On the Simplification of Generalized Conjugate-Gradient Methods for Nonsymmetrizable Linear Systems

Kang C. Jea and David M. Young*
Fu Jen University
Hsinchuang, 242 Taipei, Taiwan
and
Center for Numerical Analysis
The University of Texas at Austin
Austin, Texas 78712

Dedicated to Alexander M. Ostrowski
on the occasion of his ninetieth birthday.

Submitted by Walter Gautschi

ABSTRACT

The conjugate-gradient (CG) method, developed by Hestenes and Stiefel in 1952, can be effectively used to solve the linear system $Au = b$ when A is *symmetrizable* in the sense that ZA and Z are symmetric and positive definite (SPD) for some Z. A number of generalizations of the CG method have been proposed by the authors and by others for handling the nonsymmetrizable case. For many problems the amount of computer memory and computational effort required may be so large as to make the procedures not feasible. Truncated schemes are often used, but in some cases the truncated methods may not converge even though the nontruncated schemes converge. However, it is well known that if A is symmetric, the generalized CG schemes can be greatly simplified, even though A is not SPD, so that the truncated schemes are equivalent to the nontruncated schemes. In the present paper it is shown that such a simplification can occur if a nonsingular matrix H is available such that $HA = A^TH$. (Of course, if $A = A^T$, then H can be taken to be the identity matrix.) It is also shown that such an H always exists; however, it may not be practical to compute H. These results are used to derive three variations of the Lanczos method for solving nonsymmetrizable systems. Two of the forms are well known, but the third appears to be new. An argument is given for choosing the third form over the other two.

*The work was supported in part by the National Science Foundation under Grant MCS-7919829, and by the Department of Energy under Grant DE-AS05-81ER10954 with the University of Texas at Austin.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/$3.00

1. INTRODUCTION

In this paper we are concerned with certain iterative methods for solving the linear system

$$Au = b \tag{1.1}$$

where A is a given real nonsingular $N \times N$ matrix which is large and sparse, and b is a given $N \times 1$ column vector.

The conjugate-gradient method (CG method) developed by Hestenes and Stiefel [22] can be used to solve (1.1) in the *symmetrizable case*, where HA is symmetric and positive definite (SPD) for some SPD matrix H. The convergence of the CG method compares favorably with that of a number of other iterative methods, and moreover the method has the remarkable property that in order to determine a given iteration vector it is only necessary to use information from, at most, two preceding iterations. A number of generalizations of the CG method have been proposed for the nonsymmetrizable case. Young and Jea [42, 43] considered a method called the *idealized generalized conjugate-gradient method* (IGCG method) for this case. The IGCG method has a number of interesting theoretical properties, including convergence to the true solution in at most N iterations under very general conditions. Unfortunately, however, the determination of a given iteration vector requires information on all preceding iterations. Thus an excessive amount of computer storage and computational effort is required. Truncated procedures, where some of the information obtained from previous iterations is discarded, are often used. In this case, however, many of the theoretical properties of the IGCG method are lost. Indeed, in many cases the convergence is slowed; in some cases the process may break down or may fail to converge. In the *symmetric indefinite case*, where A is symmetric but not necessarily SPD, the formulas for the IGCG methods simplify so that, as in the symmetrizable case, a given iteration vector requires information only from one or two previous iterations. One object of this paper is to show that such a simplification can sometimes be accomplished even if A is not symmetric indefinite.

It is shown that the formulas for the IGCG method can be greatly simplified if one has available a matrix H satisfying *condition I* defined by

DEFINITION 1.1. The nonsingular matrix H satisfies *condition I with respect to the matrix A if*

$$HA = A^T H. \tag{1.2}$$

We note that if A is symmetric indefinite, we may let $H = I$ or $H = A$. We show that there always exists a matrix H satisfying condition I. However, in many cases the actual determination of H would not be feasible. For instance, the determination of H in some cases might require more work than solving (1.1).

If an SPD matrix H can be found which satisfies condition I, then the IGCG method, as simplified, is guaranteed to converge. In general, however, the process is subject to the possibility of breakdown. On the other hand, it can be shown that there exists an integer $t \leqslant N$ such that if breakdown does not occur within the first $t+1$ iterations, then $u^{(t+1)} = \bar{u}$, where $\bar{u} = A^{-1}b$ is the true solution of (1.1).

The method of Lanczos can be regarded as another generalization of the CG method for solving nonsymmetrizable systems. The Lanczos method has the attractive property that to determine a given iteration vector it is only necessary to use information from the previous one or two iterations. Unfortunately, however, there is no guarantee that the method will not break down. In Section 6, a derivation of the Lanczos method is given which is based on the application of the IGCG method to a new system derived from (1.1). A matrix H is constructed which satisfies condition I with respect to the matrix of the new system. Three forms of the Lanczos method are obtained corresponding to the three forms of the IGCG method. Convergence and breakdown conditions are given for each form. A basis is given for choosing one of the forms in preference to the other two forms, which are more commonly used.

In Section 7 an attempt is made to place this work in the perspective of the extensive work which has been done on the CG method, and its generalizations, over the past several years.

It is clear that the work described in this paper leaves open a number of questions concerning the behavior of the various methods considered. Nevertheless, it is hoped that the results obtained will prove useful in obtaining a better understanding of the methods.

The authors wish to dedicate this paper, as a token of their highest esteem, to Dr. A. M. Ostrowski on the occasion of his ninetieth birthday.

