{(m)}\mathbf{1})\, C_l^{(m)} \tag{12}$$

for $C_l^{(m)}$ normalized to unity (cf. Eqs. (8) and (9)).

Equation (11) can only be satisfied for all $k = 1, 2, \ldots, n$ for $R_l^{(m)}$ being the zero

vector (i.e., $C_l^{(m)}$ is an exact eigenvector). An approximate set of parameters $w_l^{(m)}(i)$, $i = 1, 2, \ldots, n$, can be obtained, however, by choosing $k = i$ in Eq. (11) giving

$$w_l^{(m)}(i) = R_l^{(m)}(i)/(\lambda_{ll}^{(m)} - M_{ii}). \tag{13}$$

as in the case of symmetric matrices [1, 5].

Letting $w_l^{(m)}$ represent the vector $(w_l^{(m)}(1), w_l^{(m)}(2), \ldots, w_l^{(m)}(n))$ we obtain Q_{m+1} by orthonormalizing $w_l^{(m)}$ onto V_m as

$$Q'_{m+1} = w_l^{(m)} - \sum_{k=1}^{m} (w_l^{(m)} \cdot Q_k)\, Q_k, \tag{14}$$

$$Q_{m+1} = Q'_{m+1}/\| Q'_{m+1} \|. \tag{15}$$

The procedure is now cyclically repeated starting with solving the analog of Eq. (4) with $m + 1$ replacing m.

In case the dimensionality of the matrix $\mathbf{M}^{(m)}$ becomes inconveniently large the process is restarted using as initial subspace the space V_k spanned by the current approximate k eigenvectors from Eq. (9). The new set of basis vectors in V_k is then obtained by applying a Schmidt orthonormalization to the full set of basis vectors.

The iterations are terminated when the process has converged for a specific root. This is most easily tested by requiring the norm of the residual vector to be less than a certain fixed threshold, e.g., $\| R_l^{(m)} \| < T$. If several eigensolutions are desired the iterations are continued on the remaining roots.

The proper relationship between the eigenvectors are preserved due to the matrix diagonalization in Eq. (4). This represents a significant advantage compared with previous methods. In these the eigenvalues are estimated by considering the generalized Rayleigh quotient. It requires simultaneous knowledge of both right and left eigenvectors. In the present approach only right eigenvectors need be considered.

If several eigensolutions are desired the idea due to Raffenetti [4] of optimizing simultaneously several vectors is easily incorporated in the present scheme. It is done by calculating each time the matrix is read from external storage all the residual vectors corresponding to the k solutions wanted. The additional set of k correction vectors $Q_{m+1}, Q_{m+2}, \ldots, Q_{m+k}$, is obtained using Eqs. (12)–(13) followed by a Schmidt orthonormalization of the resulting vectors $w_1^{(m)}, w_2^{(m)}, \ldots, w_k^{(m)}$, onto the space V_m using Eqs. (14)–(15). The process is then continued as in the case of a single correction vector by solving Eq. (4) with $m + k$ replacing m.

3. Computational Results

The methods outlined above have been implemented on an IBM 3033. In order to reduce the amount of input/output operations the trial vectors are all kept in core and only the matrix resides in external storage. The convergence properties of the procedures have been studied by carrying out a series of test calculations using

different matrices. In all studied cases no convergence problems did occur neither in connection with degenerate roots nor for calculation higher subdominant eigensolutions.

In order to faciliate comparison we describe in the following a series of calculations which have been performed using the same non-symmetric matrix of dimension 529. This particular matrix was obtained as $(\mathbf{A}-\mathbf{B})\ (\mathbf{A}+\mathbf{B})$, where **A** and **B** are the two square matrices (in casu 529 by 529) entering the RPA equations [7, 8, 13], and where the actual matrix elements were computed for the molecule *trans*-cyclooctene in a minimal basis atomic set [13]. The resulting non-Hermitian matrix is known to have real eigensolutions.

In all calculations the starting dimension of the subspace was chosen to be equal to the final number of solutions desired and the same dimension was used when the subspace was truncated. The convergence criterium was set to $T = 10^{-5}$.

