Reprinted from JOURNAL OF COMPUTATIONAL PHYSICS
All Rights Reserved by Academic Press, New York and London

Vol. 45, No. 2, February 1982
Printed in Belgium

A Generalization of the Davidson's Method to Large Nonsymmetric Eigenvalue Problems

K. HIRAO

Department of Chemistry, College of General Education, Nagoya University, Nagoya, Japan

AND

H. NAKATSUJI

Department of Hydrocarbon Chemistry, Faculty of Engineering, Kyoto University, Kyoto, Japan

Received June 18, 1981

The Davidson's algorithm for solving large symmetric matrices is generalized to nonsymmetric cases. The method is based on the non-variational eigenvalue equation. The algorithm can be expected to converge particularly well for the eigenvalue whose eigenvector has a desired structure (excited states as well as the ground state). The method has been applied successfully to some nonsymmetric eigenvalue problems in our cluster expansion theory.

I. INTRODUCTION

Large eigenvalue problems arise in many scientific and engineering problems. For symmetric problems a great number of methods have been developed for evaluating eigenvalues and eigenvectors [1]. The methods can be divided into two basically different groups, that is, variational and non-variational.

The variational approach is based on the fact that the Rayleigh quotient

$$\rho(c) = c^T A c / c^T c$$

has a minimum at the eigenvector corresponding to the lowest eigenvalue. An iterative algorithm has been developed by Nesbet and others [2] to find the *lowest* eigenvalue and its eigenvector. Shavitt *et al.* [3] have generalized this further to obtain higher roots but it requires all eigenvalues below the one desired. Another important development has been made by Brändas and Goscinski [4] who use a combined method of the variational principle and the perturbation theory.

The non-variational approach has also been given by several other authors. Iwata and Freed [5] have proposed a method by using Löwdin's matrix partitioning technique [6]. A very powerful method for this problem has been devised by David-

0021-9991/82/020246-09$02.00/0

son [7], which is an extension of the Lanczos' algorithm [8]. An important feature of Davidson's method is that a direct determination of *desired* higher roots is possible without a knowledge of lower ones. Another advantage is that there is no difficulty even in the case of degeneracy of roots. The potentialities of this method have been realized by the recent large scale configuration interaction (CI) calculation of a wavefunction for atoms and molecules [9].

On the other hand, known methods for solving a large nonsymmetric matrix are very limited. Most of the practical techniques presented so far are based on the power method and use a combination of iterative and direct methods for increasing the dominance of the eigenvalue of *maximum* modulus [10]. Recently the method of optimal relaxation proposed by Shavitt *et al.* [3] has been applied to the nonsymmetric problem of the equation of the motion method [11].

The purpose of a present paper is to extend the Davidson's method to nonsymmetric eigenvalue problems. Two algorithms have been presented to find *desired* eigenvectors and the associated eigenvalues. The methods have same advantages as those in Davidson's method and a convergence property is much superior to that of the power method. In the first method, left- and right-hand eigenvectors are sought in subspaces spanned by two sets of bi-orthogonal vectors. The second algorithm uses a sequence of orthonormal vectors to approximate eigenvectors. These two methods have been applied successfully to nonsymetric matrices of various sizes in the recent investigation of electron correlation problems for atoms and molecules by a cluster expansion theory [12].

II. Theoretical Background

The algorithms will be developed for the nonsymmetric problem

$$AX_i = \lambda_i X_i, \tag{1}$$

where A is a given real nonsymmetric matrix of order N, λ_i the ith eigenvalue and X_i the corresponding right-hand eigenvector. The left-hand eigenvector $\bar{X}_i$ belonging to the same eigenvalue satisfies the equation

$$A^H \bar{X}_i = \lambda_i \bar{X}_i. \tag{2}$$

A superscript H denotes a Hermitian transpose. Throughout this paper a bar above a vector or a matrix indicates that it is derived from A^H. We consider how to determine the kth eigenvalue whose eigenvectors have a desired structure.