2. THE CG METHOD FOR THE SYMMETRIZABLE CASE

We now assume that A is symmetrizable; hence HA is SPD for some SPD matrix H. To apply the CG method it is necessary to select an *auxiliary matrix* Z which can be any SPD matrix such that ZA is SPD. The CG *method* can be defined abstractly by the following two conditions:

$$u^{(n)} - u^{(0)} \in K_n(r^{(0)}) = \mathrm{Sp}\{r^{(0)}, Ar^{(0)}, \dots, A^{n-1}r^{(0)}\}, \tag{2.1}$$

$$\|u^{(n)} - \bar{u}\|_{(ZA)^{1/2}} \leqslant \|w - \bar{u}\|_{(ZA)^{1/2}} \tag{2.2}$$

for all w such that $w - u^{(0)} \in K_n(r^{(0)})$. Here the *Krylov space*, $K_n(r^{(0)})$, is the vector space $\mathrm{Sp}\{r^{(0)}, Ar^{(0)}, \dots, A^{n-1}r^{(0)}\}$ spanned by $r^{(0)}, Ar^{(0)}, \dots, A^{n-1}r^{(0)}$, and for any vector v the $(ZA)^{1/2}$-norm is defined by

$$\|v\|_{(ZA)^{1/2}} = \sqrt{(v, ZAv)}\,. \tag{2.3}$$

In Table 1 we give three alternative, but mathematically equivalent, forms of the CG method, which we refer to as ORTHODIR*, ORTHOMIN*, and ORTHORES*. ORTHOMIN* corresponds to the usual two-term form, while ORTHORES* corresponds to the three-term form considered by Engeli et al. [14], by Reid [31], and by Concus, Golub, and O'Leary [9]. For a derivation of these formulas from (2.1)–(2.2) see for instance Young, Hayes, and Jea [41].

The CG method converges substantially faster than the optimum extrapolated Richardson's method, which can be regarded as a kind of "benchmark method" which can be used as a standard of comparison. Thus, we define Richardson's method by

$$u^{(n+1)} = u^{(n)} + b - Au^{(n)} \tag{2.4}$$

Because HA is SPD for some SPD matrix H, it follows that the eigenvalues of A are real and positive. We let $m(A)$ and $M(A)$ denote, respectively, the smallest and largest eigenvalue of A. The optimum extrapolated Richardson's method is defined by

$$u^{(n+1)} = u^{(n)} + \gamma(b - Au^{(n)}) \tag{2.5}$$

where

$$\gamma = \frac{2}{M(A) + m(A)}\,. \tag{2.6}$$

It follows from the analysis of Hageman and Young [20], (see Sections 2.3, 4.2, and 7.3) that if the spectral condition number $\kappa(A)$ $[= M(A)/m(A)]$ is

TABLE 1

FORMULAS FOR THE CG METHOD: SYMMETRIZABLE CASE[a]

ORTHODIR*:

$$q^{(0)} = r^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \hat{\lambda}_n q^{(n)}$$

$$\hat{\lambda}_n = \frac{\left(Zr^{(n)}, q^{(n)}\right)}{\left(ZAq^{(n)}, q^{(n)}\right)}$$

$$q^{(n)} = Aq^{(n-1)} - a_n q^{(n-1)} - b_n q^{(n-2)}, \qquad n = 1,2,\ldots$$

$$a_n = \frac{\left(ZA^2 q^{(n-1)}, q^{(n-1)}\right)}{\left(ZAq^{(n-1)}, q^{(n-1)}\right)}, \qquad n = 1,2,\ldots$$

$$b_n = \frac{\left(ZA^2 q^{(n-1)}, q^{(n-2)}\right)}{\left(ZAq^{(n-2)}, q^{(n-2)}\right)}, \qquad n = 2,3,\ldots \quad (b_1 = 0)$$

$$r^{(n+1)} = r^{(n)} - \hat{\lambda}_n A q^{(n)}$$

ORTHOMIN*:

$$p^{(0)} = r^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}$$

$$\lambda_n = \frac{\left(Zr^{(n)}, p^{(n)}\right)}{\left(ZAp^{(n)}, p^{(n)}\right)} = \frac{\left(Zr^{(n)}, r^{(n)}\right)}{\left(ZAp^{(n)}, p^{(n)}\right)}$$

$$p^{(n)} = r^{(n)} + \alpha_n p^{(n-1)}, \qquad n = 1,2,\ldots$$

$$\alpha_n = -\frac{\left(ZAr^{(n)}, p^{(n-1)}\right)}{\left(ZAp^{(n-1)}, p^{(n-1)}\right)} = \frac{\left(Zr^{(n)}, r^{(n)}\right)}{\left(Zr^{(n-1)}, r^{(n-1)}\right)}$$

$$r^{(n+1)} = r^{(n)} - \lambda_n A p^{(n)}$$

ORTHORES*:

$$u^{(n+1)} = \rho_{n+1}\left(u^{(n)} + \gamma_{n+1} r^{(n)}\right) + \left(1 - \rho_{n+1}\right) u^{(n-1)}$$

$$\gamma_{n+1} = \frac{\left(Zr^{(n)}, r^{(n)}\right)}{\left(ZAr^{(n)}, r^{(n)}\right)}$$

$$\rho_{n+1} = \left[1 - \frac{\gamma_{n+1}}{\gamma_n} \frac{\left(Zr^{(n)}, r^{(n)}\right)}{\left(Zr^{(n-1)}, r^{(n-1)}\right)} \frac{1}{\rho_n}\right]^{-1}$$

$$= \left[1 + \gamma_{n+1} \frac{\left(ZAr^{(n)}, r^{(n-1)}\right)}{\left(Zr^{(n-1)}, r^{(n-1)}\right)}\right]^{-1}, \qquad n = 1,2,\ldots (\rho_1 = 1)$$

$$r^{(n+1)} = \rho_{n+1}\left(r^{(n)} - \gamma_{n+1} A r^{(n)}\right) + \left(1 - \rho_{n+1}\right) r^{(n-1)}$$

[a] In each case $u^{(0)}$ is arbitrary, $r^{(0)} = b - Au^{(0)}$, Z and ZA are SPD.

large, then the CG method requires only about $\kappa(A)^{-1/2}$ as many iterations to achieve a given convergence accuracy as does the optimum extrapolated Richardson's method. Thus, for large $\kappa(A)$, there is an enormous advantage to be gained by using the CG method. Another (less important) property of the CG method is that, in theory, it converges to the true solution in at most N iterations.

