MATHEMATICS OF COMPUTATION
VOLUME 44, NUMBER 170
APRIL 1985, PAGES 417-424

Conjugate Gradient-Like Algorithms for Solving Nonsymmetric Linear Systems*

By Youcef Saad and Martin H. Schultz

Abstract. This paper presents a unified formulation of a class of the conjugate gradient-like algorithms for solving nonsymmetric linear systems. The common framework is the Petrov-Galerkin method on Krylov subspaces. We discuss some practical points concerning the methods and point out some of the interrelations between them.

1. Introduction. In the recent few years, a large number of generalizations of the conjugate gradient and conjugate residual methods, which are very successful in solving symmetric positive-definite linear systems, have been proposed for solving nonsymmetric linear systems [3], [6], [5], [7], [12]. In this paper we present an abstract framework which includes most of these methods and many new ones. Our goal is to understand the relationships among the methods and to synthesize.

Consider the general linear system:

$$Ax = f, \tag{1}$$

where A is a large sparse nonsymmetric matrix. If x_0 is an initial approximation to x and $r_0 = f - Ax_0$, we define the mth Krylov subspace $K_m \equiv \text{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$. For symmetric positive-definite systems, the conjugate gradient method and conjugate residual methods each compute approximations to x in the affine space $x_0 + K_m$.

Two distinct points of view have been followed to develop methods for nonsymmetric problems. The first is a variational approach generalizing the conjugate residual method, which minimizes the norm of the residual vector over a Krylov subspace. This class of methods includes among others the algorithms ORTHOMIN [14], GCR, [6], ORTHODIR [7], Axelsson's method [3], and GMRES [13].

The second point of view is to regard the conjugate gradient method as a Galerkin process and to derive a generalization in this sense for nonsymmetric problems. This class includes the generalized conjugate gradient method (GCG) [4], [15], the full orthogonalization method [10], the ORTHORES algorithm [7], and the Axelsson-Galerkin method [2].

Users of these methods want to know which of them performs best. Unfortunately, it seems to us at present that none of these different generalizations emerges as a clear winner. A given method will perform better than the others on a

Received November 28, 1983.
1980 *Mathematics Subject Classification*. Primary 65F05.
* This work was supported by ONR Grant number N000014-82-K-0184 and NSF Grant number MCS-8106181.

0025-5718/85 $1.00 + $.25 per page

particular problem, sometimes by a wide margin, but when applied to other problems, it can be disappointingly slow or may diverge. However, many of the methods are mathematically equivalent in the sense that if they start with the same approximation and if they do not break down then they will produce the same sequence of approximations. Among these equivalent versions some can be more reliable or less expensive than others and this knowledge may facilitate the choice.

2. The Petrov-Galerkin Methods for Solving Linear Systems. Many methods for solving large sparse linear systems can be regarded as Petrov-Galerkin methods or oblique projection methods [11]. Let $(\cdot,\cdot)$ be the l_2-inner-product on $\mathbf{R}^N$ and $\|\cdot\|$ be the corresponding l_2-norm. Let K_m and L_m be two subspaces of dimension m of $\mathbf{R}^N$ and x_0 any initial approximate solution to Problem (1). The Petrov-Galerkin method seeks an approximation to (1) of the form $x_0 + z$ where z belongs to the subspace K_m by imposing the condition that the residual vector of the approximation is orthogonal to the subspace L_m. In other words, the Petrov-Galerkin approximate problem can be defined as:

$$\text{find } x = x_0 + z, \text{ where } z \in K_m \text{ such that}$$
$$(f - Ax, v) = 0 \quad \text{for all } v \in L_m.$$

Note that we can avoid referring to affine subspaces by simply observing that Problem (1) is equivalent to solving the system $Az = r_0$ for $z \in K_m$, where $r_0 = f - Ax_0$ is the initial residual. Then the corresponding Petrov-Galerkin problem for z becomes

$$\text{find } z \in K_m \text{ such that } (r_0 - Az, v) = 0 \quad \text{for all } v \in L_m. \tag{2}$$

Suppose now that we have a basis $V_m \equiv [v_1, v_2, \ldots, v_m]$ of K_m and a basis $W_m \equiv [w_1, w_2, \ldots, w_m]$ of L_m. Then in order to realize the Petrov-Galerkin approximation (2), we may write the unknown z in the form $z = V_m y$ and solve the linear system

$$W_m^T(r_0 - AV_m y) = 0.$$

Assuming $W_m^T A V_m$ is nonsingular, this leads to the solution

$$y = \left[W_m^T A V_m\right]^{-1} W_m^T r_0$$
$$\text{and } x = x_0 + V_m y = x_0 + \left[W_m^T A V_m\right]^{-1} W_m^T r_0.$$

The above Petrov-Galerkin approximation is well defined if and only if $W_m^T A V_m$ is nonsingular.