Let $N \times m$ matrices

$$\bar{B}^{(m)} = (\bar{b}_1, \bar{b}_2, ..., \bar{b}_m) \qquad \text{and} \qquad B^{(m)} = (b_1, b_2, ..., b_m)$$

represent two sets of vectors which obey a bi-orthogonal relation

$$(\bar{B}^{(m)})^H B^{(m)} = I^{(m)}, \tag{3}$$

where $I^{(m)}$ is an $m \times m$ unit matrix. Assume that subspaces spanned by columns of $\bar{B}^{(m)}$ and $B^{(m)}$ contain a good approximation to desired left- and right-hand eigenvectors, $\bar{X}_k$ and X_k, respectively. These trial vectors are determined simultaneously in order to give a fast convergence to the desired eigenvectors. Since A is nonsymmetric, complex basis vectors may appear at some stage of the iteration even when one starts initially with real vectors. A projection of A onto subspaces gives a so-called interaction matrix

$$\tilde{A}^{(m)} = (\bar{B}^{(m)})^H A B^{(m)}. \tag{4}$$

Suppose that $\tilde{A}^{(m)}$ can be diagonalized such that

$$(C^m)^{-1} \tilde{A}^{(m)} C^{(m)} = (\bar{C}^{(m)})^H \tilde{A}^{(m)} C^{(m)} = \lambda^{(m)}, \tag{5}$$

where left- and right-hand eigenvectors of $\tilde{A}^{(m)}$ are bi-orthogonal

$$(\bar{C}^{(m)})^H C^{(m)} = I^{(m)}. \tag{6}$$

This involves a direct solution of a comparatively small $m \times m$ nonsymmetric eigenvalue problem. Equation (5) is equivalent to eigenvalue equations

$$(\tilde{A}^{(m)} - \lambda_k^{(m)})\, C_k^{(m)} = 0, \tag{7a}$$

$$((\tilde{A}^{(m)})^H - \lambda_k^{(m)})\, \bar{C}_k^{(m)} = 0. \tag{7b}$$

The approximation consists in replacing an eigenvalue problem for A by a problem of the same type of $\tilde{A}^{(m)}$. When $m = N$, eigenvalues of A are the same eigenvalues of $\tilde{A}^{(N)}$. Even when m is very much less than N, we could obtain desired eigenvalues of A to a good approximation if subspaces are well chosen. Suppose that the desired kth eigenvalue of the full matrix of A corresponds to the k'-th eigenvalue of $\tilde{A}^{(m)}$. Then expansions

$$\bar{X}_k^{(m)} = \bar{B}^{(m)} \bar{C}_{k'}^{(m)} \qquad \text{and} \qquad X_k^{(m)} = B^{(m)} C_{k'}^{(m)}$$

will converge to left- and right-hand eigenvectors of A, respectively.

We can improve the approximate eigenvalue $\lambda_{k'}^{(m)}$ by extending subspaces, that is, by adding new vectors $\bar{b}_{m+1}$ and b_{m+1} to the original sets. New vectors are generated such that residual vectors will converge to zero. Residual vectors $\bar{\xi}$ and ξ satisfy

$$(A - \lambda_{k'}^{(m)})(X_k^{(m)} + \xi) = 0, \tag{8a}$$

$$(A^H - \lambda_{k'}^{(m)})(\bar{X}_k^{(m)} + \bar{\xi}) = 0. \tag{8b}$$

Then we can approximate residual vectors as

$$\xi_{I,m+1} = (\lambda_{k'}^{(m)} - A_{II})^{-1} q_{I,m},$$

$$\bar{\xi}_{I,m+1} = (\lambda_{k'}^{(m)} - A_{II})^{-1} \bar{q}_{I,m}, \qquad I = 1, 2, \ldots, N, \tag{9}$$

with

$$q_m = (A - \lambda_{k'}^{(m)})\, X_k^{(m)},$$
$$\bar{q}_m = (A^H - \lambda_{k'}^{(m)})\, \bar{X}_k^{(m)}. \tag{10}$$

If $\lambda_{k'}^{(m)}$ and $X_k^{(m)}$ are an exact eigenvalue and eigenvector pair, then $q_m = 0$. Thus, the size of q_m measures the accuracy of $\lambda_{k'}^{(m)}$ and $X_k^{(m)}$.