We remark that it is often possible to obtain even faster convergence by applying the CG method, not to the original system (1.1), but instead to a system of the form

$$Q^{-1}Au = Q^{-1}b. \tag{2.7}$$

This is sometimes called a "preconditioned system"; see, e.g., Evans [15] and Axelsson [2]. If the matrix $Q^{-1}A$ is symmetrizable and if the condition of the matrix $Q^{-1}A$ is substantially less than that of A, then the CG method applied to (2.7) will converge much faster than the CG method applied to (1.1). We note that if A and Q are SPD, then H and $HQ^{-1}A$ are both SPD if $H = Q$ or $H = A$.

3. THE IGCG METHOD FOR THE NONSYMMETRIZABLE CASE

We now consider a generalized CG method for solving (1.1) in the nonsymmetrizable case. We choose an auxiliary matrix Z such that ZA is positive real[1] (PR). We define the *idealized generalized CG method* (IGCG method) by the conditions

$$u^{(n)} - u^{(0)} \in K_n(r^{(0)}), \tag{3.1}$$

$$(Zr^{(n)}, v) = 0 \qquad \text{for all} \quad v \in K_n(r^{(0)}), \tag{3.2}$$

where $r^{(n)}$ is the residual vector

$$r^{(n)} = b - Au^{(n)}. \tag{3.3}$$

We remark that if ZA is SPD, then the conditions (3.1)–(3.2) are equivalent to the conditions (2.1)–(2.2); see, e.g., [10].

[1] The real matrix K is positive real if $K + K^T$ is SPD. Note that if $Z = A^T Y$ for some PR matrix Y, then ZA is PR.

TABLE 2

FORMULAS FOR THE IGCG METHOD: NONSYMMETRIZABLE CASE[a]

ORTHODIR:

$$q^{(0)} = r^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \hat{\lambda}_n q^{(n)}$$

$$\hat{\lambda}_n = \frac{(Zr^{(n)}, q^{(n)})}{(ZAq^{(n)}, q^{(n)})}$$

$$q^{(n)} = Aq^{(n-1)} + \beta_{n,n-1}q^{(n-1)} + \cdots + \beta_{n,0}q^{(0)}, \qquad n = 1,2,\ldots$$

$$\beta_{n,i} = -\frac{(ZA^2q^{(n-1)}, q^{(i)}) + \sum_{j=0}^{i-1} \beta_{n,j}(ZAq^{(j)}, q^{(i)})}{(ZAq^{(i)}, q^{(i)})}, \qquad i = 0,1,\ldots,n-1, \quad n = 1,2\ldots$$

$$r^{(n+1)} = r^{(n)} - \hat{\lambda}_n Aq^{(n)}$$

ORTHOMIN:

$$p^{(0)} = r^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}$$

$$\lambda_n = \frac{(Zr^{(n)}, p^{(n)})}{(ZAp^{(n)}, p^{(n)})} = \frac{(Zr^{(n)}, r^{(n)})}{(ZAp^{(n)}, p^{(n)})}$$

$$p^{(n)} = r^{(n)} + \alpha_{n,n-1}p^{(n-1)} + \cdots + \alpha_{n,0}p^{(0)}, \qquad n = 1,2,\ldots$$

$$\alpha_{n,i} = -\frac{(ZAr^{(n)}, p^{(i)}) + \sum_{j=0}^{i-1} \alpha_{n,j}(ZAp^{(j)}, p^{(i)})}{(ZAp^{(i)}, p^{(i)})}, \qquad i = 0,1,\ldots,n-1, \quad n = 1,2,\ldots$$

ORTHORES:

$$u^{(n+1)} = \lambda_n r^{(n)} + f_{n+1,n}u^{(n)} + \cdots + f_{n+1,0}u^{(0)}$$

$$\lambda_n = (\sigma_{n+1,0} + \sigma_{n+1,1} + \cdots + \sigma_{n+1,n})^{-1}$$

$$f_{n+1,i} = \lambda_n \sigma_{n+1,i}, \qquad i = 0,1,\ldots,n$$

$$\sigma_{n+1,i} = \frac{(ZAr^{(n)}, r^{(i)}) - \sum_{j=0}^{i-1} \sigma_{n+1,j}(Zr^{(j)}, r^{(i)})}{(Zr^{(i)}, r^{(i)})}, \qquad i = 0,1,\ldots,n$$

$$r^{(n+1)} = -\lambda_n Ar^{(n)} + f_{n+1,n}r^{(n)} + \cdots + f_{n+1,0}r^{(0)}$$

[a] In each case $u^{(0)}$ is arbitrary; $r^{(0)} = b - Au^{(0)}$.

Three forms of the IGCG method, namely, ORTHODIR, ORTHOMIN, and ORTHORES are given in Table 2. Derivations of these formulas are given by Young and Jea [42, 43]. For the derivation of ORTHODIR, only the assumption that ZA is PR is required. For the derivation of ORTHOMIN and ORTHORES it is assumed that Z and ZA are PR. (An example where Z and ZA are PR would be if A is PR and $Z = I$ or A^T.) It should be noted that if Z and ZA are SPD, then the formulas of Table 2 reduce to those of Table 1.

It can be shown (see Young and Jea [42, 43]) that if ZA is PR, then ORTHODIR converges to the true solution $\bar{u}$ in at most N steps. Moreover, if Z and ZA are PR, then ORTHOMIN and ORTHORES converge to $\bar{u}$ in at most N steps. In fact, in this case all three forms of the IGCG method are equivalent.

Let us now consider the situation where neither ZA nor Z is assumed to be PR. It was shown by Jea [23] that for ORTHODIR, ORTHOMIN, and ORTHORES there exists an integer $t \leqslant N$ such that if "breakdown" does not occur within $t+1$ iterations, then $u^{(t+1)} = \bar{u}$, i.e., we have convergence. It is also shown in [23] that ORTHOMIN converges if and only if ORTHORES converges and if both converge, then ORTHODIR converges. In that case all three methods are equivalent in the sense that $u^{(n)}$ is the same in all cases for each n.

