JOURNAL OF COMPUTATIONAL PHYSICS 80, 232–242 (1989)

Iterative Methods for the Solution of Large Systems of Linear Equations

K. HIRAO

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

Received November 30, 1987; revised February 29, 1988

New methods are described for the solution of large systems of linear equations for nonsymmetric matrices which arise in molecular calculations. The methods use the interaction matrix procedure in Lanczos-like algorithm previously applied to the eigenvalue problems. The efficiency of the methods is examined using test calculations. © 1989 Academic Press, Inc.

I. INTRODUCTION

The solution of linear systems of large dimensionality arises in many molecular calculations. Our symmetry adapted cluster (SAC) theory [1] for calculating electron correlations requires the solution of linear equations. In the SAC approach, the exact wave function Ψ is expressed as a cluster expansion in the neighborhood of an independent particle wave function Φ,

$$\Psi = e^S \Phi = \exp\left(\sum_i S_i\right) \| \varphi_1 \bar{\varphi}_1 \cdots \varphi_N \bar{\varphi}_N \| .$$

Here, S_i is a linked cluster operator which produces i-fold symmetry adapted excited configurations when operating on Φ. The Schrödinger equation, $H\Psi = E\Psi$, is projected against a sufficient set of excited functions to generate a series of coupled linear equations

$$\langle S_i | e^{-S} H e^S | \rangle = 0$$

or, in the matrix notation,

$$\mathbf{A} + \mathbf{B}\alpha + \mathbf{C}\alpha\alpha = 0.$$

Here, the matrices **A**, **B**, and **C** arise, respectively, from the zero-, one-, and two-commutator terms and $\{\alpha_i\}$ are now viewed as the unknown coefficients. The total energy is given by

$$\langle | e^{-S} H e^S | \rangle = E$$

0021-9991/89 $3.00

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.

or, in the matrix notation,

$$E = \mathbf{A}\alpha.$$

The solution of linear systems of small dimensionality is carried out by Gauss elimination for arbitrary matrices or Cholesky decomposition for symmetric, positive definite matrices. These approaches become unsuitable, however, for large systems, whose coefficient matrix cannot be kept in fast core.

In such a case, iterative algorithms, which require only one or a few matrix elements at a time in core and converge reasonably fast, are preferable. The most appropriate procedure must satisfy the following conditions: (1) It preserves the sparseness of the coefficient matrix throughout the calculation. (2) It can easily be handled in pieces. Matrices of very large dimension must be kept on external storage in case they are not extremely sparse. Thus, the elements of one row (or one column) are needed at a time and different rows (columns) are needed in a sequential order. (3) The original coefficient matrix is not modified in the course of the calculation. (4) The expense of the arithmetic operation increases with twice the effective order of the coefficient matrix times the number of iterations needed to achieve convergence.

Recently, two algorithms have been proposed. The method of Pople, Krishnan, Schlegel, and Binkley [2] has been formulated for coupled Hartree–Fock theory and successfully applied in Newton–Raphson MCSCF and CI perturbation theory. The reduced linear equation method of Purvice and Bartlett [3] has been exploited in their coupled cluster calculations. It has been proved [4] that these two methods are different version of the conjugate gradient method [5]. When the coefficient matrix is symmetric, the conjugate gradient method must be one of the most efficient procedures. Unfortunately when the coefficient matrix is nonsymmetric, the conjugate gradient method cannot be applied directly because it is based on the variational theorem and its convergence is often too slow.

In view of the importance of the algorithm in correlation problems and in molecular properties, we derive alternative methods for a solution of a large system of linear equations. We have adapted the interaction matrix procedure previously applied to the eigenvalue problems by Davidson *et al.* [6]. The solution vector is sought in the subspace spanned by a Krylov sequence. The improved trial vector can be found by solving a linear problem of small order involving the interaction matrix in Lanczos-like algorithm. The method has been applied successfully to linear problems of various sizes of the recent investigation of the electron correlation problems for molecules by the cluster expansion theory.