There are two important special cases. The first choice $L_m = K_m$ leads to the well-known Galerkin method. The second choice $L_m = AK_m$ leads to the least-squares method which finds the approximate solution to (1) of the form $x_0 + z$, having the smallest possible residual l_2-norm. In fact this observation may be formalized as follows [8].

THEOREM 1. *If $L_m = AK_m$, $\bar{x}$ is the approximate solution provided by the Petrov-Galerkin method if and only if it minimizes the l_2-norm of the residual vector $f - Ax$, for all x in the affine subspace $x_0 + K_m$.*

Proof. Let x be any vector in $x_0 + K_m$, i.e., $x = x_0 + z$, where $z \in K_m$. Then we have

$$\begin{aligned}\|f - Ax\|^2 &= \|f - A[(x - \bar{x}) + \bar{x}]\|^2 \\ &= (f - A[(x - \bar{x}) + \bar{x}], f - A[(x - \bar{x}) + \bar{x}]) \\ &= (f - A\bar{x}, f - A\bar{x}) - 2(f - A\bar{x}, A(x - \bar{x})) + (A(x - \bar{x}), A(x - \bar{x})).\end{aligned}$$

By the Petrov-Galerkin condition, the middle term in the above expression is equal to zero. Hence,

$$\|f - Ax\|^2 = (f - A\bar{x}, f - A\bar{x}) + (A(x - \bar{x}), A(x - \bar{x})) \geqslant \|f - A\bar{x}\|^2.$$

Conversely, if $\bar{x} = x_0 + \bar{z} \in x_0 + K_m$ minimizes $\|f - Ax\|$ over $x \in x_0 + K_m$, then for any $z \in K_m$, the quadratic polynomial $Q(\alpha) = \|r_0 - A\bar{z} + \alpha Az\|^2$ is minimized at $\alpha = 0$. Setting $dQ(0)/d\alpha = 0$, we get the condition that $(r_0 - A\bar{z}, Az) = 0$, for all z in K_m, which is the Petrov-Galerkin condition when $L_m = AK_m$. □

An interesting question is whether a similar optimality property is also satisfied for the Galerkin method, i.e., when $L_m = K_m$. The answer is known only in the case of a symmetric positive-definite matrix A. For a symmetric positive-definite matrix A we will denote by $\|x\|_A$ the A-norm of x, defined by $\|x\|_A \equiv (Ax, x)^{1/2}$.

THEOREM 2 [9]. *If A is symmetric positive-definite and $L_m = K_m$, $\bar{x}$ is the approximate solution produced by the Petrov-Galerkin method if and only if it minimizes the A^{-1}-norm of the residual vector over $x_0 + K_m$, i.e., we have*

$$\|f - A\bar{x}\|_{A^{-1}} = \min_{x \in x_0 + K_m} \|f - Ax\|_{A^{-1}}.$$

Alternatively, $\bar{x}$ is the Petrov-Galerkin approximate solution if and only if it minimizes the A-norm of the error vector $x - \bar{x}$ over the same affine space.

The lack of an extension of Theorem 2 to the nonsymmetric case seems to be one of the main reasons for the interest in the least-squares methods. The optimality property of Theorem 1 is an important tool in theoretical analysis of the methods.

On the practical side, for the above formulation of the Petrov-Galerkin method, several serious difficulties may occur for a general pair of bases V_m and W_m:

- If m is large, the matrix $W_m^T A V_m$ may be dense and expensive to form;
- The matrix $W_m^T A V_m$ may be ill-conditioned;
- The formation of the approximate solution x requires that all the vectors v_i, $i = 1, 2, \ldots, m$, be saved.

These difficulties usually occur if one attempts to use *directly* the above formulation when K_m is the Krylov subspace $\operatorname{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$. Nevertheless, there are a number of methods which use Krylov subspaces in a more elegant way, leading to some of the most successful techniques for solving nonsymmetric linear systems. A general formulation of these techniques is considered next.