For ORTHODIR breakdown occurs if $(ZAq^{(n)}, q^{(n)}) = 0$ but $r^{(n)} \neq 0$. Note that if $r^{(n)} \neq 0$ and if $(ZAq^{(i)}, q^{(i)}) \neq 0$, $i = 0, 1, \ldots, n-1$, then $q^{(n)} \neq 0$. For ORTHOMIN breakdown occurs if either $(ZAp^{(n)}, p^{(n)}) = 0$ or $\lambda_n = 0$ but $r^{(n)} \neq 0$. Note that if $r^{(n)} \neq 0$ and if $(ZAp^{(i)}, p^{(i)}) \neq 0$ and $\lambda_i \neq 0$ for $i = 0, 1, \ldots, n-1$, then $p^{(n)} \neq 0$. Finally, for ORTHORES breakdown occurs if either $(Zr^{(n)}, r^{(n)}) = 0$ or $\sigma_{n+1,0} + \sigma_{n+1,1} + \cdots + \sigma_{n+1,n} = 0$ but $r^{(n)} \neq 0$.

Unfortunately, as can be seen from Table 2, the three forms of the IGCG method, in their idealized, or nontruncated, forms, often require too much computer memory and too many arithmetic operations to be practical. Truncated versions have been developed (see, e.g., [42, 43]) which require much less storage and many fewer operations. However, these procedures may, in some cases, converge much slower than the idealized procedures and may even break down or fail to converge in situations where the idealized procedures would converge.

In Section 5 we will show that in some cases certain truncated versions of the IGCG method are *equivalent* to the idealized versions.

4. CONDITION I

In this section we show the existence of a matrix satisfying condition I with respect to a given matrix A. We prove

THEOREM 4.1. *For any real square matrix A there exists a matrix H satisfying condition I with respect to A.*

We remark that H may be complex.

To prove Theorem 4.1 we first prove the following lemma:

LEMMA 4.2. *Let J_A be a Jordan canonical form of the matrix A. There exists a nonsingular matrix S such that*

$$S^{-1}A^TS = J_A.$$

Proof. Let P be any nonsingular matrix which reduces A to J_A. Thus

$$P^{-1}AP = J_A$$

and

$$P^TA^TP^{-T} = J_A^T.$$

We now construct a permutation matrix R such that

$$R^{-1}J_A^TR = J_A.$$

This can be done in two stages. First we let

$$R_1 = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & \cdot & \cdot & \cdot & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}.$$

Then

$$R_1^{-1}J_A^TR_1$$

is a Jordan canonical form of A except that the ordering of the blocks is just opposite to that in J_A.

Evidently by a permutation of the blocks of $R_1^{-1}J_A^TR_1$ we can obtain J_A. Thus, for some permutation matrix R_2 we have

$$R_2^{-1}(R_1^{-1}J_A^TR_1)R_2 = J_A$$

and

$$\left(R_2^{-1}R_1^{-1}P^T\right)A^T\left(P^{-T}R_1R_2\right) = J_A.$$

Thus the lemma follows with $S = P^{-T}R_1R_2$. ∎

To prove the theorem we let P be any nonsingular matrix such that $P^{-1}AP = J_A$, and we let S be any nonsingular matrix such that $S^{-1}A^TS = J_A$. Thus $A = PJ_AP^{-1}$ and $A^T = SJ_AS^{-1}$. Evidently (1.2) holds with $H = SP^{-1}$, and the theorem follows.

We remark that Theorem 4.1 is a special case of a result given in [19] concerning the solvability of the matrix equation $AX = XB$ where A and B are given.

We also remark that if H satisfies condition I with respect to A then H^T, A^TH, HA, A^TH^T, and $H + H^T$ satisfy (1.2). Therefore H^T satisfies condition I and, if A is nonsingular, so do A^TH, HA, A^TH^T, and H^TA. Moreover, if $H + H^T$ is nonsingular, then it also satisfies condition I.

An important special case is where an SPD matrix H exists which satisfies condition I. This is equivalent to assuming that A is similar to a symmetric matrix. For, if A is similar to a symmetric matrix K, then $W^{-1}AW = K$ for some nonsingular matrix W. In this case $A = WKW^{-1}$ and $A^T = W^{-T}K^TW^T$. The matrix $H = W^{-T}W^{-1}$ is evidently nonsingular and satisfies condition I, since $HA = W^{-T}KW^{-1}$ and $A^TH = W^{-T}K^TW^{-1}$. On the other hand, if an SPD matrix H exists satisfying condition I, then A is similar to the symmetric matrix $H^{1/2}AH^{-1/2}$. This follows because $(H^{1/2}AH^{-1/2})^T = H^{-1/2}A^TH^{1/2} = H^{-1/2}(A^TH)H^{-1/2} = H^{-1/2}(HA)H^{-1/2} = H^{1/2}AH^{-1/2}$.

We remark that if a PR matrix H exists which satisfies condition I, then there also exists an SPD matrix, namely $\frac{1}{2}(H + H^T)$, which also satisfies condition I.

By Theorem 4.1, there always exists a matrix H satisfying condition I. As we have seen, the symmetric matrix $H' = \frac{1}{2}(H + H^T)$ satisfies (1.2). However, H' will not in general satisfy condition I unless H' is nonsingular. To our knowledge, the question of whether or not there always exists a symmetric matrix satisfying condition I is open.

We remark that if a symmetric matrix H which satisfies condition I is available, one could consider replacing the original system (1.1) with the modified system

$$HAu = Hb. \tag{4.1}$$

Here the matrix HA of the system is symmetric but not necessarily SPD. [The symmetry follows because $(HA)^T = A^TH^T = A^TH = HA$ because of (1.2) and the symmetry of H.] In general, the matrix HA is symmetric indefinite. As we shall see, effective methods are available for dealing with such systems. The idea of replacing (1.1) by (4.1) seems reasonable if the condition number of HA is smaller than, or at least not substantially larger than, the condition of A.