II. Theoretical Background

The algorithm will be developed for the general linear equation

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

where $\mathbf{x}$ and $\mathbf{b}$ are column vectors and $\mathbf{A}$ is a diagonally dominant non-singular matrix of order N. We also consider the left-hand linear equation

$$\mathbf{A}^{\mathrm{T}}\mathbf{y} = \mathbf{b}, \tag{2}$$

where the superscript T denotes the transpose. In the CI and SAC calculations, the correlation energy is expressed as

$$E = -\mathbf{b}^{\mathrm{T}}\mathbf{x} = -\mathbf{b}^{\mathrm{T}}\mathbf{y}. \tag{3}$$

Let the $N * m$ matrices

$$\mathbf{U}^{(m)} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m) \qquad \text{and} \qquad \mathbf{V}^{(m)} = (\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m)$$

represent two sets of vectors which obey the bi-orthogonal reltion

$$(\mathbf{V}^{(m)})^{\mathrm{T}}\,\mathbf{U}^{(m)} = \mathbf{I}^{(m)}, \tag{4}$$

where $\mathbf{I}^{(m)}$ is a $m * m$ unit matrix. Assume that the subspaces spanned by the columns of $\mathbf{U}^{(m)}$ and $\mathbf{V}^{(m)}$ contain a good approximation to the solution vectors, $\mathbf{x}$ and $\mathbf{y}$, respectively. These trial vectors are determined simultaneously as to give a fast convergence to the desired solution vectors. The projection of $\mathbf{A}$ onto subspaces gives the so-called interaction matrix defined by

$$\tilde{\mathbf{A}}^{(m)} = (\mathbf{V}^{(m)})^{\mathrm{T}}\,\mathbf{A}\mathbf{U}^{(m)}. \tag{5}$$

Suppose that the linear problems on the subspaces can be solved such that

$$\tilde{\mathbf{A}}^{(m)}\tilde{\mathbf{x}}(m) = \tilde{\mathbf{b}}^{(m)}, \qquad (\tilde{\mathbf{A}}^{(m)})^{\mathrm{T}}\,\tilde{\mathbf{y}}^{(m)} = \tilde{\mathbf{c}}^{(m)}, \tag{6}$$

where

$$\tilde{\mathbf{b}}^{(m)} = (\mathbf{V}^{(m)})^{\mathrm{T}}\,\mathbf{b}, \qquad \tilde{\mathbf{c}}^{(m)} = (\mathbf{U}^{(m)})^{\mathrm{T}}\,\mathbf{b}. \tag{7}$$

This involves the direct solution of comparatively small $m * m$ linear problems. Equations in (6) are easily seen to be equivalent to

$$\begin{aligned} (\mathbf{V}^{(m)})^{\mathrm{T}}\,\mathbf{A}\mathbf{U}^{(m)}\,\tilde{\mathbf{x}}^{(m)} &= (\mathbf{V}^{(m)})^{\mathrm{T}}\,\mathbf{b} \\ (\mathbf{U}^{(m)})^{\mathrm{T}}\,\mathbf{A}^{\mathrm{T}}\mathbf{V}^{(m)}\tilde{\mathbf{y}}^{(m)} &= (\mathbf{U}^{(m)})^{\mathrm{T}}\,\mathbf{b}. \end{aligned} \tag{8}$$

The approximation consists in replacing the linear problems for (1) and (2) by problems of the same type of (6). When $m = N$, two linear problems become identical. Even when m is much smaller than N, we could obtain the desired solutions of (1) and (2) to a good approximation if the subspaces are well chosen.

We can improve the approximate solutions $\mathbf{x}^{(m)}$ and $\mathbf{y}^{(m)}$,

$$\mathbf{x}^{(m)} = \mathbf{U}^{(m)}\tilde{\mathbf{x}}^{(m)}, \qquad \mathbf{y}^{(m)} = \mathbf{V}^{(m)}\tilde{\mathbf{y}} \tag{9}$$

by expanding the subspaces, that is, by adding new vectors $\mathbf{u}^{(m+1)}$ and $\mathbf{v}^{(m+1)}$ to the original sets. The new vectors are generated such that the residual vectors will converge to zero. The residual vectors $\mathbf{q}_r$ and $\mathbf{q}_l$ satisfy

$$\mathbf{A}(\mathbf{x}^{(m)} + \mathbf{q}_r^{(m)}) = \mathbf{b}, \qquad \mathbf{A}^{\mathrm{T}}(\mathbf{y}^{(m)} + \mathbf{q}_l^{(m)}) = \mathbf{b}. \tag{10}$$