3. Petrov-Galerkin-Krylov Algorithms. Let $b(\cdot, \cdot)$ be a bilinear form on $\mathbf{R}^N$ and θ be a monotonically increasing, integer-valued function defined on the nonnegative integers such that $0 \leqslant \theta(i) \leqslant i + 1$. We can define the following general class of

algorithms:

Algorithm: PGK(b, θ).

1. *Start*: Set $p_0 = r_0 = f - Ax_0$.
2. *Iterate*: For $i = 0, 1, 2, \ldots$ until convergence do:

(a) Compute:

$$x_{i+1} = x_0 + \sum_{j=0}^{i} \alpha_j^{(i)} p_j, \tag{3}$$

$$r_{i+1} = r_0 - \sum_{j=0}^{i} \alpha_j^{(i)} A p_j, \tag{4}$$

where $\{\alpha_j^{(i)}\}$ are chosen so that either

$$\text{(i)} \quad (r_{i+1}, p_j) = 0, \qquad 0 \leqslant j \leqslant i, \tag{5}$$

or

$$\text{(ii)} \quad (r_{i+1}, Ap_j) = 0, \qquad 0 \leqslant j \leqslant i; \tag{6}$$

(b) Compute p_{i+1} by either of the following:

$$\text{(i)} \quad p_{i+1} = r_{i+1} + \sum_{j=\theta(i)}^{i} \beta_j^{(i)} p_j \tag{7}$$

or

$$\text{(ii)} \quad \beta_{i+1}^{(i)} p_{i+1} = Ap_i + \sum_{j=\theta(i)}^{i} \beta_j^{(i)} p_j, \tag{8}$$

where $\{\beta_j^{(i)}\}$ are chosen so that

$b(p_{i+1}, p_j) = 0, \qquad \theta(i) \leqslant j \leqslant i.$

Clearly, (3) and (4) of PGK could be recast as

$$x_{i+1} = x_i + \sum_{j=0}^{i} \alpha_j^{(i)} p_j \quad \text{and} \tag{9}$$

$$r_{i+1} = r_i - \sum_{j=0}^{i} \alpha_j^{(i)} A p_j. \tag{10}$$

This opens the possibility of a truncated Petrov-Galerkin-Krylov method or TPGK method. More precisely, if ϕ satisfies the same hypothesis as θ we can define the following class of methods.

Algorithm: TPGK(b, θ, ϕ).

1. *Start*: Set $p_0 = r_0 - Ax_0$.
2. *Iterate*: For $i = 0, 1, 2, \ldots,$ until convergence do:

$$\text{(a)} \quad x_{i+1} = x_i + \sum_{j=\phi(i)}^{i} \alpha_j^{(i)} p_j, \tag{11}$$

$$r_{i+1} = r_i - \sum_{j=\phi(i)}^{i} \alpha_j^{(i)} A p_j, \tag{12}$$

where $\{\alpha_j^{(i)}\}$ are chosen so that either

(13) (i) $(r_{i+1}, p_j) = 0, \quad \phi(i) \leqslant j \leqslant i,$

or

(14) (ii) $(r_{i+1}, Ap_j) = 0, \quad \phi(i) \leqslant j \leqslant i;$

(b) same as before.

We may also think of doing either PGK or TPGK for a fixed number, m, of iterations, then examining convergence, and then either stopping if the iteration has converged or restarting otherwise. We call these schemes either the restarted PGK (RPGK) or the restarted, truncated PGK (RTPGK). They may be simulated by an appropriate choice of θ and ϕ in TPGK(b, θ, ϕ). Thus the general form of RPGK is as follows.

Algorithm: RPGK(b, θ).