It is convenient to bi-orthogonalize predicted vectors before a further round of iteration. Current vectors can be bi-orthogonalized according to

$$d_{m+1} = \left[\prod_{i=1}^{m} (I - b_i \bar{b}_i^H)\right] \xi_{m+1},$$
$$\bar{d}_{m+1} = \left[\prod_{i=1}^{m} (I - \bar{b}_i b_i^H)\right] \bar{\xi}_{m+1}. \tag{11}$$

Arbitrary scaling factors can be applied to each vector to give

$$\bar{b}_{m+1}^H b_{m+1} = 1. \tag{12}$$

In practice, $\bar{b}_{m+1}$ and b_{m+1} have been chosen such that

$$b_{m+1} = d_{m+1}/(g_{m+1})^{1/2}, \qquad \bar{b}_{m+1} = \bar{d}_{m+1}/(g_{m+1})^{1/2}, \tag{13}$$

where

$$g_{m+1} = \bar{d}_{m+1}^H d_{m+1}. \tag{14}$$

It is necessary to change the sign of one of the vectors if g_{m+1} becomes negative. However, this may only happen in the initial round of iteration.

Convergence is achieved if $\|q_m\|$ and $\|\bar{q}_m\|$ become less than a given threshold. Left- and right-hand eigenvectors obtained are bi-orthogonal

$$(\bar{X}^{(m)})^H X^{(m)} = (\bar{C}^{(m)})^H (\bar{B}^{(m)})^H B^{(m)} C^{(m)} = I \tag{15}$$

but they are not self-orthonormal. So it is necessary to renormalize eigenvectors after convergence is reached. As suggested by Davidson [7], when m becomes large, current sets of $\bar{B}^{(m)}\bar{C}_{k'}^{(m)}$ and $B^{(m)}C_{k'}^{(m)}$ can be taken as new initial sets and the calculation restarted. We call this a refreshment process.

If A is symmetric, two sets of $\bar{B}^{(m)}$ and $B^{(m)}$ are identical and the method becomes essentially Davidson's method since columns of $B^{(m)}$ itself are made orthogonal at each stage.

The present method is only based on the eigenvalue equations (1) and (2) (i.e., (8a) and (8b)). Thus, theoretically any eigenvalues and eigenvectors of the matrix can be found. In practice the convergence for each eigenvalue is obtained by means of the root-homing-procedure discussed by Butscher and Kammer [13]. The eigenvector

matrix $X^{(m)}$ is inspected after each iteration such that $X_{k'}^{(m)}$ has the largest overlap with the prediagonalization vector which represents the structure of the eigenvector wanted. The algorithm enforces convergence into just the eigenvalue whose eigenvector has the desired structure. The use of the root- homing procedure is essential for the calculation of higher roots and degenerate roots.

An alternative approach for solving the eigenvalue problem for nonsymmetric matrices is as follows. Consider a matrix $B^{(m)}$ having orthonormal vectors

$$B^{(m)} = (b_1, b_2, ..., b_m)$$

with

$$(B^{(m)})^H B^{(m)} = I^{(m)}. \tag{16}$$

Assume that the desired right-hand eigenvector, X_k, can be expressed to a good approximation as a linear combination of these trial vectors. In this case, we concentrate our attention on the right-hand eigenvalue problem of (1). Form the nonsymmetric intraction matrix of order m,

$$\tilde{A}^{(m)} = (B^{(m)})^H A B^{(m)} \tag{17}$$

by projecting A onto the subspace spanned by columns of $B^{(m)}$ and find a unitary transformation $U^{(m)}$ such that $T^{(m)}$ is upper triangular

$$T^{(m)} = (U^{(m)})^H \tilde{A}^{(m)} U^{(m)} \tag{18}$$

with diagonal elements $\lambda_i^{(m)}$. The matrix $T^{(N)}$ is similar to the full matrix A and its eigenvalues coincide with those of A. Now suppose that $T^{(m)}$ can be diagonalized by $P^{(m)}$

$$(P^{(m)})^{-1} T^{(m)} P^{(m)} = \lambda^{(m)}. \tag{19}$$

Let us define $C^{(m)} = U^{(m)} P^{(m)}$. Then, the column $X_k^{(m)} = B^{(m)} C_{k'}^{(m)}$ will converge to the desired right-hand eigenvector of A.