Then we can approximate

$$\begin{aligned} \mathbf{q}_r^{(m+1)} &= \mathbf{D}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{U}^{(m)}\tilde{\mathbf{x}}^{(m)}) \\ \mathbf{q}_l^{(m+1)} &= \mathbf{D}^{-1}(\mathbf{b} - \mathbf{A}^{\mathrm{T}}\mathbf{V}^{(m)}\tilde{\mathbf{y}}^{(m)}), \end{aligned} \tag{11}$$

where $\mathbf{D}$ represents the diagonal part of $\mathbf{A}$, If $\mathbf{x}^{(m)}$ and $\mathbf{y}^{(m)}$ are exact solutions, then $\mathbf{q}^{(m)} = 0$. Thus, the size of $\mathbf{q}^{(m)}$ measures the accuracy of the solutions. The energy in the mth iteration is given by

$$E^{(m)} = -\tilde{\mathbf{b}}^{\mathrm{T}}\tilde{\mathbf{x}}^{(m)} = -\tilde{\mathbf{c}}^{\mathrm{T}}\tilde{\mathbf{y}}^{(m)}. \tag{12}$$

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

$$\begin{aligned} \mathbf{d}_r^{m+1} &= \left[\prod_{i=1}^{n} (\mathbf{I} - \mathbf{u}_i(\mathbf{v}_i)^{\mathrm{T}}\right] \mathbf{q}_r^{(m)} \\ \mathbf{d}_l^{m+1} &= \left[\prod_{i=1}^{m} (\mathbf{I} - \mathbf{v}_i(\mathbf{u}_i)^{\mathrm{T}}\right] \mathbf{q}_l^{(m)}. \end{aligned} \tag{13}$$

Arbitrary scaling factors can be applied to each vector to give

$$(\mathbf{v}^{(m+1)})^{\mathrm{T}}\,\mathbf{u}^{(m+1)} = 1 \tag{14}$$

in practice, $\mathbf{u}^{(m+1)}$ and $\mathbf{v}^{(m+1)}$ can be chosen such that

$$\begin{aligned} \mathbf{u}^{(m+1)} &= \mathbf{d}_r^{(m+1)}/\|\mathbf{g}^{m+1}\| \\ \mathbf{v}^{(m+1)} &= \mathbf{d}_l^{(m+1)}/\|\mathbf{g}^{m+1}\|, \end{aligned} \tag{15}$$

where

$$\mathbf{g}^{(m+1)} = (\mathbf{d}_l^{m+1})^{\mathrm{T}}\,\mathbf{d}_r^{m+1}. \tag{16}$$

It is necessary to change the sign of one of the vectors if $\mathbf{g}^{m+1}$ becomes negative. However, the bi-orthogonalization procedure is not essential since the trial vectors $\mathbf{u}^{(m)}$ and $\mathbf{v}^{(m)}$ constructed as above are Krylov sequences and are linearly independent.

If the $\mathbf{A}$ is symmetric, the two sets $\mathbf{U}^{(m)}$ and $\mathbf{V}^{(m)}$ are identical and $\mathbf{U}^{(m)}$ is made orthogonal at each stage.

Convergence is achieved if $\|\mathbf{q}^{(m)}\|$ becomes less than a given threshold. When m becomes inconveniently large, the current sets of $\mathbf{x}^{(m)}$ and $\mathbf{y}^{(m)}$ in (9) can be taken as the new initial vectors and the calculation is restarted. This is an update process.