1. *Start*: Set $p_0 = r_0 = f - Ax_0$.
2. *Iterate*: Perform m steps of PGK.
3. *Restart*: if $\|r_m\|$ is sufficiently small STOP else set

$$x_0 = x_m, \quad p_0 = r_0 = r_m \quad \text{and go to STEP 2.}$$

Clearly, TPGK($b, \theta, 0$) = PGK(b, θ). Furthermore, in PGK(b, θ) there is little reason for computing x_{i+1} on each iteration. We compute it only after convergence. If we use the option 2-(b)-(ii) in PGK, then it is unnecessary to compute the residual at each step. Instead, we compute it in step 3 of RPGK to check for convergence after every m iterations. In TPGK(b, θ, ϕ) we need not compute x_{i+1} (or r_{i+1} if option 2-(b)-(ii) is selected) on each iteration as long as we save the array of coefficients $\alpha_j^{(i)}$ and direction vectors so that it can be computed at every lth step, where l is some fixed integer. Moreover, we may use the following form which generates the same Krylov subspace as 2-(b)-(ii):

$$p_{i+1} = A^i p_0 + \sum_{j=\theta(i)}^{i} \beta_j^{(i)} p_j.$$

This enables us to compute the direction vectors $\{A^i p_0\}$ all at once, which might be advantageous for some vector system architectures.

In the next section we show that many published methods are examples of the general formulation presented in this section. By varying the bilinear form b and choosing among the restarted, truncated or restarted-truncated versions, we see that there are infinitely many possibilities.

4. Some Classical Petrov-Galerkin-Krylov Methods.

4.1. *The Generalized Conjugate Residual, GCR, Method.* This method proposed in [5], [6], is of the form PGK(b, θ) with formulation 2-a-(ii), 2-b-(i), $\theta(i) = 0$, $\forall i$, and the bilinear form

(15) $$b(u, v) \equiv (Au, Av).$$

The iterative form (9), (10) is the most common formulation. In that case it is known that $\alpha_j^{(i)} = 0$ for $j \neq i$. Clearly, the algorithm becomes costly as the step number i increases. Therefore, the restarted method is more realistic in practice.

It can be shown by induction that r_{i+1} is orthogonal to all the previous Ar_j's, $j = 1, 2, \ldots, i$. From Theorem 1, it is therefore clear that x_{i+1} minimizes the residual

norm over the affine subspace $x_0 + \text{span}\{r_0, Ar_0, \ldots, A^i r_0\}$. Elman [6] has used this property to derive convergence results for the case of matrices with positive-definite symmetric parts.

4.2. ORTHOMIN. In its original version the ORTHOMIN algorithm presented by Vinsome [14] corresponds to TPGK(b, θ, ϕ), where b is the same as for GCR (i.e., definition (15)), $\phi(i) = i$, $\theta(i) = i - k + 1$, where k is some fixed integer, with options 2-(a)-(ii), 2-(b)-(i). Thus, ORTHOMIN(k) is the truncated version of GCR. Again, it is possible to show that r_{i+1} is orthogonal to Ar_j, $i - k \leqslant j \leqslant i$. An interesting particular case is $k = 0$ which leads to the steepest descent algorithm for minimizing the function $\|f - Ax\|$.

4.3. ORTHODIR. This algorithm presented by Jea and Young [7] corresponds to PGK(b, θ) with b again defined by (15) and options 2-(a)-(ii), 2-(b)-(ii). In the original version $\beta_{i+1}^{(i)} \equiv 1$. This may lead to overflow situations and a scaling is sometimes necessary (although it need not be done at each step).

The restarted ORTHODIR is mathematically equivalent to the restarted GCR [6]. However, the truncated algorithms ORTHOMIN(k) and ORTHODIR(k) are not mathematically equivalent. The truncated ORTHODIR may diverge [6].

4.4. *The* GMRES *Algorithm*. The Generalized Minimum Residual (GMRES) method introduced by Saad and Schultz [13], is a PGK(b, θ) method with the bilinear form b defined as the l_2-inner-product, $\theta(i) = 0$, for all i, and $\beta_{i+1}^{(i)}$ chosen to normalize p_{i+1}. The vectors p_i, $i = 0, 1, \ldots, m$, form an l_2-orthonormal basis of the Krylov subspace. In this case (8) is the well-known Arnoldi process [1]. Once this orthonormal basis is generated, the approximate solution is computed from (3) by requiring that the α's yield the solution with minimum residual. According to Theorem 1, this corresponds to applying option 2-(a)-(ii) in the general formulation PGK(b, θ).

The restarted GMRES method is mathematically equivalent to the restarted versions of both GCR and ORTHODIR [6]. GMRES has several advantages over these two methods, with respect to cost and reliability [13]. There has been no study of the truncated GMRES method.