The subsequent vector b_{m+1} can be found by Schmidt-orthogonalization of a residual vector ξ_{m+1} to all previous vectors. The ξ_{m+1} is defined similarly as above by

$$\xi_{I,m+1} = (\lambda_{k'}^{(m)} - A_{II})^{-1} q_{I,m}, \qquad I = 1, 2, ..., N, \tag{20}$$

with

$$q_m = (A - \lambda_{k'}^{(m)}) X_k^{(m)} \tag{21}$$

and then orthogonalized with respect to the other vectors

$$\begin{aligned} d_{m+1} &= \left[\prod_{i=1}^{m} (I - b_i b_i^H) \right] \xi_{m+1}, \\ b_{m+1} &= d_{m+1} / \| d_{m+1} \|. \end{aligned} \tag{22}$$

The algorithm is the same as Davidson's except for solving a nonsymmetric interaction matrix. The method gives only the right-hand eigenvector but the expansion $\bar{X}_k^{(m)} = B^{(m)}(C^{(m)})_{k'}^{-H}$ at the end of finding the right-hand eigenvector may provide a good approximation for the corresponding left-hand eigenvector.

Two algorithms presented here can also be used for the general eigenvalue problems $AX = \lambda SX$ with a slight modification. Here S is a given metric which is real, symmetric and positive definite. Every $(A - \lambda_i)$, $\bar{b}_i^H b_j$, $b_i \bar{b}_i^H$,..., should be replaced by the corresponding $(A - \lambda_i S)$, $\bar{b}_i^H S b_j$, $b_i (S\bar{b}_i)^H$,..., respectively.

When several eigenvalues are wanted, it is very useful to extend the present methods to find several eigenvalues at a time. In this case several eigenvectors are expanded in a block sequence of orthonormal or bi-orthogonal vector basis. Each iteration yields several vectors simultaneously resulting in fast convergence to the desired eigenvectors of the matrix. These extended algorithms may have more powerful convergence properties than the original ones because the relative interference of vectors can be removed at each interation and poor convergence due to the near degeneracy of roots can be improved.

III. Test Results

Since the method requires only the concentration of Ab_i $(A^H \bar{b}_i)$ for any given vectors b_i $(\bar{b}_i)$ and the orthogonal or bi-orthogonal basis can be generated by sequential vector matrix multiplication, the algorithm is well suited to large nonsymmetric eigenvalue problems. In addition, the method can handle sparse matrices efficiently.

The present algorithms have been applied to several nonsymmetric matrices of various sizes in the recent investigation of electron correlation problems by a cluster expansion theory [12]. The methods turn out to be very effective and no convergence difficulties arise even for a direct determination of desired higher roots. As an example, we summarized in Tables I and II results of the convergence process for the lowest and second lowest roots of the non-variational SAC-CI matrix for the triplet A_2 state of H_2O (almost symmetric). Starting vectors are determined by the CI calculation. The convergence properties are similar to those in the symmetric case. The time per iteration of the the method with bi-orthogonal basis was 1.2 times as much as that of the Davidson's procedure for the corresponding symmetric matrix (variational SAC-CI matrix). The nonsymmetric algorithm often converges in slightly more iterations than the symmetric procedure. In these examples two algorithms (two sets of bi-orthogonal vectors and one set of orthogonal vectors) were found to have very similar convergence properties. In both cases, the convergence was obtained in a reasonable number of iterations, generally in fewer than 25 iterations with the convergence criteria $\|q_m\| < 10^{-6}$.