An alternative approach is possible with the orthonormal basis. Consider the matrix $\mathbf{U}^{(m)}$ having the orthonormal vectors

$$\mathbf{U}^{(m)} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m) \tag{17}$$

with

$$(\mathbf{U}^{(m)})^{\mathrm{T}} \mathbf{U}^{(m)} = \mathbf{I}^{(m)}. \tag{18}$$

Assume that the desired solution $\mathbf{x}$ can be expressed to a good approximation as a linear combination of these trial vectors. In this case, we will pay attention to only the right-hand linear equation of (1). Form the interaction matrix of order m,

$$\tilde{\mathbf{A}}^{(m)} = (\mathbf{U}^{(m)})^{\mathrm{T}} \mathbf{A} \mathbf{U}^{(m)}. \tag{19}$$

Suppose that the linear problem on the subspace can be solved such that

$$\tilde{\mathbf{A}}^{(m)} \tilde{\mathbf{x}}^{(m)} = \tilde{\mathbf{b}}^{(m)}, \tag{20}$$

where

$$\tilde{\mathbf{b}}^{(m)} = (\mathbf{U}^{(m)})^{\mathrm{T}} \mathbf{b}. \tag{21}$$

The subsequent vector $\mathbf{u}^{(m+1)}$ can be found by Schmidt-orthogonalization of the residual vector $\mathbf{q}^{(m+1)}$ to all the previous vectors. The residual vector $\mathbf{q}^{(m+1)}$ is defined similarly as above by

$$\mathbf{q}_r^{(m+1)} = \mathbf{D}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{U}^{(m)} \tilde{\mathbf{x}}^{(m)}). \tag{22}$$

It is useful to orthogonalize $\mathbf{u}^{(m+1)}$ to the other vectors,

$$\begin{aligned} \mathbf{d}^{(m+1)} &= \left[\prod_{i=1}^{N} (\mathbf{I} - \mathbf{u}_i (\mathbf{u}_i)^{\mathrm{T}}) \right] \mathbf{q}^{(m+1)} \\ \mathbf{b}^{(m+1)} &= \mathbf{d}^{(m+1)} / \| \mathbf{d}^{(m+1)} \|. \end{aligned} \tag{23}$$

This approach becomes equivalent to the reduced linear equation method [3] when we precondition the linear equation and use the Jacobi recursion relation.

The convergence rate depends on the diagonal dominance of $\mathbf{A}$. One of the advantages of this method is that the elements of $\mathbf{A}$ need not be used in any specific sequence. $\mathbf{A}$ is defined as an operator without giving its explicit matrix representation, which makes this method useful for those direct SAC techniques [7] which proceed directly from orbital integrals to the solution without forming an explicit $\mathbf{A}$ matrix. In addition, the method can handle the sparsity of the matrix efficiently.

When $\mathbf{b}$ is a block sequence of column vectors and several solutions are wanted, it is very useful to extend the present method to find several solutions at a time. In this case several solutions are expanded in a block sequence of orthonormal (bi-orthogonal) vectors simultaneously as to give a fast convergence to the desired solutions. These extended algorithms may have more powerful convergence properties than the original one because the relative interference of the solution vectors can be removed at each iteration.

III. Computational Details

(i) Algorithm. With two sets of bi-orthogonal vectors.

A. Select zero-order two sets of bi-orthogonal vectors

$$\mathbf{U}^{(m)} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m) \qquad \text{and} \qquad \mathbf{V}^{(m)} = (\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m)$$

spanning the dominant components of the full solution of the right-hand and left-hand linear equations. Form and save $\mathbf{A}\mathbf{u}_i$, $\mathbf{A}^{\mathrm{T}}\mathbf{v}_i = (\mathbf{v}_i^{\mathrm{T}}\mathbf{A})^{\mathrm{T}}$, $\mathbf{b}_i = (\mathbf{v}_i)^{\mathrm{T}}\mathbf{b}$, $\tilde{\mathbf{c}}_i = (\mathbf{u}_i)^{\mathrm{T}}\mathbf{b}$ and $\mathbf{A}_{ij} = (\mathbf{v}_i)^{\mathrm{T}}\mathbf{A}\mathbf{u}_j$ $(1 \leqslant i, j \leqslant m)$. Solve the small linear problem $\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ and $\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{y}} = \tilde{\mathbf{c}}$ using a direct method, say, the Gauss elimination or the Cholesky decomposition.