4.5. *The Axelsson Least-Squares Method*. Axelsson has proposed two generalizations of the conjugate gradient method. One of them is a least-squares method, i.e., the solution minimizes the residual norm over a Krylov subspace. The Axelsson least-squares (Axel-LS) method is a PGK(b, θ) method with $b(u, v)$ defined by (15) and option 2-(b)-(ii) with $\theta(i) = i$. The α's in (3) are computed to minimize the residual norm $\|f - Ax\|$, which corresponds to choosing option 2-(a)-(ii). This method is mathematically equivalent to GCR and to ORTHODIR [6]. The truncated version has been defined by Axelsson, but the restarted version was not given much consideration.

4.6. *The Axelsson-Galerkin Method*. All of the methods described so far have the common feature that they minimize the residual norm in some subspace. In fact GCR, ORTHODIR, GMRES, and Axel-LS methods are equivalent for this reason. A second class of methods is more closely related to the Galerkin methods.

The method proposed by Axelsson in [2] is an example. It corresponds to taking the bilinear form

$$b(u, v) = (u, Av) \tag{16}$$

which makes the direction vectors $\{p_i\}$ semiconjugate. Again $\theta(i)$ is defined as $\theta(i) = i$. The coefficients $\alpha_j^{(i)}$ are defined so that r_{i+1} is orthogonal to the vectors p_j, $j = 0, 1, \ldots, i$, i.e., option 2-(a)-(i) is used. It can then be shown that r_{i+1} is orthogonal to all previous p_i's [2]. The truncated version of this algorithm has been mentioned by Axelsson but not studied. The restarted version has not been considered.

4.7. *The Full Orthogonalization Method.* This method introduced in [10] uses the l_2-inner-product for $b(\cdot,\cdot)$. As in GMRES, $\theta(i) = 0$, for all i, and $\beta_{i+1}^{(i)}$ is chosen to normalize p_{i+1}, which leads to the Arnoldi algorithm for constructing an orthonormal basis $\{p_i\}$, $i = 0, 1, \ldots, m$, of the Krylov subspace $K_m = \text{span}\{r_0, Ar_0, \ldots, A^m r_0\}$. Once this orthonormal basis is generated, the approximate solution is computed from (3) by imposing the condition 2-(a)-(i) on the α's.

This method is equivalent to Axelsson-Galerkin's method, because the residual vectors satisfy the same condition for both methods. The restarted version was defined and tested in [10]. The truncated version of this method as defined by TPGK(b, θ) with the l_2-inner-product for $b(\cdot,\cdot)$, $\theta(i) = \phi(i) = i - k + 1$, and