5. SIMPLIFICATION OF THE IGCG METHOD

In this section we show that the formulas for the IGCG method given in Table 2 can be greatly simplified if a matrix H is available which satisfies condition I with respect to A. To accomplish this simplification, or reduction, we choose an auxiliary matrix Z to be any matrix, such as H or A^TH for example, such that Z satisfies condition I with respect to A. We obtain the formulas given in Table 3. We refer to the simplified forms of ORTHODIR,

TABLE 3

SIMPLIFIED FORMULAS FOR THE IGCG METHOD[a]

ORTHODIR(2): Same as ORTHODIR* except that

$$a_n = \frac{\left(ZA^2q^{(n-1)}, q^{(n-1)}\right) - b_n\left(ZAq^{(n-2)}, q^{(n-1)}\right)}{\left(ZAq^{(n-1)}, q^{(n-1)}\right)}$$

(note that if ZA is symmetric the formula for a_n is the same as for ORTHODIR*)

ORTHOMIN(1): Same as ORTHOMIN*; use the first form given for λ_n and α_n for ORTHOMIN*.

ORTHORES(1): Same as ORTHORES* except that[b]

$$\gamma_{n+1} = \frac{\left(Zr^{(n)}, r^{(n)}\right)}{\left(ZAr^{(n)}, r^{(n)}\right) - \dfrac{\left(ZAr^{(n)}, r^{(n-1)}\right)}{\left(Zr^{(n-1)}, r^{(n-1)}\right)}\left(Zr^{(n-1)}, r^{(n)}\right)}$$

$$\rho_{n+1} = \left[1 + \gamma_{n+1}\frac{\left(ZAr^{(n)}, r^{(n-1)}\right)}{\left(Zr^{(n-1)}, r^{(n-1)}\right)}\right]^{-1}, \qquad n = 1,2,\ldots \quad (\rho_1 = 1)$$

[a]In each case $u^{(0)}$ is arbitrary and $r^{(0)} = b - Au^{(0)}$; Z is nonsingular and $ZA = A^TZ$.

[b]Note that if Z is symmetric the formula for γ_{n+1} is the same as for ORTHORES* while either formula for ρ_{n+1} for ORTHORES* can be used.

ORTHOMIN, and ORTHORES as ORTHODIR(2), ORTHOMIN(1), and ORTHORES(1), respectively. Details of the simplification process are given in [23].

We remark that if ZA is symmetric, then ORTHODIR(2) and ORTHOMIN(1) reduce to ORTHODIR* and ORTHOMIN*, respectively. Similarly, if Z is symmetric, then ORTHORES(1) reduces to ORTHORES*.

Let us now consider the case where (1.2) holds for some SPD matrix H. This, of course, includes the symmetric indefinite case, where $H = I$. In this case Chandra [6] and Chandra et al. [7] use the auxiliary matrix

$$Z = A^T H. \tag{5.1}$$

Thus ZA is SPD, and ORTHODIR(2) and ORTHOMIN(1) reduce to ORTHODIR* and ORTHOMIN*, respectively. We remark that if we let $Z = H$, then ORTHORES(1) reduces to ORTHORES*.

In any case, with the choice (5.1) ORTHODIR(2) is guaranteed to converge, since ZA is SPD. On the other hand, we have no assurance that ORTHOMIN(1) or ORTHORES(1) will not fail. We remark that one could use ORTHOMIN(1) [which requires somewhat less work and storage than ORTHODIR(2)] and switch to ORTHODIR(2) if trouble occurs; see [6] and [7].

6. THE LANCZOS METHOD

In this section we show that the method of Lanczos [24, 25] for solving (1.1) in the nonsymmetrizable case can be derived from the results in Section 5. To do this we consider an expanded system of the form

$$\textcircled{A}\,\textcircled{u} = \textcircled{b}, \tag{6.1}$$

where

$$\textcircled{A} = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}, \qquad \textcircled{u} = \begin{bmatrix} u \\ \tilde{u} \end{bmatrix}, \qquad \textcircled{b} = \begin{bmatrix} b \\ \tilde{b} \end{bmatrix} \tag{6.2}$$

For the present we will not specify $\tilde{b}$. We note that if u satisfies (1.1) and if $\tilde{u}$ satisfies $A^T\tilde{u} = \tilde{b}$, then $\textcircled{u} = [u \;\; \tilde{u}]^T$ satisfies (6.1). Conversely if $\textcircled{u} = [u \;\; \tilde{u}]^T$ satisfies (6.1), then u satisfies (1.1).

Evidently the matrix $\textcircled{H}$ given by

$$\textcircled{H} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \tag{6.3}$$

satisfies condition I with respect to Ⓐ. This follows because

$$Ⓗ Ⓐ = \begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix} = Ⓐ^T Ⓗ. \tag{6.4}$$

We therefore can apply the ORTHODIR(2), ORTHOMIN(1), and ORTHORES(1) procedures of Section 5 with Ⓩ = Ⓗ,[2] thus obtaining the three methods given in Table 4. (It should be noted that Ⓩ Ⓐ is symmetric; hence the simpler formulas of Table 3 can be used.) For convenience here we refer to the procedures as Lanczos/ORTHODIR, Lanczos/ORTHOMIN, and Lanczos/ORTHORES, respectively. We remark that the "biconjugate gradient" algorithm considered by Fletcher [16, 17] is essentially the Lanczos/ORTHOMIN method.