Second the method is compared with a full diagonalization. Consider the nonsymmetric matrix [14]

TABLE I

Convergence Process to the Lowest Root ($^3A_2$ State) of H_2O Matrix ($N = 1129$)[a,b]

M	E (au)	ΔE (au)	$\|q_M\|$
1	0.22005917	—	0.4058984
2	0.14810541	−0.07195377	0.1442040
3	0.13395789	−0.01414752	0.0765616
4	0.13041188	−0.00354600	0.0440078
5	0.12923475	−0.00127713	0.0234006
6	0.12881130	−0.00032345	0.0128173
7	0.12871985	−0.00009145	0.0082276
8	0.12866675	−0.00005310	0.0049747
9	0.12864683	−0.00001992	0.0033006
10	0.12863951	−0.00000732	0.0018201
11	0.12863712	−0.00000238	0.0009033
12	0.12863679	−0.00000033	0.0004092
13	0.12863667	−0.00000012	0.0002910
14	0.12863662	−0.00000005	0.0001624
15	0.12863660	−0.00000002	0.0000911
16	0.12863659	−0.00000001	0.0000504
17	0.12863659	0.00000000	0.0000357
18	0.12863659	0.00000000	0.0000218
19	0.12863659	0.00000000	0.0000140
20	0.12863659	0.00000000	0.0000076
21	0.12863659	0.00000000	0.0000053
22	0.12863659	0.00000000	0.0000024
23	0.12863659	0.00000000	0.0000016
24	0.12863659	0.00000000	0.0000010

[a] The energies are relative to the Hartree–Fock energy, −76.03731 au.

[b] The right-hand eigenvector; $0.8355 \times \Psi(5 \to 7) + 0.4482 \times \Psi(5 \to 10) - 0.2402 \times \Psi(5 \to 14) + \cdots$. The left-hand eigenvector; $0.8267 \times \Psi(5 \to 7) + 0.4421 \times \Psi(5 \to 10) - 0.2360 \times \Psi(5 \to 14) + \cdots$.

$$\begin{aligned} A_{ij} &= i\delta_{ij} - (i - j - k^2) \qquad 1 \leqslant j \leqslant k, \\ &= i\delta_{ij} + (i - j - k^2) \qquad k + 1 \leqslant j \leqslant N, \end{aligned} \tag{23}$$

where

$$N = 2k.$$

The matrix A has real eigenvalues $1, 2, \ldots, N$, right-hand eigenvectors which are columns of $(I + UV^T)$ and left-hand eigenvectors which are rows of $(I - UV^T)$. Here $U = (1, 1, \ldots, 1)$ and $V^T = (1, 1, \ldots, 1, -1, -1, \ldots, -1)$ with k components of 1 and k components of -1. First a full diagonalization was done by a double QR method on a FACOM M200 computer and all eigenvalues were calculated. It takes 2.5 and 17.7 sec to solve the matrix of $N = 100$ and of $N = 200$, respectively. Second the

TABLE II

Convergence Process to the Second Lowest Root ($^3A_2$ State) of H_2O Matrix ($N = 1129$)[a,b]

M	E (au)	ΔE (au)	$\|q_M\|$
1	0.29430328	—	0.3149432
2	0.24700941	−0.04729387	0.1447755
3	0.23429804	−0.01271137	0.0760444
4	0.23081265	−0.00348540	0.0425022
5	0.22983184	−0.00098080	0.0221159
6	0.22952958	−0.00030226	0.0153357
7	0.22943993	−0.00008965	0.0082444
8	0.22938817	−0.00005157	0.0062884
9	0.22936302	−0.00002515	0.0037299
10	0.22935281	−0.00001021	0.0023221
11	0.22934977	−0.00000304	0.0012189
12	0.22934897	−0.00000080	0.0007334
13	0.22934853	−0.00000044	0.0005549
14	0.22934827	−0.00000025	0.0003860
15	0.22934819	−0.00000009	0.0002177
16	0.22934816	−0.00000003	0.0001788
17	0.22934815	−0.00000001	0.0000940
18	0.22934814	−0.00000001	0.0000734
19	0.22934814	0.00000000	0.0000389
20	0.22934814	0.00000000	0.0000238
21	0.22934814	0.00000000	0.0000179
22	0.22934814	0.00000000	0.0000113
23	0.22934814	0.00000000	0.0000084

[a] The energies are relative to the Hartree–Fock energy, −76.03731 au.