B. Form

$$\mathbf{q}_r^M = \sum_{i=1}^{M} \mathbf{D}^{-1}[\mathbf{b}_i - (\mathbf{A}\mathbf{u}_i)\,\tilde{\mathbf{x}}_i]$$

$$\mathbf{q}_l^M = \sum_{i=1}^{M} \mathbf{D}^{-1}[\mathbf{b}_i - (\mathbf{A}^{\mathrm{T}}\mathbf{v}_i)\,\tilde{\mathbf{y}}_i].$$

Here M is the dimension of $\tilde{\mathbf{A}}$.

C. Form $\|\mathbf{q}_r^M\|$ and $\|\mathbf{q}_l^M\|$ and check convergence.

D. Form

$$\mathbf{d}_r^{M+1} = \left[\prod_{i=1}^{M} (\mathbf{I} - \mathbf{u}_i(\mathbf{v}_i)^{\mathrm{T}})\right]\mathbf{q}_r^M$$

$$\mathbf{d}_l^{M+1} = \left[\prod_{i=1}^{M} (\mathbf{I} - \mathbf{v}_i(\mathbf{u}_i)^{\mathrm{T}})\right]\mathbf{q}_l^M.$$

E. Form and save

$$\mathbf{u}_{M+1} = \mathbf{d}_r^{M+1}/[\mathbf{d}_r^{M+1}\mathbf{d}_l^{M+1}]^{1/2}$$

$$\mathbf{v}_{M+1} = \mathbf{d}_l^{M+1}/[\mathbf{d}_r^{M+1}\mathbf{d}_l^{M+1}]^{1/2}.$$

F. Form and save $\mathbf{A}\mathbf{u}_{M+1}$, $\mathbf{A}^{\mathrm{T}}\mathbf{v}_{M+1}$, $\tilde{\mathbf{b}}_{M+1} = (\mathbf{v}_{M+1})^{\mathrm{T}}\mathbf{b}$, and $\tilde{\mathbf{c}}_{M+1} = (\mathbf{u}_{M+1})^{\mathrm{T}}\mathbf{b}$.

G. Form and save

$$\tilde{\mathbf{A}}_{M+1,M+1} = (\mathbf{v}_{M+1})^{\mathrm{T}}\mathbf{A}\mathbf{u}_{M+1}$$

$$\tilde{\mathbf{A}}_{i,M+1} = (\mathbf{v}_i)^{\mathrm{T}}\mathbf{A}\mathbf{u}_{M+1}$$

$$\tilde{\mathbf{A}}_{M+1,i} = (\mathbf{v}_{M+1})^{\mathrm{T}}\mathbf{A}\mathbf{u}_i, \qquad i = 1, 2, ..., M.$$

H. Solve the $(M+1)$-dimensional linear equations $\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ and $\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{y}} = \tilde{\mathbf{c}}$ and obtain solutions $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$. Update the vector if necessary and return to step B.

When the convergence is achieved, the solution vector is given by

$$\mathbf{x} = \sum_{i=1}^{M} \mathbf{v}_i \tilde{\mathbf{x}}_i, \qquad \mathbf{y} = \sum_{i=1}^{M} \mathbf{u}_i \tilde{\mathbf{y}}_i.$$

Steps D and E can be skipped. In this case, $\mathbf{q}_r^{(M)}$ and $\mathbf{q}_l^{(M)}$ become new trial vectors.

(ii) ALGORITHM. With orthonormal vectors.

A. Select a zero-order orthonormal subspace

$$\mathbf{U}^{(m)} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m)$$

spanning the dominant components of the full solution of the linear equation. Form and save $\mathbf{Au}_i$, $\tilde{\mathbf{b}}_i = (\mathbf{u}_i)^{\mathrm{T}} \mathbf{b}$, and $\tilde{\mathbf{A}}_{ij} = (\mathbf{u}_i)^{\mathrm{T}} \mathbf{Au}_j$ $(1 \leqslant i, j \leqslant m)$. Solve the small linear problem $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$.