Since Ⓗ Ⓐ is not necessarily PR, there is no guarantee that any of the three methods will not break down. From the conditions for breakdown given in Section 3 for ORTHODIR, ORTHOMIN and ORTHORES, we obtain the following conditions for the breakdown of the Lanczos methods. For the Lanczos /ORTHODIR method breakdown occurs if $(Aq^{(n)}, \tilde{q}^{(n)}) = 0$ but $r^{(n)} \neq 0$. [We no longer can guarantee that $q^{(n)} \neq 0$ if $(Aq^{(i)}, \tilde{q}^{(i)}) \neq 0$, $i = 0, 1, \ldots, n-1$ and if $r^{(n)} \neq 0$, as we could for ORTHODIR(2).] For the Lanczos/ORTHOMIN method, breakdown occurs if $(Ap^{(n)}, \tilde{p}^{(n)}) = 0$ or $\lambda_n = 0$, but $r^{(n)} \neq 0$. For the Lanczos/ORTHORES method, breakdown occurs if $(Ar^{(n)}, \tilde{r}^{(n)}) = 0$ or $(r^{(n)}, \tilde{r}^{(n)}) = 0$ but $r^{(n)} \neq 0$. There exists an integer $t \leqslant N$ such that if any one of the three methods does not break down within $t+1$ iterations then $u^{(t+1)} = \bar{u}$. Also, the Lanczos/ORTHOMIN method converges if and only if the Lanczos/ORTHORES method converges, and if both converge, then the Lanczos/ORTHODIR method converges and all three methods are equivalent. From this it would appear that the Lanczos/ORTHODIR method is the safest of the three.

We remark that in the symmetrizable case, if Z and ZA are SPD and if we let

$$\tilde{r}^{(0)} = Zr^{(0)}, \tag{6.5}$$

then for $n = 0, 1, \ldots$ we have

$$\begin{aligned} \tilde{r}^{(n)} &= Zr^{(n)}, \\ \tilde{p}^{(n)} &= Zp^{(n)}, \\ \tilde{q}^{(n)} &= Zq^{(n)}. \end{aligned} \tag{6.6}$$

[2]Other choices of Ⓩ, such as Ⓩ = Ⓐ^{T}Ⓗ, are possible. We have not yet studied such choices in detail.

TABLE 4

THE LANCZOS METHOD FOR THE NONSYMMETRIZABLE CASE[a]

Lanczos/ORTHODIR:

$$q^{(0)} = r^{(0)}$$

$$\tilde{q}^{(0)} = \tilde{r}^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \hat{\lambda}_n q^{(n)}$$

$$\hat{\lambda}_n = \frac{\left(\tilde{r}^{(n)}, q^{(n)}\right) + \left(r^{(n)}, \tilde{q}^{(n)}\right)}{2\left(Aq^{(n)}, \tilde{q}^{(n)}\right)}$$

$$q^{(n)} = Aq^{(n-1)} - a_n q^{(n-1)} - b_n q^{(n-2)}, \qquad n = 1,2,\ldots$$

$$\tilde{q}^{(n)} = A^T\tilde{q}^{(n-1)} - a_n \tilde{q}^{(n-1)} - b_n \tilde{q}^{(n-2)}, \qquad n = 1,2,\ldots$$

$$a_n = \frac{\left(Aq^{(n-1)}, A^T\tilde{q}^{(n-1)}\right)}{\left(Aq^{(n-1)}, \tilde{q}^{(n-1)}\right)}, \qquad n = 1,2,\ldots$$

$$b_n = \frac{\left(Aq^{(n-2)}, A^T\tilde{q}^{(n-1)}\right) + \left(Aq^{(n-1)}, A^T\tilde{q}^{(n-2)}\right)}{2\left(Aq^{(n-2)}, \tilde{q}^{(n-2)}\right)},$$

$$= \frac{\left(Aq^{(n-1)}, \tilde{q}^{(n-1)}\right)}{\left(Aq^{(n-2)}, \tilde{q}^{(n-2)}\right)}, \qquad n = 2,3,\ldots, \quad (b_1 = 0)$$

$$r^{(n+1)} = r^{(n)} - \hat{\lambda}_n A q^{(n)}$$

$$\tilde{r}^{(n+1)} = \tilde{r}^{(n)} - \hat{\lambda}_n A^T \tilde{q}^{(n)}$$

Lanczos/ORTHOMIN:

$$p^{(0)} = r^{(0)}$$

$$\tilde{p}^{(0)} = \tilde{r}^{(0)}$$

$$u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}$$

$$\lambda_n = \frac{\left(r^{(n)}, \tilde{r}^{(n)}\right)}{\left(Ap^{(n)}, \tilde{p}^{(n)}\right)}$$

$$p^{(n)} = r^{(n)} + \alpha_n p^{(n-1)}, \qquad n = 1,2,\ldots$$

$$\tilde{p}^{(n)} = \tilde{r}^{(n)} + \alpha_n \tilde{p}^{(n-1)}, \qquad n = 1,2,\ldots$$

$$\alpha_n = \frac{\left(r^{(n)}, \tilde{r}^{(n)}\right)}{\left(r^{(n-1)}, \tilde{r}^{(n-1)}\right)}, \qquad n = 1,2,\ldots$$

$$r^{(n+1)} = r^{(n)} - \lambda_n A p^{(n)}$$

$$\tilde{r}^{(n+1)} = \tilde{r}^{(n)} - \lambda_n A^T \tilde{p}^{(n)}$$

TABLE 4 (*Continued*)

Lanczos/ORTHORES:

$$u^{(n+1)} = \rho_{n+1}\left(u^{(n)} + \gamma_{n+1} r^{(n)}\right) + (1-\rho_{n+1}) u^{(n-1)}$$

$$\gamma_{n+1} = \frac{(r^{(n)}, \tilde{r}^{(n)})}{(Ar^{(n)}, \tilde{r}^{(n)})}$$

$$\rho_{n+1} = \left[1 - \frac{\gamma_{n+1}}{\gamma_n} \frac{(r^{(n)}, \tilde{r}^{(n)})}{(r^{(n-1)}, \tilde{r}^{(n-1)})} \frac{1}{\rho_n}\right]^{-1}, \qquad n = 1, 2, \dots \quad (\rho_1 = 1)$$

$$r^{(n+1)} = \rho_{n+1}\left(r^{(n)} - \gamma_{n+1} A r^{(n)}\right) + (1-\rho_{n+1}) r^{(n-1)}$$

$$\tilde{r}^{(n+1)} = \rho_{n+1}\left(\tilde{r}^{(n)} - \gamma_{n+1} A^T \tilde{r}^{(n)}\right) + (1-\rho_{n+1}) \tilde{r}^{(n-1)}$$

[a] In each case $u^{(0)}$ is arbitrary, $r^{(0)} = b - Au^{(0)}$, and $\tilde{r}^{(0)}$ is arbitrary.