In the first series of calculations the single vector optimization algorithm is used starting from the lowest root calculating successively higher solutions. Table I shows the number of iterations required for obtaining the five lowest roots for various choices of the maximum dimension of the subspace, P. Essentially it demonstrates the drastic reduction in the number of iterations when P is increased. This is in particular marked when calculating higher roots.

Table II shows the equivalent results obtained when utilizing the multiple vector optimization method. In this case we observe a similar behaviour as in the first series of calculations. The most important result, however, is that the matrix is now only required very few times. For example for $P = 30$ it is only necessary to read the matrix 13 times from external storage.

In order to investigate the possibility of obtaining also higher roots using the multiple vector optimization method we carried out a calculation of the 10 lowest roots of the same matrix with $P = 30$ and $T = 10^{-5}$. The number of iterations

TABLE I

The Number of Iterations Required for Obtaining the Five Lowest Eigensolutions Using Single Vector Optimization as a Function of the Maximum Dimension of the Subspace

Maximum dimension of subspace P	Iterations per root					Total No. of iterations	No. of times matrix is required
	1	2	3	4	5		
8	14	14	15	133	124	300	300
10	12	10	11	59	51	143	143
15	11	9	9	41	42	112	112
20	11	9	7	33	21	81	81
25	11	7	8	27	18	71	71
30	11	7	6	28	19	71	71

Table II

The Number of Iterations Required for Obtaining the Five Lowest Eigensolutions Using Multiple Vector Optimization as a Function of the Maximum Dimension of the Subspace

Maximum dimension of subspace P	Iterations per root					Total No. of iterations	No. of times matrix is required
	1	2	3	4	5		
10	14	37	15	24	107	197	107
15	10	14	11	15	16	66	16
20	9	11	10	12	15	57	15
25	9	10	10	12	15	56	15
30	9	10	9	12	13	53	13

required per root was found to be respectively 9, 14, 9, 11, 12, 17, 14, 28, 17, 31, which implies that the matrix was only read from external storage 31 times. The CPU time used was approximately 15 times less than the time used when applying the standard method [8]. Finally we conclude that the present method maintains the fast convergence rate also in calculating relatively many of the subdominant roots. In addition the method requires only a limited amount of input/output processing.

At present a version of the method is being developed which utilizes directly the specific form of the RPA equations [14].

ACKNOWLEDGMENTS

The author is grateful to Dr. Aa. E. Hansen and Professor T. D. Bouman for enlightening discussions and for making the used test matrices available. Professor E. R. Davidson is acknowledged for bringing Ref. [5b] to the author's attention and for clarifying suggestions.

REFERENCES

1. R. K. NESBET, *J. Chem. Phys.* **43** (1965), 311.
2. I. SHAVITT, *J. Comput. Phys.* **6** (1970), 124.
3. I. SHAVITT, C. F. BENDER, A. PIPANO, AND R. P. HOSTENY, *J. Comput. Phys.* **11** (1973), 90.
4. R. C. RAFFENETTI, *J. Comput. Phys.* **32** (1979), 403.
5. (a) E. R. DAVIDSON, *J. Comput. Phys.* **17** (1975), 87; (b) *J. Phys. A* **13** (1980), L179.
6. J. M. NORBECK AND R. MCWEENY, *Chem. Phys. Lett.* **34** (1975), 206.
7. J. ODDERSHEDE, "Advances in Quantum Chemistry" (P–O. Löwdin, Ed.), Vol. 11, Academic Press, London, 1978.
8. P. JØRGENSEN AND J. LINDERBERG, *Int. J. Quantum Chem.* **4** (1970), 587.
9. C. F. BENDER AND I. SHAVITT, *J. Comput. Phys.* **6** (1970), 146.
10. R. M. NISBET, *J. Comput. Phys.* **10** (1972), 614.

11. J. P. FLAMENT AND H. G. GERVAIS, *Int. J. Quantum Chem.* **16** (1979), 1347.
12. C. G. HALL, "Matrices and Tensors," Pergamon, London, 1963.
13. AA. E. HANSEN AND T. D. BOUMAN, "Advances in Chemical Physics" (I. Prigogine and S. A. Rice, Eds.), Vol. 44, Wiley, New York, 1980.
14. AA. E. HANSEN, T. D. BOUMAN, AND S. RETTRUP, to be published.