B. Form

$$\mathbf{q}^M = \sum_{i=1}^{M} \mathbf{D}^{-1}[\mathbf{b}_i - (\mathbf{Ab}_i)\, \tilde{\mathbf{x}}_i].$$

Here M is the dimension of $\tilde{\mathbf{A}}$.

C. Form $\|\mathbf{q}^M\|$ and check convergence.

D. Form

$$\mathbf{d}^{M+1} = \left[\prod_{i=1}^{M} (\mathbf{I} - \mathbf{u}_i(\mathbf{u}_i)^{\mathrm{T}}) \right] \mathbf{q}^M.$$

E. Form and save

$$\mathbf{u}_{M+1} = \mathbf{d}^{M+1} / \|\mathbf{d}^{M+1}\|.$$

F. Form and save $\mathbf{Au}_{M+1}$ and $\tilde{\mathbf{b}}_{M+1} = (\mathbf{u}_{M+1})^{\mathrm{T}} \mathbf{b}$.

G. Form and save

$$\tilde{\mathbf{A}}_{M+1,\,M+1} = (\mathbf{u}_{M+1})^{\mathrm{T}} \mathbf{Au}_{M+1}$$

$$\tilde{\mathbf{A}}_{i,\,M+1} = (\mathbf{u}_i)^{\mathrm{T}} \mathbf{Au}_{M+1}$$

$$\tilde{\mathbf{A}}_{M+1,\,i} = (\mathbf{u}_{M+1})^{\mathrm{T}} \mathbf{Au}_i, \quad i = 1, 2, ..., M.$$

H. Solve the $(M+1)$-dimensional linear equation $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ and obtain the solution $\tilde{\mathbf{x}}$. Update the vector if necessary and return to step B.

When the convergence is obtained, the solution vector is given by

$$\mathbf{x} = \sum_{i=1}^{M} \mathbf{u}_i \tilde{\mathbf{x}}_i.$$

Steps D and E can be skipped. In this case, $\mathbf{q}^{(M)}$ becomes a new trial vector.

IV. Test Results

Since the method requires only to form $\mathbf{Au}_i(\mathbf{A}^T\mathbf{v}_i)$ and $\mathbf{u}_i^T\mathbf{b}$ $(\mathbf{v}_i^T\mathbf{b})$ for any given vector $\mathbf{u}_i$ $(\mathbf{v}_i)$ and the orthogonal (bi-orthogonal) basis can be generated by sequential vector matrix multiplication, the algorithm is well suited to large or gigantic linear problems.

The present algorithm has been applied to several linear problems of various sizes in the recent investigation of the electron correlation problems by the cluster expansion of the wavefunction. We found that the present method is very effective and no convergence difficulties arise. As an example, we summarized in Tables I and II the results of the convergence process for the ground states of H_2O SAC problems with single, double, and triple excitations at two geometries, at the equilibrium bond distance, R_e and at the stretched bond distance, $2.0 * R_e$. The coefficient matrices are both nonsymmetric with dimension $N = 1041$ and 1052, respec-

TABLE I

The Solution of 1041-Dimensional SAC problem for H_2O with $R = R_e$ Using Double Zeta Basis Set[a,b]

Cycle number	Orthonormal basis		Bi-orthogonal basis		
	- E (au)	$\|q\|$	- E (au)	$\|q_r\|$	$\|q_l\|$
1	0.14719229	0.01350357	0.14718667	0.01350357	0.01248906
2	0.14719935	0.00382285	0.14719987	0.00389647	0.00220823
3	0.14720137	0.00157605	0.14720182	0.00139455	0.00096340
4	0.14720233	0.00069811	0.14720232	0.00064449	0.00040754
5	0.14720254	0.00029178	0.14720245	0.00027922	0.00019912
6	0.14720250	0.00012399	0.14720247	0.00010811	0.00007535
7	0.14720246	0.00005966	0.14720247	0.00005194	0.00003426
8	0.14720247	0.00002945	0.14720247	0.00002819	0.00001598
9	0.14720248	0.00001563	0.14720247	0.00001355	0.00000852
10	0.14720247	0.00000848	0.14720247	0.00000799	0.00000457
11	0.14720247	0.00000370	0.14720247	0.00000386	0.00000202
12	0.14720247	0.00000180	0.14720247	0.00000181	0.00000097
13	0.14720247	0.00000099	0.14720247	0.00000087	0.00000054
14	0.14720247	0.00000047	0.14720247	0.00000050	0.00000029
15	0.14720247	0.00000023	0.14720247	0.00000035	0.00000019
16	0.14720247	0.00000010	0.14720247	0.00000038	0.00000020
17	0.14720247	0.00000005	0.14720247	0.00000067	0.00000038
18	0.14720247	0.00000002	0.14720247	0.00000007	0.00000004
19	0.14720247	0.00000001	0.14720247	0.00000002	0.00000001
20	0.14720247	0.000000005	0.14720247	0.000000008	0.000000004