Evidently in this case the formulas of Table 4 reduce to those of Table 1 for the CG method for the symmetrizable case.

If a matrix Z is available which satisfies condition I with respect to A, then we can let $\tilde{r}^{(0)} = Z^T r^{(0)}$. It is easy to show that $\tilde{r}^{(n)} = Z^T r^{(n)}$, $\tilde{p}^{(n)} = Z^T p^{(n)}$, and $\tilde{q}^{(n)} = Z^T q^{(n)}$. Moreover, the formulas of Table 4 reduce to those given in Table 1, except that for the Lanczos/ORTHODIR procedure we have[3]

$$\hat{\lambda}_n = \frac{\left((Z + Z^T) r^{(n)}, q^{(n)}\right)}{2\left(ZAq^{(n)}, q^{(n)}\right)} \tag{6.7}$$

Even if a matrix Z is available which satisfies condition I with respect to A, it does not necessarily follow that one should let $\tilde{r}^{(0)} = Z^T r^{(0)}$. Other choices of $\tilde{r}^{(0)}$ are discussed in [42, 43, 23].

7. RELATION TO OTHER WORK

In this section we attempt to place the work described in this paper in the perspective of work which has been done and is now being done on the CG method and its generalizations. An indication of some of this work has been given earlier.

The CG method was developed by Hestenes and Stiefel [22] in 1952. Generalizations were given by Hestenes [21] in 1956. For various reasons, the method was not widely used for many years. However, beginning in the mid

[3] We conjecture that the Lanczos/ORTHODIR procedure is equivalent to the ORTHODIR procedure given in Table 3.

1960s there was a resurgence of interest with the appearance of papers by Daniel [10, 11], J. K. Reid [31, 32], Axelsson [2], O'Leary [28], Concus, Golub, and O'Leary [9], and many others.

A number of people have worked on extending the CG method to nonsymmetrizable systems. Concus and Golub [8] and Widlund [39] gave a generalized CG method for the case where the matrix of (1.1) is PR. Vinsome [38] considered a procedure which he called ORTHOMIN and which is a truncated version of the method which we call ORTHOMIN here. Similar methods were also considered by Axelsson [3, 4] and by Eisenstat, Elman, and Schultz [12]; see also Elman [13]. Saad [33, 35, 36] considered a procedure based on a method of Arnoldi [1] which is closely related to ORTHORES.

The work of Chandra, Eisenstat, and Schultz [7] and of Chandra [6] on the symmetric indefinite case includes the case where an SPD matrix H is available satisfying condition I with respect to A. Earlier work on the symmetric indefinite case has been done by Luenberger [26, 27]; by Paige and Saunders [29], who introduced a scheme called SYMMLQ; and by Bunch and Kaufman [5], who considered SYMMBK. Other work on the symmetric indefinite case was done by Fridman [18] and by Stoer and Freund [37].

Fletcher [16, 17] considered a biconjugate-gradient algorithm which is essentially the Lanczos/ORTHOMIN scheme given in Section 6. Numerical experiments based on this procedure are given by Wong [40]. For further discussion on the Lanczos method see the papers by Paige and Saunders [29], by Parlett and Scott [30], and by Saad [34].

We now summarize briefly what we believe to be the most significant contributions of this paper. The first result is given in Section 3 and concerns the behavior of the IGCG method when an auxiliary matrix Z is used where Z and ZA are not necessarily PR. Second, we have not seen the result given in Section 4 about the existence of a matrix H satisfying condition I with respect to A, though it may well be known. Third, the simplification of the IGCG method to the case where there is available a matrix H satisfying condition I appears to be new. Finally, we believe that the derivation of various forms of the Lanczos method given in Section 6 is new. While the ORTHOMIN and ORTHORES forms of the Lanczos method are well known, we have not seen the ORTHODIR form previously. This is the form which the theory indicates should be most reliable, and we plan to run numerical experiments to test this result. We note that all of the above results, except that concerning the Lanczos method, are given in the thesis of Jea [23].

REFERENCES

1 W. E. Arnoldi, The principle of minimized iterations in the solution of the matrix eigenvalue problem, *Quart. Appl. Math.* 9:17–29 (1951).

2 O. Axelsson, On reconditioning and convergence acceleration in sparse matrix problems, CERN 74-10, European Organization for Nuclear Research (CERN), Data Handling Division, 1974.

3 O. Axelsson, A generalized conjugate direction method and its application on a singular perturbation problem, in *Proceedings, 8th Biennial Numerical Analysis Conference, Dundee, Scotland, June 26–29, 1979* (G. A. Watson, Ed.), Lecture Notes in Mathematics, No. 773, Springer, Berlin, 1980.

4 O. Axelsson, Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations, *Linear Algebra Appl.* 29:1–16 (1980).

5 J. R. Bunch and L. Kaufman, Some stable methods for calculating inertia and solving symmetric linear systems, Report CU-CS-063-75, Dept. of Mathematics, Univ. of California at San Diego, 1975.

6 R. Chandra, Conjugate gradient methods for partial differential equations, Research Rep. 129, Dept. of Computer Science, Yale Univ., New Haven, Conn., 1978.

7 R. Chandra, S. C. Eisenstat, and M. H. Schultz, The modified conjugate residual method for partial differential equations, in *Advances in Computer Methods for Partial Differential Equations* (R. Vichnevetsky, Ed.), Vol. II, IMACS, 1977.

8 P. Concus and G. H. Golub, A generalized conjugate gradient method for nonsymmetric systems of linear equations, Report STAN-CS-76-535, Computer Science Dept., Stanford Univ., Stanford, Calif., 1976.

9 P. Concus, G. H. Golub and D. P. O'Leary, A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations, in *Sparse Matrix Computation* (J. R. Bunch and D. J. Rose, Eds.), Academic, New York, 1976, pp. 309–332.