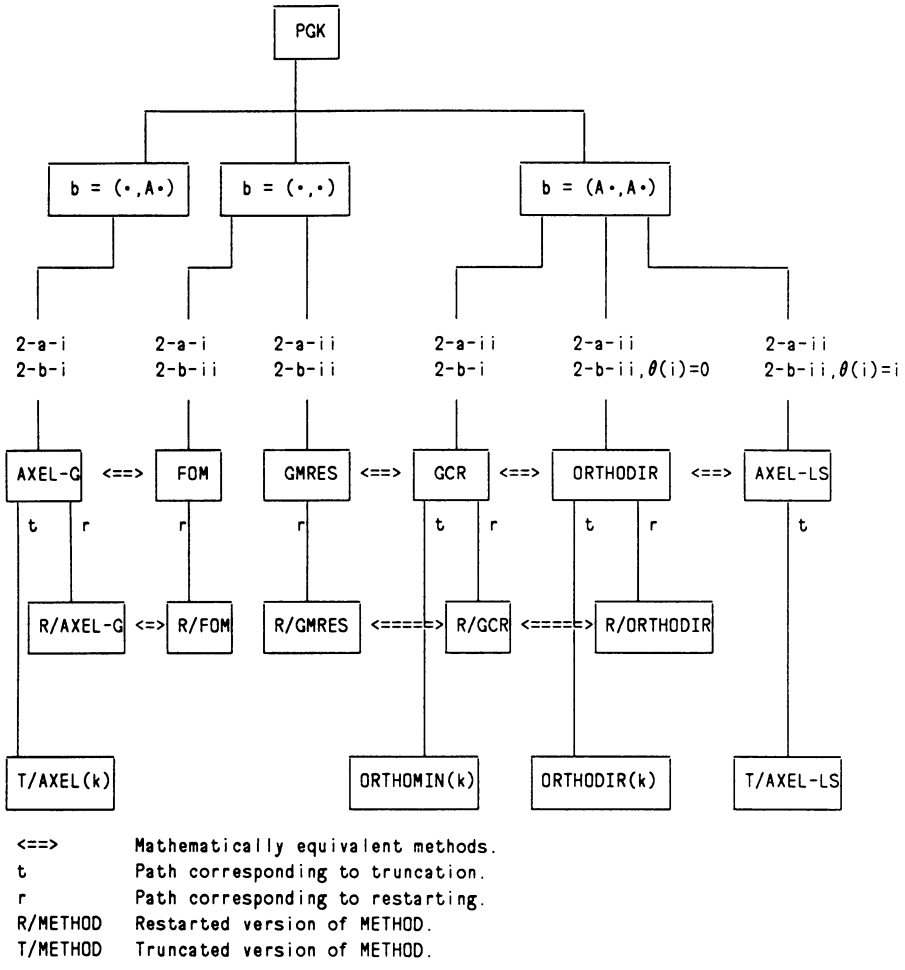

FIGURE 5-1

The conjugate gradient-like methods for nonsymmetric problems

option 2-(a)-(i) has not been considered in the literature. It is *not* the incomplete orthogonalization method (IOM) introduced in [10], [12]. The incomplete orthogonalization method does not fit in the general framework described in Section 3.

5. Synthesis. The chart shown in Figure 5-1 summarizes the derivations of the main methods from the general framework of Algorithm PGK. Note that some of the truncated methods have either been vaguely considered as a possibility by their authors or not been considered at all. Most often the restarted methods have been defined because the basic methods become expensive as the number of steps required for convergence increases. The combinations Restarted-Truncated have been omitted from the chart.

Department of Computer Science
Yale University
New Haven, Connecticut 06520

1. W. E. ARNOLDI, "The principle of minimized iteration in the solution of the matrix eigenvalue problem," *Quart. Appl. Math.*, v. 9, 1951, pp. 17–29.

2. O. AXELSSON, "A generalized conjugate direction method and its application to a singular perturbation problem," in *Proc. 8th biennial Numerical Analysis Conference* (Dundee, Scotland, June 26–29, 1979), (G. A. Watson, ed.), Lecture Notes in Math., vol. 773, Springer-Verlag, Berlin, 1980, pp. 1–11.

3. O. AXELSSON, "Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations," *Linear Algebra Appl.*, v. 29, 1980, pp. 1–16.

4. P. CONCUS & G. H. GOLUB, *A Generalized Conjugate Gradient Method for Nonsymmetric Systems of Linear Equations*, Technical Report STAN-CS-76-535, Stanford University, 1976.

5. S. C. EISENSTAT, H. C. ELMAN & M. H. SCHULTZ, "Variational iterative methods for nonsymmetric systems of linear equations," *SIAM J. Numer. Anal.*, v. 20, 1983, 345–357.

6. H. C. ELMAN, *Iterative Methods for Large Sparse Nonsymmetric Systems of Linear Equations*, Ph.D. Thesis, Computer Science Dept., Yale University, 1982.

7. D. M. YOUNG & K. C. JEA, "Generalized conjugate-gradient acceleration of nonsymmetrizable iterative methods," *Linear Algebra Appl.*, v. 34, 1980, 159–194.

8. M. A. KRASNOSELSKII et al., *Approximate Solutions of Operator Equations*, Wolters-Noordhoff, Groningen, 1972.

9. D. G. LUENBERGER, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1965.

10. Y. SAAD, "Krylov subspace methods for solving large unsymmetric linear systems," *Math. Comp.*, v. 37, 1981, pp. 105–126.

11. Y. SAAD, "The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems," *SIAM J. Numer. Anal.*, v. 19, 1982, pp. 470–484.

12. Y. SAAD, "Practical use of some Krylov subspace methods for solving indefinite and unsymmetric linear systems," *SIAM J. Sci. Statist. Comput.*, v. 5, 1984, pp. 203–228.

13. Y. SAAD & M. H. SCHULTZ, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, Technical Report #254, Yale University, 1983.

14. P. K. W. VINSOME, "ORTHOMIN, an iterative method for solving sparse sets of simultaneous linear equations," in *Proc. Fourth Symposium on Reservoir Simulation*, Society of Petroleum Engineers of AIME, 1976, pp. 149–159.

15. O. WIDLUND, "A Lanczos method for a class of non-symmetric systems of linear equations," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 801–812.