[a] The energies are relative to the SCF energy, -76.00984 au.

[b] It takes 40 iterations to achieve an 8-digit accuracy of $-E$ and 52 iterations to achieve $\|\mathbf{q}\| < 10^{-8}$ by a preconditioned conjugate gradient method.

TABLE II

The Solution of 1052-Dimensional SAC Problem for H_2O with $R = 2.0 * R_e$ Using Double Zeta Basis Set[a,b]

Cycle number	Orthonormal basis -E (au)	Orthonormal basis $\|q\|$	Bi-orthogonal basis -E (au)	Bi-orthogonal basis $\|q_r\|$	Bi-orthogonal basis $\|q_\ell\|$
1	0.33209822	0.15591557	0.33316821	0.15591557	0.08067728
2	0.33389499	0.07856317	0.33307836	0.11311221	0.08128108
3	0.33465386	0.03857268	0.33472402	0.08868100	0.05897184
4	0.33472254	0.01604181	0.33467527	0.04036891	0.03086299
5	0.33473138	0.00885434	0.33475697	0.01950013	0.01635421
6	0.33473468	0.00418533	0.33475657	0.01844502	0.01699863
7	0.33474718	0.00205262	0.33476277	0.01441450	0.01306801
8	0.33475621	0.00103402	0.33476336	0.00185523	0.00199128
9	0.33476173	0.00055670	0.33476378	0.00180426	0.00192779
10	0.33476366	0.00031799	0.33476383	0.00093607	0.00073097
11	0.33476438	0.00014297	0.33476383	0.00030991	0.00027678
12	0.33476427	0.00006449	0.33476384	0.00125184	0.00109523
13	0.33476393	0.00003396	0.33476384	0.00011048	0.00011070
14	0.33476395	0.00001920	0.33476384	0.00008223	0.00007033
15	0.33476385	0.00001025	0.33476384	0.00003208	0.00003969
16	0.33476383	0.00000479	0.33476384	0.00001564	0.00001631
17	0.33476384	0.00000217	0.33476384	0.00002664	0.00003347
18	0.33476384	0.00000112	0.33476384	0.00000421	0.00000411
19	0.33476384	0.00000054	0.33476384	0.00000614	0.00000627
20	0.33476384	0.00000027	0.33476384	0.00000622	0.00000634
21	0.33476384	0.00000015	0.33476384	0.00000227	0.00000204
22	0.33476384	0.00000009	0.33476384	0.00000035	0.00000025
23	0.33476384	0.00000005	0.33476384	0.00000018	0.00000011
24	0.33476384	0.00000002	0.33476384	0.00000022	0.00000018
25	0.33476384	0.00000001	0.33476384	0.00000006	0.00000005
26	0.33476384	0.000000007	0.33476384	0.00000004	0.00000003
27			0.33476384	0.00000002	0.00000001
28			0.33476384	0.000000005	0.000000004

[a] The energies are relative to the SCF energies, -75.59519 au.

[b] It takes 64 iterations to achieve an 8-digit accuracy of $-E$ and 75 iterations to reach $\|\mathbf{q}\| < 10^{-8}$ by a preconditioned conjugate gradient method.

tively. We used as the convergence test $\|\mathbf{q}^M\| < 10^{-8}$. We started with only one trial vector, $\mathbf{u}_1$ ($\mathbf{v}_1$) which is obtained by the SDT–CI calculation. If there is no good initial guess, the initial approximate vector can be generated by $\mathbf{b} * \mathbf{D}^{-1}$. For the sake of comparison, solutions were also computed by the conjugate gradient method for nonsymmetric cases. That is, the product $\mathbf{A}^T\mathbf{A}$, which is symmetric and positive definite, is constructed and the symmetric linear equation, $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$, is solved by a preconditioned conjugate gradient method. The tables show that the

algorithms proposed here have very good convergence properties. In the method with the orthonormal basis, $\|\mathbf{q}\|$ decreases monotonically and apparently becomes less than a threshold in fewer iterations than that of the method with a bi-orthogonal basis. However, as to the energy, the latter method converges faster than the former method. It takes 6 (12) iterations to achieve an 8-digit accuracy in the latter method and 10 (17) iterations in the former method in Table I (Table II). The method with bi-orthogonal basis solves the right-hand and left-hand linear equations simultaneously. Thus, the subspaces generated are well balanced, which may accelerate the convergence.

In Table III results are shown for the two SAC problems for SiH_2 and C_2H_4. The order of the matrices are 16800 and 45678, respectively (we did not use spatial symmetry). The linear equations are solved by a present method with orthonormal vectors. The initial trial vector $\mathbf{u}_1$ is generated by the SD–CI vector. Large scale calculations show a similar behavior. Convergence of the iterative step was essentially immediate (5 iterations to get an 8-digit accuracy in energy).

In summary, we feel that the present algorithms represent both a practical and efficient approach to solving the large scale symmetric and nonsymmetric linear equation problems.

TABLE III

The Solutions of the SAC Problems for 16800-dimensional SiH_2 and 45678-dimensional C_2H_4[a]

Cycle number	SiH_2 (5s3p1d/2s) basis		C_2H_4 (4s2p1d/3s1p) basis	
	- E (au)	$\|q\|$	- E (au)	$\|q\|$
1	0.12796328	0.02533800	0.24434121	0.01391084
2	0.13425669	0.00595869	0.25548304	0.00295669
3	0.13429248	0.00141847	0.25549774	0.00051597
4	0.13429457	0.00032688	0.25549839	0.00010278
5	0.13429461	0.00007244	0.25549843	0.00001753
6	0.13429461	0.00001548	0.25549843	0.00000403
7	0.13429461	0.00000382	0.25549843	0.00000063
8	0.13429461	0.00000095	0.25549843	0.00000014
9	0.13429461	0.00000021	0.25549843	0.00000002
10	0.13429461	0.00000004	0.25549843	0.000000005
11	0.13429461	0.000000008		

[a] The energies are relative to the SCF energy, $SiH_2 = -289.683340$ au and $C_2H_4 = -74.064833$ au.

ACKNOWLEDGMENTS

The author wishes to thank H. Wasada for computations by the conjugate gradient method. The calculations were carried out on the VP200 computer at the Computational Center of the Nagoya University. 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. H. NAKATSUJI AND K. HIRAO, *J. Chem. Phys.* **68**, 2053, 4279 (1978).
2. J. A. POPLE, R. KRISHNAN, H. B. SCHEGEL, AND J. S. BINKLEY, *Int. J. Quantum Chem. Symp.* **13**, 225 (1979).
3. G. D. PURVIS AND R. J. BARTLETT, *J. Chem. Phys.* **75**, 1284 (1981).
4. P. E. S. WORMER, F. VISSER, AND J. PALDUS, *J. Comput. Phys.* **48**, 23 (1982).
5. M. R. HESTENES AND E. STIEFEL, *N. B. S. J. of Res.* **49**, 409 (1952).
6. E. R. DAVIDSON, *J. Comput. Phys.* **17**, 87 (1975); K. HIRAO AND H. NAKATSUJI, *J. Comput. Phys.* **45**, 246 (1982).
7. K. HIRAO, *J. Chem. Phys.* **79**, 5000 (1983).