10 J. W. Daniel, The conjugate gradient method for linear and nonlinear operator equations, Ph.D. Thesis, Stanford Univ., Stanford, Calif., 1965.

11 J. W. Daniel, The conjugate gradient method for linear and nonlinear operator equations, *SIAM J. Numer. Anal.* 4:10–26 (1967).

12 Stanley C. Eisenstat, Howard C. Elman, and Martin H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations, Research Report 209, Yale Univ., Dept. of Computer Sciences, Aug. 1981.

13 Howard C. Elman, Iterative methods for large, sparse, nonsymmetric systems of linear equations, Research Report 229, Yale Univ., Dept. of Computer Sciences, Apr. 1982.

14 M. Engeli, T. Ginsburg, H. Rutishauser, and E. Stiefel, Refined iterative methods for computation of the solution and the eigenvalues of self-adjoint boundary value problems, *Mitt. Inst. Angew. Math. ETH* (Zurich), No. 8, Basel, Stuttgart, 1959.

15 D. J. Evans, The use of preconditioning in iterative methods for solving linear equations with symmetric positive definite matrices, *J. Inst. Math. Appl.* 4:295–314 (1967).

16 R. Fletcher, Conjugate gradient methods for indefinite systems, in *Proceedings of the Dundee Biennial Conference on Numerical Analysis* (G. A. Watson, Ed.), Springer, New York, 1975, pp. 73–89.

17 R. Fletcher, *Conjugate Gradient Methods for Indefinite Systems*, Lecture Notes in Mathematics 506, Springer, New York, 1976.

18 V. M. Fridman, The method of minimum iterations with minimum errors for a system of linear algebraic equations with a symmetrical matrix, *USSR Computational Math. and Math. Phys.* 2:362–363 (1963).

19 F. R. Gantmacher, *Matrix Theory*, Vol. I, Chelsea, New York, 1959.

20 L. A. Hageman and D. M. Young, *Applied Iterative Methods*, Academic, New York, 1981.

21 M. R. Hestenes, The conjugate-gradient method for solving linear systems, in *Numerical Analysis*, Vol. VI (J. Curtiss, Ed.), McGraw-Hill, New York, 1956.

22 M. R. Hestenes and E. L. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards* 49:409–436 (1952).

23 K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, Ph.D. Thesis, Dept. of Mathematics, Univ. of Texas at Austin, 1982; CNA-176, Center for Numerical Analysis, Univ. of Texas at Austin, 1982.

24 C. Lanczos, An iterative method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards* 45:255–282 (1950).

25 C. Lanczos, Solution of systems of linear equations by minimized iterations, *J. Res. Nat. Bur. Standards* 49:33–53 (1952).

26 D. G. Luenberger, Hyperbolic pairs in the method of conjugate gradients, *SIAM J. Appl. Math.* 17:1263–1267 (1969).

27 D. G. Luenberger, The conjugate residual method for constrained minimization problems, *SIAM J. Numer. Anal.* 7:390–398 (1970).

28 D. P. O'Leary, Hybrid conjugate gradient algorithms, Doctoral thesis, Stanford Univ., Palo Alto, Calif., 1975.

29 C. C. Paige and M. A. Saunders, Solution of sparse indefinite systems of linear equations, *SIAM J. Numer. Anal.* 12:617–629 (1975).

30 B. N. Parlett and D. S. Scott, The Lanczos algorithm with selective orthogonalization, *Math. Comp.* 33:217–238 (1979).

31 J. K. Reid, On the method of conjugate gradients for the solution of large sparse systems of linear equations, in *Proceedings of the Conference on Large Sparse Sets of Linear Equations*, Academic, New York, 1971, pp. 231–254.

32 J. K. Reid, The use of conjugate gradients for systems of linear equations possessing property A, *SIAM J. Numer. Anal.* 9:325–332 (1972).

33 Y. Saad, Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices, *Linear Algebra Appl.* 34:269–295 (1980).

34 Y. Saad, The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems, Technical Report UIUCDCSS-R-1036, Univ. of Illinois at Urbana-Champaign, 1980.

35 Y. Saad, Krylov subspace methods for solving large unsymmetric linear systems, *Math. Comp.* 37:105–126 (1981).

36 Y. Saad, Practical use of some Krylov subspace methods for solving indefinite and unsymmetric linear systems, Technical Report 214, Yale Univ., Dept. of Computer Science, Jan. 1982.

37 J. Stoer and R. Freund, On the solution of large indefinite systems of linear equations by conjugate gradient algorithms, Institut für Angewandte Mathematik und Statistik, Univ. Würzburg, Würzburg, West Germany, 1981.

38 P. K. W. Vinsome, ORTHOMIN, an iterative method for solving sparse sets of simultaneous linear equations, in *4th Symposium of Numerical Simulation of Reservoir Performance of the Society of Petroleum Engineers of the AIME*, Los Angeles, Calif., 1976, Paper SPE 5739.

39 O. Widlund, A Lanczos method for a class of nonsymmetric systems of linear equations, *SIAM J. Numer. Anal.* 15:801–812 (1978).

40 Y. S. Wong, Conjugate gradient type methods for unsymmetric matrix problems, Technical Report No. 79-36, The Institute of Applied Mathematics and Statistics, Univ. of British Columbia, Canada, 1979.

41 David M. Young, Linda J. Hayes, and Kang C. Jea, Generalized conjugate gradient acceleration of iterative methods, Part I: The symmetrizable case, CNA-162, Center for Numerical Analysis, Univ. of Texas at Austin, 1981.

42 D. M. Young and K. C. Jea, Generalized conjugate-gradient acceleration of nonsymmetrizable iterative methods, *Linear Algebra Appl.* 34:159–194 (1980).

43 D. M. Young and K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, Part II: the Nonsymmetrizable case, CNA-163, Center for Numerical Analysis, Univ. of Texas at Austin, 1981.

Received 27 September 1982; revised 20 February 1983.