[b] The right-hand eigenvector;$0.4906 \times \Psi(5 \to 7) - 0.5862 \times \Psi(5 \to 10) + 0.6033 \times \Psi(5 \to 14) + \cdots$. The left-hand eigenvector; $0.4873 \times \Psi(5 \to 7) - 0.5793 \times \Psi(5 \to 10) + 0.5973 \times \Psi(5 \to 14) + \cdots$.

present algorithm with orthogonal vectors was applied and the four lowest roots were computed simultaneously. The four starting vectors are unit vectors with 1.0 in position of 1, 2, 3 and 4, respectively. The computing time for the matrix of $N = 100$ is 4.5 seconds and 10.7 sec for the matrix of $N = 200$. Thus, the present algorithm is preferred for large N if one wants only a few eigenvalues.

To conclude, Davidson's algorithm for solving large symmetric matrices is generalized to the nonsymmetric cases. The method is based on the non-variational eigenvalue equation. Therefore, the algorithm can be expected to converge particularly well for the eigenvalue whose eigenvector has a desired structure and there is no difficulty even in the case of degeneracy of roots.

ACKNOWLEDGMENTS

The authors wish to thank Drs. P. Dacre and S. Iwata for useful discussions. We also thank Professors T. Yonezawa and H. Kato for the encouragement. This study was supported in part by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science and Culture.

REFERENCES

1. J. H. WILKINSON, "The Algebraic Eigenvalue Problem," Oxford Univ. Press, London, 1965; G. W. STEWERT, *in* "Information Processing 74," Proceedings of IFIP Congress 74, Stockholm, p. 666, North-Holland, Amsterdam, 1974; I. SHAVITT, *in* "Modern Theoretical Chemistry" (H. F. Schaefer, III, Ed.), Vol. 3, p. 189, Plenum, New York, 1977.
2. R. K. NESBET, *J. Chem. Phys.* **43** (1965), 311; I. SHAVITT, *J. Comput. Phys.* **6** (1970), 124; C. F. BENDER AND E. R. DAVIDSON, *Phys. Rev.* **183** (1969), 23.
3. I. SHAVITT, C. F. BENDER, A. PIPANO, AND R. P. HOSTENY, *J. Comput. Phys.* **11** (1973), 90.
4. E. BRÄNDAS AND O. GOSCINSKI, *Phys. Rev. A* **1** (1970), 552.
5. S. IWATA AND K. F. FREED, *Chem. Phys.* **11** (1975), 433. See also G. A. SEGAL AND R. W. WETMORE, *Chem. Phys. Lett.* **32** (1975), 556.
6. P. O. LÖWDIN, *Advan. Chem. Phys.* **2** (1959), 207.
7. E. R. DAVIDSON, *J. Comput. Phys.* **17** (1975), 87.
8. C. LANCZOS, *J. Res. Nat. Bur. Stand.* **45** (1950), 255.
3. B. ROOS AND P. SIEGBAHN, *in* "Modern Theoretical Chemistry" (H. F. Schaefer, III, Ed.), Vol. 3, p. 277, Plenum, New York, 1977.
10. M. CLINT AND A. JENNINGS, *J. Inst. Maths. Appl.* **8** (1971), 111.
11. J. P. FLAMENT AND H. G. GERVAIS, *Internat. J. Quantum Chem.* **16** (1979), 1347.
12. K. HIRAO AND H. NAKATSUJI, *Chem. Phys. Lett.* **79** (1981), 292; H. NAKATSUJI AND K. HIRAO, *Internat. J. Quantum Chem.*, in press; H. NAKATSUJI, K. OHTA, AND K. HIRAO, *J. Chem. Phys.* **75** (1981), 2952.
13. W. BUTSCHER AND W. E. KAMMER, *J. Comput. Phys.* **20** (1976), 313.
14. R. T. GREGORY AND D. L. KARNEY, "A Collection of Matrices for Testing Computational Algorithms," Wiley–Interscience, New York, 1969.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium