On IGMRES(q): Incomplete Generalized Minimal Residual Methods for Large Unsymmetric Linear Systems*

Zhongxiao Jia†

Abstract

In this paper, we analyze in detail incomplete generalized minimal residual methods (IGMRES(q)), proposed by Brown et al., for solving large unsymmetric linear systems, which are truncated versions of the generalized minimal residual method (GMRES), and establish a convergence theory of the methods. We derive the relationship between the residual norms for IGMRES(q) and GMRES, and expose that IGMRES(q) can work well and behave much like GMRES under the assumption that the basis vectors of Krylov subspace generated by IGMRES(q) are strongly linearly independent. Also, we establish some relationships between the residual norms for IOMs(q) and IGMRES(q). Theoretical analysis shows that IGMRES(q) may not behave like GMRES since the basis vectors generated by IGMRES(q) can be nearly linearly dependent for any unsymmetric matrix and the parameter q. As for practical considerations, we suggest one strategy of choosing q. Finally, some numerical experiments are reported to show convergence behaviors of IGMRES(q) and their restarted versions.

Key words. unsymmetric, Krylov subspace, IGMRES(q), IOMs(q), GMRES, convergence, restarted

AMS(MOS) subject classifications. 65F10

1 Introduction

Consider the solution of linear system

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

*This work was supported by the Graduiertenkolleg Mathematik at the University of Bielefeld

†Fakultät für Mathematik, Universität Bielefeld, email: jia@mathematik.uni-bielefeld.de, Postfach 10 03 01, 33501 Bielefeld, Germany

where $A \in C^{N\times N}$ is nonsingular, large and unsymmetric (non-Hermitian), and b an $N \times 1$ column vector. A major class of methods for solving (1) is Krylov subspace or conjugate gradient type methods. So far, much work has been done; see, e.g. [12, 2, 3, 4, 41, 31, 32, 33, 35, 8, 9, 13, 18, 19, 29, 24, 25, 27] and many others. For survey papers, refer to [16, 36, 17, 14]; for a numerically comparative study of some of them, see [37, 38]; for a synthesis and a general description of available conjugate gradient type methods, see [34].

Of these methods, one kind of the most successful schemes is based on the orthogonal projection process, typified by GCR or ORTHOMIN [39, 41, 9], FOM or Arnoldi's method [31] and GMRES [35]. There are rich references concerning these methods, as mentioned previously. Here we want to stress that there is an intimate connection between FOM and GMRES. Brown [7] has shown that a singular upper Hessenberg matrix H_m in FOM corresponds to the "stagnation" of approximate solutions in GMRES, while the occurrence of the "stagnation" of approximate solutions in GMRES is due to a singular upper Hessenberg matrix H_m. If successful, GMRES performs better than FOM since the residual norms of the former are characterized by a minimization property over the Krylov subspace generated by A. In contrast to those algorithms based on the Lanczos like process [26, 12, 13], GCR, FOM and GMRES use long recurrences, which means that work and storage per step grow as the number of steps increases. As a consequence, one has to restart these algorithms in practice, which may result in very slow convergence for some problems. Except for those matrices of very special forms, however, several researchers have shown that long recurrences cannot be simplified [10, 11, 23].

Even so, noticing the disadvantage of long recurrences, some researchers have proposed truncated strategies, which use only a few rather than all the previously computed vectors in recurrences and can thus be significantly less expensive. For example, Vinsome [39] suggested the truncated version ORTHOMIN(q) of ORTHOMIN or GCR, Young et al. [41] presented the corresponding truncated versions ORTHODIR(q) and ORTHORES(q) of ORTHODIR and ORTHORES, Saad [31] proposed the truncated version IOMs(q) of FOM, and Brown et al. [5, 6] suggested the truncated version IGMRES(q) of GMRES for solving (1). At about the same time, Saad applied IOMs(q) for solving eigenvalue problems [30]. Later, Saad [32, 33, 35], Brown et al. [5, 6] and Gear et al. [15] discussed and used IOMs(q), and a number of researchers described or mentioned them, e.g. [8, 17, 27]. Moreover, Saad [33] conjectured that IOMs(q) should converge in general provided that the parameter q is sufficiently large. Besides, there are some relationships between various truncated versions, for example, Saad [33] has shown that IOMs(q) are equivalent to ORTHORES(q) in the mathematical sense.

A convergence analysis of the truncated ORTHOMIN(q) has been made by [8], and a convergence theory of IOMs(q) has been established by [21]. For all the other truncated versions, however, no convergence theory has been given.

We devote this paper to the study of IGMRES(q), and establish a convergence theory.

In Section 2 we briefly review GMRES and IGMRES(q); in Section 3 we study the convergence theory of IGMRES(q), and derive the relationship between the residual norms for IGMRES(q) and GMRES. The results show that IGMRES(q) can work well and behave much like GMRES under the assumption that the basis vectors of Krylov subspace generated by IGMRES(q) are strongly linearly independent. Theoretical analysis shows that IGMRES(q) may not behave like GMRES since the basis vectors generated by IGMRES(q) can be nearly linearly dependent for any unsymmetric matrix A and the parameter q; in Section 4 we establish some relationships between the residual norms for IOMs(q) and IGMRES(q); as for practical considerations, we suggest in Section 5 one strategy of choosing the parameter q; in Section 6 we report some numerical experiments to show convergence behaviors of IGMRES(q) and their restarted versions.

2 GMRES and IGMRES(q)

Let x^* be the solution of (1), $x^{(0)}$ an initial guess to x^*, and $r^{(0)} = b - Ax^{(0)}$ the initial residual, and let $V_m = (v_1, v_2, \ldots, v_m)$ be a matrix consisting of m linearly independent columns and the Krylov subspace $\mathcal{K}_m(r^{(0)}, A) = span\{r^{(0)}, Ar^{(0)}, \ldots, A^{m-1}r^{(0)}\} = span\{V_m\}$. Throughout the paper, the norm used is the 2-norm unless stated otherwise. The generalized minimal residual method (GMRES) seeks an approximate solution $x_G^{(m)}$ of the form $x_G^{(m)} = x^{(0)} + z_G^{(m)}$ such that $z_G^{(m)}$ satisfies the following orthogonal projection:

$$\begin{cases} z_G^{(m)} & \in & \mathcal{K}_m(r^{(0)}, A), \\ r_G^{(m)} & \perp & A\mathcal{K}_m(r^{(0)}, A), \end{cases} \tag{2}$$

where $r_G^{(m)} = b - Ax_G^{(m)} = r^{(0)} - Az_G^{(m)}$, or equivalently, seeks $z_G^{(m)}$ in $\mathcal{K}_m(r^{(0)}, A)$ such that

$$\|r_G^{(m)}\| = \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|b - A(x^{(0)} + z)\| = \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|r^{(0)} - Az\|. \tag{3}$$

In implementations, GMRES uses Arnoldi's process [1, 30] to generate an orthonormal basis $\{v_i\}_1^m$ of $\mathcal{K}_m(r^{(0)}, A)$ and restarts per m steps. Refer to [35] for details.

Algorithm 1. The restarted GMRES

1. Start: Choose the number of steps m, a tolerance tol and an initial guess $x^{(0)}$. Then compute $r^{(0)} = b - Ax^{(0)}$, and take $v_1 := r^{(0)}/\|r^{(0)}\|$.

2. For $i = 1, 2, \ldots, m$ do Arnoldi's process

$$w := Av_i - \sum_{j=1}^{i} h_{ji} v_j \qquad \text{with } h_{ji} := v_j^H A v_i, \tag{4}$$

$$h_{i+1i} := \|w\|, \qquad v_{i+1} := w/h_{i+1i}. \tag{5}$$

3. Form the approximate solution $x_G^{(m)} = x^{(0)} + V_m y_G^{(m)}$, where $y_G^{(m)}$ solves the least squares problem

$$\min_{y \in C^m} \|\beta e_1 - \bar{H}_m y\|. \tag{6}$$

Here $\bar{H}_m$ is an $(m+1) \times m$ upper Hessenberg matrix whose nonzero entries are h_{ji} generated by Step 2, $\beta = \|r^{(0)}\|$, and e_1 is the first standard unit vector of dimension $m+1$.

4. Decide if $r_G^{(m)} = b - Ax_G^{(m)}$ satisfies the accuracy *tol*. If satisfied, then stop; otherwise, set $x^{(0)} := x_G^{(m)}$, and return to Step 1.

Several notes are made in order. First, it can be shown that solving (6) is equivalent to solving (3); second, after m steps of Arnoldi's process we have an orthonormal matrix $V_{m+1} = (v_1, v_2, \ldots, v_{m+1})$ if $h_{m+1m} \neq 0$; third, if $h_{m+1m} = 0$, then GMRES would find the exact solution x^* of (1) [35].

IGMRES(q) presented by Brown et al. [5] are the algorithms which are but Step 2 the same as Algorithm 1. In Step 2 of IGMRES(q), Av_i is only orthogonalized against the previous q vectors $v_{i-q+1}, \ldots, v_i$.

Algorithm 2. The restarted IGMRES(q)

1. Start: Choose the number of steps m, a parameter q satisfying $2 \leq q \leq m$, a tolerance *tol* and an initial guess $x^{(0)}$. Then compute $r^{(0)} = b - Ax^{(0)}$, and set $v_1 := r^{(0)}/\|r^{(0)}\|$.

2. For $i = 1, 2, \ldots, m$ do the incomplete orthogonalization process

$$\begin{aligned} i_0 &:= \max\{1, i-q+1\}, \\ w &:= Av_i - \sum_{j=i_0}^{i} h_{ji} v_j \text{ with } h_{ji} := v_j^H A v_i, \qquad (7) \\ h_{i+1i} &:= \|w\|, \; v_{i+1} := w/h_{i+1i}. \qquad (8) \end{aligned}$$

3. Form the approximate solution $x_{IG}^{(m)} = x^{(0)} + z_{IG}^{(m)} = x^{(0)} + V_m y_{IG}^{(m)}$, where $y_{IG}^{(m)}$ solves the least squares problem

$$\min_{y \in C^m} \|\beta e_1 - \bar{H}_m y\|. \tag{9}$$

Here $\bar{H}_m$ is an $(m+1) \times m$ banded upper Hessenberg matrix whose nonzero entries are h_{ji} generated by Step 2, and $\beta = \|r^{(0)}\|$.

4. Decide if $r_{IG}^{(m)} = b - Ax_{IG}^{(m)}$ satisfies the accuracy *tol*. If satisfied, then stop; otherwise, set $x^{(0)} := x_{IG}^{(m)}$, and return to Step 1.

Here we want to stress that (9) is not equivalent to (3) since the matrix $V_{m+1} = (v_1, v_2, \ldots, v_{m+1})$ generated by Step 2 is now not orthonormal in general. As in Algorithm 1, $\bar{H}_m$ generated by Algorithm 2 has still full column rank provided that $h_{m+1m} \neq 0$. So (9) has the unique solution. Analogously, as will be seen later, we would find the exact solution x^* of (1) if $h_{m+1m} = 0$.

Two notes are made in order: First, we only need to save $v_{i-q+1}, \ldots, v_i$ in primary memory at the ith step of Step 2, and recall all the $\{v_i\}_1^m$ to compute $x_{IG}^{(m)}$ until the convergence of Algorithm 2 occurs; second, when $q = m$, Algorithm 2 is identical with Algorithm 1.

It is important to observe that Step 2 of Algorithm 2 can be written in matrix form

$$AV_m = V_{m+1}\bar{H}_m, \tag{10}$$

which is very useful later.

3 Convergence analysis

In this section, we establish a convergence theory of IGMRES(q), and prove how IGMRES(q) converge and behave like GMRES.

Proposition 1 *The residual norms for IGMRES(q) and GMRES satisfy*

$$\|r_{IG}^{(m)}\| \leq \kappa(V_{m+1})\|r_G^{(m)}\|, \tag{11}$$

where V_{m+1} is generated by Algorithm 2, $\kappa(V_{m+1}) = \|V_{m+1}\|\|V_{m+1}^+\|$, and V_{m+1}^+ is the pseudoinverse of V_{m+1}.

If $h_{m+1m} = 0$, then Algorithm 2 would find the exact solution x^*.

Proof. Due to $r^{(0)} = \beta V_{m+1}e_1$ and (10), the residual associated with $x_{IG}^{(m)}$ is

$$\begin{aligned} r_{IG}^{(m)} &= b - Ax_{IG}^{(m)} = r^{(0)} - Az_{IG}^{(m)} \\ &= r^{(0)} - AV_m y_{IG}^{(m)} = r^{(0)} - V_{m+1}\bar{H}_m y_{IG}^{(m)} \\ &= V_{m+1}(\beta e_1 - \bar{H}_m y_{IG}^{(m)}), \end{aligned}$$

which gives

$$\|r_{IG}^{(m)}\| = \|V_{m+1}(\beta e_1 - \bar{H}_m y_{IG}^{(m)})\|. \tag{12}$$

Since $V_{m+1}^H V_{m+1}\beta e_1 = V_{m+1}^H r^{(0)}$, we have

$$\beta e_1 = (V_{m+1}^H V_{m+1})^{-1} V_{m+1}^H r^{(0)} = V_{m+1}^+ r^{(0)}.$$

Premultiplying the two hand sides of (10) by V_{m+1}^H yields

$$V_{m+1}^H AV_m = V_{m+1}^H V_{m+1}\bar{H}_m,$$

which results in

$$\bar{H}_m = V_{m+1}^+ A V_m$$

after premultiplying the two hand sides by $(V_{m+1}^H V_{m+1})^{-1}$. Therefore, (12) now becomes

$$\|r_{IG}^{(m)}\| = \|V_{m+1}(V_{m+1}^+ r^{(0)} - V_{m+1}^+ A V_m y_{IG}^{(m)})\|$$

and (9) amounts to solving the least squares problem

$$\min_{y \in C^m} \|V_{m+1}^+ r^{(0)} - V_{m+1}^+ A V_m y\| = \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|V_{m+1}^+ r^{(0)} - V_{m+1}^+ A z\|, \tag{13}$$

where C^m denotes the vector space of dimension m. We thus have

$$\begin{aligned}
\|r_{IG}^{(m)}\| &\leq \|V_{m+1}\| \|V_{m+1}^+ r^{(0)} - V_{m+1}^+ A V_m y_{IG}^{(m)}\| \\
&= \|V_{m+1}\| \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|V_{m+1}^+ r^{(0)} - V_{m+1}^+ A z\| \\
&\leq \|V_{m+1}\| \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|V_{m+1}^+\| \|r^{(0)} - A z\| \\
&= \|V_{m+1}\| \|V_{m+1}^+\| \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|r^{(0)} - A z\| \\
&= \kappa(V_{m+1}) \min_{z \in \mathcal{K}_m(r^{(0)}, A)} \|r^{(0)} - A z\|.
\end{aligned}$$

Hence combining with (3) proves that (11) holds.

If $h_{m+1\,m} = 0$, then (10) becomes $A V_m = V_m \tilde{H}_m$, where $\tilde{H}_m$ is an $m \times m$ banded upper Hessenberg matrix by deleting the last row of $\bar{H}_m$, which means that the columns of V_m span an invariant subspace of A. So the eigenvalues of $\tilde{H}_m$ are all the eigenvalues of A. Therefore, $\tilde{H}_m$ is nonsingular. Now

$$\begin{aligned}
\|r_{IG}^{(m)}\| &= \|r^{(0)} - A V_m y_{IG}^{(m)}\| \\
&= \|r^{(0)} - V_m \tilde{H}_m y_{IG}^{(m)}\| \\
&= \|V_m(\beta e_1 - \tilde{H}_m y_{IG}^{(m)})\|,
\end{aligned}$$

while we have $\min_{y \in C^m} \|\beta e_1 - \tilde{H}_m y\| = 0$ if taking $y_{IG}^{(m)} = \beta \tilde{H}_m^{-1} e_1$. Hence we get $\|r_{IG}^{(m)}\| = 0$, which implies that Algorithm 2 finds the exact solution x^*. □

Remark. If q is taken as m, Algorithm 2 is identical with Algorithm 1 and V_{m+1} is theoretically orthonormal. So $\kappa(V_{m+1}) = 1$ and the equality of (11) holds. We can thus say that (11) cannot be improved.

Now it is necessary for us to explain the meaning of convergence. In the finite N dimensional space, IGMRES(q) must converge at some step $m \leq N$, which is trivial. Therefore, our aim is to attempt to establish some properties, which show that in what conditions $\|r_{IG}^{(m)}\|$ is very near zero for certain $m < N$ if $\|r_G^{(m)}\|$ is such a case, that is, in what conditions IGMRES(q) can behave like GMRES.

$\|r_G^{(m)}\|$ has been analyzed in [35, 20] in detail for the diagonalizable and defective matrix A, respectively, and the results show that $\|r_G^{(m)}\|$ usually approaches zero rapidly provided that the distribution of the eigenvalues of A is favorable, A not highly defective and the eigenvector or Jordan basis of A not ill-conditioned. Therefore, in terms of (11), how IGMRES(q) converge depends strongly on $\kappa(V_{m+1})$. Noticing that the columns of V_{m+1} all have unit norms and the incomplete orthogonality to be exposed later, we have $1 \leq \|V_{m+1}\| \leq \sqrt{m+1-q}$, which is a slowly increasing factor with m. We then conclude how large $\kappa(V_{m+1})$ mainly depends on $\|V_{m+1}^+\|$, namely, the extent of linear independence of the columns of V_{m+1}. The stronger the columns of V_{m+1} have linear independence, the smaller $\|V_{m+1}^+\|$ is. If the columns of V_{m+1} are almost linearly dependent, $\|V_{m+1}^+\|$ will be very (arbitrarily) large, say like 10^{10}. In this circumstance, $\kappa(V_{m+1}) \geq 10^{10}$, which implies, in terms of Proposition 1, that IGMRES(q) may not converge, unless $\|r_G^{(m)}\| = 0$, or at least not behave like GMRES since the right hand side of (11) may not decrease as m increases though $\|r_G^{(m)}\|$ tends to zero. If $\kappa(V_{m+1})$ is moderate, say less than 10^3, then IGMRES(q) surely converge as long as GMRES does.

On behaviors of V_{m+1} and some related quantites, the following two theorems have been proved in [21].

Theorem 1 *Assume $q \geq 2$. Then for all $i - j \geq q + 1$, if Step 2 of Algorithm 2 does not break down, the following relation holds:*

$$| v_j^H v_i | \leq C_i \|A - A^H\| h_{ii-1}^{-1}, \tag{14}$$

where C_i is defined by

$$C_i = 1 + \sum_{k=1}^{i-q-2} f_{k,i}(\bar{H}_m)\Big(\prod_{l=i-k}^{i-1} h_{ll-1}^{-1}\Big), \tag{15}$$

and $f_{k,i}(\bar{H}_m), k = 1, 2, \ldots, i-q-2$ are combinations of the products of k entries of $\bar{H}_m$. If $k \leq 0$, define $f_{k,i}(\bar{H}_m) = 0$.

Theorem 2 *Define $B_{m+1} = V_{m+1}^H V_{m+1}$. Then all singular values $\sigma(B_{m+1})$ of B_{m+1} satisfy*

$$\max\{0, 1 - (m-q)c\|A - A^H\|\} \leq \sigma(B_{m+1}) \leq 1 + (m-q)c\|A - A^H\|, \tag{16}$$

where c is a function depending on $\bar{H}_m$.

Corollary 1 *Under the assumptions of Theorem 1, we have*

$$\|V_{m+1}\| \leq \sqrt{m+1-q}; \tag{17}$$

if $(m-q)c\|A-A^H\| < 1$,

$$\|V_{m+1}^+\| \leq \sqrt{\frac{1}{1-(m-q)c\|A-A^H\|}}, \tag{18}$$

where c *is as in Theorem 1.*

Proof. It is direct from the incomplete orthonormality of V_{m+1} [30, Theorem 3.2] and $\|V_{m+1}\| = \sqrt{\|B_{m+1}\|}$ that

$$\|V_{m+1}\| \leq \sqrt{1+(m-q)} = \sqrt{m+1-q}.$$

From (16), we have

$$\|B_{m+1}^{-1}\| \leq \frac{1}{1-(m-q)c\|A-A^H\|}.$$

On the other hand, noticing that $\|V_{m+1}^+\| = \sqrt{\|B_{m+1}^{-1}\|}$, we obtain

$$\|V_{m+1}^+\| \leq \sqrt{\frac{1}{1-(m-q)c\|A-A^H\|}},$$

which is just (18). □

We make a few comments on Theorem 1 and Corollary 1.

Remarks.

1. It is obvious from Step 2 of Algorithm 2 that the left hand side of (14) is invariant under a scaling of A. In view of (15), we can verify that C_i and the right hand side of (14) indeed has nothing to do with scaling.

2. When $A = A^H$, it follows from Theorem 1 that V_{m+1} is orthonormal provided that $q \geq 2$. In this case, Algorithm 2 coincides with MINRES developed by Paige et al. [28].

3. Theorem 1 and Corollary 1 reveal that V_{m+1} might deviate from orthonormality completely even though we take $q = m-1$ if A is very unsymmetric, which means that $\|A-A^H\|$ or $\|A-A^H\|/\|A\|$ is not a number near zero, e.g. $\|A-A^H\| > 1$ or $\|A-A^H\|/\|A\| > 1$. Therefore, when A is very unsymmetric, it is possible that V_{m+1} will never be close to orthonormality except for $q = m$ in which case V_{m+1} are theoretically orthonormal. Furthermore, for very unsymmetric matrices, the columns of V_{m+1} even become almost linearly dependent such that $\|V_{m+1}^+\|$ will be very large except for $q = m$ in theory. On the other hand, keep in mind that for any $|\, i-j \,| \geq q+1$ we all have $|\, v_i^H v_j \,| < \|v_i\|\|v_j\| = 1$ if Step 2 of Algorithm 2 does not break down. Therefore, (14) and will possibly lose practical value if A is very unsymmetric.

4. From the process of proof, we can see that (14) is essentially of order $\|A - A^H\|$ usually if A is nearly symmetric and all $h_{ii-1}, i = q+2, \ldots, m+1$ are not small. However, in view of (14), if some h_{ii-1} for $q+2 \leq i \leq m$ are small, then the columns of B_{m+1} may become nearly linearly dependent and B_{m+1} thus nearly singular.

An example has been given in [21] to illustrate the above comments.

4 Properties, relationships between residual norms for IOMs(q) and IGMRES(q)

Now let us further analyze IGMRES(q) and give more insights into them. Meanwhile, we compare IGMRES(q) with IOMs(q).

(9) can be solved via the QR decomposition. Borrowing a tool used in [5], we let $F_{m,i} \in C^{(m+1)\times(m+1)}$ be the Givens rotation and

$$F_{m,i} = \begin{pmatrix} I_{i-1} & & & \\ & c_i & -\bar{s}_i & \\ & s_i & \bar{c}_i & \\ & & & I_{m-i} \end{pmatrix},$$

where c_i and s_i are chosen to annihilate the subdiagonal elements of $\bar{H}_m$. Now suppose that the product $F_{m,m}F_{m,m-1}\cdots F_{m,1}$ has been applied to the banded upper Hessenberg matrix $\bar{H}_m$, giving

$$F_{m,m}F_{m,m-1}\cdots F_{m,1}\bar{H}_m = \bar{R}_m \in C^{(m+1)\times m},$$

where $\bar{R}_m$ is upper triangular with its last row containing all zeros. Let $\bar{R}_m^H = (\tilde{R}_m, 0)^H$. Then we have the following result due to Brown et al. [5, 7].

Theorem 3 *Let $P_m = (p_1, p_2, \ldots, p_{m+1}) = F_{m,m}F_{m,m-1}\cdots F_{m,1}$. Then*

$$\begin{aligned} \|r_{IG}^{(m)}\| &= \beta \mid p_{m+1}^H e_1 \mid \|V_{m+1}p_{m+1}\| && (19)\\ &= \beta \mid s_1 s_2 \cdots s_m \mid \|V_{m+1}p_{m+1}\|, && (20)\\ \|r_{IOM}^{(m)}\| &= \|b - Ax_{IOM}^{(m)}\| = \beta \mid s_1 s_2 \cdots s_m \mid \sqrt{1 + (h_{m+1m}/\tilde{r}_{mm})^2}, && (21) \end{aligned}$$

where $\tilde{r}_{mm}$ is the mth diagonal entry of $\tilde{R}_m$.

As stated previously, GMRES performs better than FOM, while their truncated versions do not enjoy this nice property. Combining Theorem 2 with Corollary 1, however, we can establish the following result.

Theorem 4 *Let P_m as above. Then*

$$\frac{\beta \mid s_1 s_2 \cdots s_m \mid}{\|V_{m+1}^+\|} \leq \|r_{IG}^{(m)}\| \leq \beta \mid s_1 s_2 \cdots s_m \mid \sqrt{m+1-q}, \tag{22}$$

$$\|r_{IOM}^{(m)}\| \frac{1}{\|V_{m+1}^+\| \sqrt{1+(h_{m+1m}/\tilde{r}_{mm})^2}} \leq \|r_{IG}^{(m)}\| \leq \|r_{IOM}^{(m)}\| \frac{\sqrt{m+1-q}}{\sqrt{1+(h_{m+1m}/\tilde{r}_{mm})^2}}. \tag{23}$$

Proof. Let us look at $\|V_{m+1}p_{m+1}\|$. It is clear from (17) that $\|V_{m+1}p_{m+1}\| \leq \|V_{m+1}\|\|p_{m+1}\| \leq \sqrt{m+1-q}$. On the other hand, we have $p_{m+1} = V_{m+1}^+ V_{m+1} p_{m+1}$, which results in

$$\|p_{m+1}\| \leq \|V_{m+1}^+\|\|V_{m+1}p_{m+1}\|,$$

namely,

$$\|V_{m+1}p_{m+1}\| \geq \|p_{m+1}\|/\|V_{m+1}^+\| = 1/\|V_{m+1}^+\|.$$

It then follows from (20) that (22) holds.

Dividing (20) by (21) gives

$$\|r_{IG}^{(m)}\| = \|r_{IOM}^{(m)}\| \frac{\|V_{m+1}p_{m+1}\|}{\sqrt{1+(h_{m+1m}/\tilde{r}_{mm})^2}},$$

which, combining (22), shows that (23) holds. □

Remarks.

1. (22) shows that $\|r_{IG}^{(m)}\| \leq \beta\sqrt{m+1-q}$, which implies that they cannot go up arbitrarily. Therefore, compared with Proposition 1, it can be seen that the upper bound of (11) is unobtainable if $\|r_G^{(m)}\|$ converges quite slowly and $\|V_{m+1}^+\|$ is very large as m increases. However, the upper bound of (11) is realistic and almost attainable even when $\|V_{m+1}^+\|$ is very large, if GMRES converges very rapidly. Later numerical experiments will confirm this claim.

2. (23) shows that IGMRES(q) will converge provided that IOMs(q) converge, and the residual norms for IGMRES(q) and IOMs(q) differ by a factor no larger than $\sqrt{m+1-q}$.

We know that $\|r_G^{(m)}\|$ is a nonincreasing function with respect to m in GMRES. Unfortunately, this property cannot extend to IGMRES(q). However, we can present a weaker result.

Theorem 5 *Let P_m as above. Then*

$$\|r_{IG}^{(m)}\| \leq \|r_{IG}^{(m-1)}\| \mid s_m \mid (\mid s_m \mid + \mid c_m \mid \|V_m^+\|). \tag{24}$$

Proof. Let $\tilde{P}_m = F_{m-1,m-1}F_{m-1,m-2}\cdots F_{m-1,1}$ be the product of the Givens rotations applied to $\bar{H}_{m-1}$ such that

$$\tilde{P}_m\bar{H}_{m-1} = \bar{R}_{m-1} \in C^{m\times(m-1)},$$

where $\bar{R}_{m-1}$ is an upper triangular matrix with its last row containing all zeros. Let $\tilde{P}_m = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_m)$. Then we have

$$V_{m+1}p_{m+1} = s_m V_m\tilde{p}_m + c_m v_{m+1} \tag{25}$$

due to Brown et al. [5]. Based on (20) and the proof of Theorem 3, we obtain

$$\begin{aligned}
\|r_{IG}^{(m)}\| &= \|r_{IG}^{(m-1)}\| \mid s_m \mid \frac{\|s_m V_m\tilde{p}_m + c_m v_{m+1}\|}{\|V_m\tilde{p}_m\|} \\
&\leq \|r_{IG}^{(m-1)}\| \mid s_m \mid (\mid s_m \mid + \mid c_m \mid /\|V_m\tilde{p}_m\|) \\
&\leq \|r_{IG}^{(m-1)}\| \mid s_m \mid (\mid s_m \mid + \mid c_m \mid \|V_m^+\|),
\end{aligned}$$

which proves (24). □

Remarks.

1. It can be seen from (24) that the larger $\|V_m^+\|$ is, the more slowly $\|r_{IG}^{(m)}\|$ may decrease.

2. In terms of Brown's Theorem 3.1 [7] and (20)(25), it is easy to see that $\|r_{IG}^{(m)}\| = \|r_{IG}^{(m-1)}\|$ if $\tilde{H}_m$ is singular, since in such a case $s_m = 1$ and $c_m = 0$; conversely, however, it cannot follow from $\|r_{IG}^{(m)}\| = \|r_{IG}^{(m-1)}\|$ that $\tilde{H}_m$ is singular, which is different from the counterpart of the relationships between FOM and GMRES. Nonetheless, we can claim that $\tilde{H}_m$ is singular provided that $\min_{y\in C^m}\|\beta e_1 - \bar{H}_m y\| = \min_{y\in C^{m-1}}\|\beta e_1 - \bar{H}_{m-1}y\|$ holds, which is direct from the derivation of Brown's Theorem 3.1 [7] and (20)(25).

5 Implementations

We are only concerned with choices of q. First of all, we have to point out that q is problem dependent and has something to do with the process of building $\bar{H}_m$. In particular, Theorems 1–2 and Corollary 1 will possibly lose practical values if A is generally unsymmetric. Therefore, it is impossible for us to give a global strategy which covers all real problems and treat all potential cases. Based on the theoretical analysis in Section 3, to some extent we attempt to suggest a reliable strategy. To this end, we will make the following assumptions:

- A is nearly symmetric, namely, $\|A - A^H\|$ is about 10^{-1} or less.

- No occurrence of small h_{ii-1}, $i = q+2, \ldots, m+1$, which means that the right hand side of (14) will be small and of order $\|A - A^H\|$. Note that a very small h_{m+1m} implies the convergence of Algorithm 2 [5].

Under these assumptions, we can roughly simplify the results of Theorem 1 and Corollary 1 as follows:

$$\begin{aligned} \mid v_i^H v_j \mid &\approx \|A - A^H\|, \text{ for } \mid i-j \mid \geq q+1, \\ \|V_{m+1}^+\| &\approx \sqrt{\frac{1}{1-(m-q)\|A-A^H\|}}. \end{aligned}$$

In order to make Algorithm 2 comparable to Algorithm 1, we have to choose q such that $\|V_{m+1}^+\|$ is moderate, namely, $\|V_{m+1}^+\| \leq tol_1$. On the other hand, keep in mind that $q \geq 2$. Remembering these considerations, we now solve the following problem:

$$\begin{cases} \frac{1}{1-(m-q)\|A-A^H\|} & \leq & tol_1^2 \\ q & \geq & 2. \end{cases} \tag{26}$$

The minimal integer q_{min} satisfying (26) is

$$q_{min} = [\max\{m - \frac{tol_1^2 - 1}{tol_1^2 \|A - A^H\|}, 2\}], \tag{27}$$

where the symbol $[\cdot]$ denotes the integer part of a scalar.

We test (27) for $m = 30$, $\|A - A^H\| = 0.1$ and $tol_1 = 10$. It is easy to justify $q_{min} = 19$. If $\|A - A^H\| = 0.01$ and other parameters remain unchanged, then $q_{min} = 2$. Numerical experiments show that (27) is reliable and usually conservative.

There remains an important problem to be solved. (27) involves the computation of $\|A - A^H\|$, which is impractical to obtain. To get rid of this, we simply replace it by using

$$\|A - A^H\| \leq \sqrt{\|A - A^H\|_1 \|A - A^H\|_\infty} = \|A - A^H\|_1, \tag{28}$$

where $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote, respectively, the 1-norm and the ∞-norm. Since A is usually very sparse, the cost of computing $\|A - A^H\|_1$ is negligible, compared with the total cost of solving (1).

6 Numerical experiments

We report some numerical results to show convergence behaviors of IGMRES(q) and their restarted versions, i.e. Algorithm 2. All the experiments were run using *Matlab*4.1 on DEC RISC workstation with the machine precision $eps \approx 2.22 \times 10^{-16}$. For simplicity we do not use any preconditioner.

Example 1. Consider the following two dimensional convection-diffusion partial equations:

$$-\triangle u(x,y) + \alpha \frac{\partial}{\partial x} u(x,y) = f(x,y) \text{ on } \Omega = [0,1] \times [0,1] \tag{29}$$

with the boundary condition $u(x,y) = 0$.

Discretizing (29) by centered differences with mesh size $h = 1/(n+1)$ gives rise to an unsymmetric matrix $A(\alpha)$ of dimension $N = n^2$. The parameter α is used to control the unsymmetry of $A(\alpha)$, ranging from nearly symmetric to very unsymmetric. $A(\alpha)$ is of the block tridiagonal form $A(\alpha) = \text{tri}(I, B(\alpha), I)$, where $B(\alpha) = \text{tri}(-1-\delta, 4, -1+\delta)$ is a tridiagonal matrix with subdiagonal entries $-1-\delta = -1-\alpha/2(n+1)$, main diagonal entries 4 and superdiagonal entries $-1+\delta = -1+\alpha/2(n+1)$. The right hand side b of (1) was set to $A(\alpha)x^*$, where $x^* = (1, 1, \ldots, 1)^H$, and the initial vector $x^{(0)}$ was all taken zero vector. First, we examined Algorithm 2 for various m, q, α and the fixed $n = 22$, respectively. In the following, *it* denotes the number of restarts, $CPU(s)$ and CPU_a, respectively, denote the total CPU consuming with the unit "second" and the average CPU cost per restart, and *ratio* is the ratio of CPU time of Algorithm 2 to that of Algorithm 1. We required that the stopping criterion be $\|r_{IG}^{(m)}\| \leq 10^{-5}$. Also, we remind readers that the secondary storage was not exploited for the simplicity of comparison and Algorithm 2 reduces to Algorithm 1 for $q = m$. Tables 1–3 display the results obtained by Algorithm 2.

For $\alpha = 1$, the matrix $A(\alpha)$ is nearly symmetric. Hence, according to the theoretical analysis, IGMRES(q) behave much like GMRES usually, and Algorithm 2 can thus be expected to be more efficient than Algorithm 1. These expectations were confirmed by Table 1, and it was observed that the columns of all the V_{m+1} kept strong linear independence and strategy (27) worked well in these experiments if tol_1 was all set to 100.

For $\alpha = 25$ and $\alpha = 1000$, $A(\alpha)$ are mild unsymmetric and very unsymmetric. In terms of the a-priori convergence theory presented by [35, 20], it can be expected that Algorithm 1 may use many restarts to achieve the desired accuracy. So if successful, Algorithm 2 may also converge slowly and are possibly less efficient than Algorithm 1 for some q. Compared with Table 1, Tables 2–3 indeed coincided with the theoretical analysis. We found that in most cases Algorithm 2 was competitive with and more efficient than Algorithm 1, this is because all the $\|V_{m+1}^+\|$ observed were moderate. Further, there occurred such cases: Algorithm 2 with smaller q may result in faster convergence than those with larger q, e.g. $m = 30, q = 10$ and $\alpha = 1000$ versus $m = 30, q = 15$ and $\alpha = 1000$. Moreover, Algorithm 2 may also use fewer restarts than Algorithm 1, e.g. $m = 20, q = 5$ and $\alpha = 25$ versus the corresponding Algorithm 1.

Below we take $n = 9$, namely $N = 81$, and then show how the non-restarted IGMRES(q) converge with the steps k, $k = 1, 2, \ldots, m$ and various q, where m was set to 35. Table 4 displays the final computed results with $k = m$.

m	q	it	$CPU(s)$	CPU_a	$ratio$	Residual norm
10	2	15	94.54	6.36	0.7528	$9.8D-6$
10	5	15	93.52	6.23	0.7369	$7.1D-6$
10	8	15	97.46	6.50	0.7679	$5.9D-6$
10	10	15	126.91	8.46	1	$6.2D-6$
20	2	5	57.89	11.60	0.7092	$4.9D-6$
20	5	5	59.45	11.89	0.7272	$5.3D-6$
20	10	5	61.60	12.32	0.7535	$9.D-6$
20	15	5	65.05	13.01	0.7957	$9.7D-6$
20	20	6	81.75	13.62	1	$7.9D-7$
30	2	3	48.21	16.07	0.7086	$1.9D-6$
30	5	3	50.78	16.93	0.7463	$1.7D-6$
30	10	3	55.97	18.66	0.8226	$7.1D-7$
30	15	3	59.74	19.91	0.8780	$6.1D-7$
30	20	3	62.75	20.92	0.9223	$5.9D-7$
30	25	3	66.25	22.08	0.9737	$4.3D-7$
30	30	3	68.04	22.68	1	$3.7D-7$

Table 1: *The results obtained by Algorithm 2 for* $N = n^2 = 484$ *and* $\alpha = 1$, $\|A(1) - A(1)^H\|_1 = 0.8696$

m	q	it	$CPU(s)$	CPU_a	$ratio$	Residual norm
10	5	13	82.75	6.36	1.1133	$3.6D-6$
10	8	11	72.77	6.61	0.9791	$5.4D-6$
10	10	11	74.33	6.76	1	$7.4D-6$
20	5	5	63.36	12.67	0.7877	$6.1D-6$
20	10	7	86.45	12.35	1.0747	$8.D-7$
20	15	6	78.44	13.07	0.9752	$1.5D-6$
20	20	6	80.44	13.41	1	$3.5D-6$
30	5	4	69.28	17.32	0.7989	$5.3D-7$
30	10	5	99.45	19.89	1.1468	$7.7D-8$
30	15	4	88.16	22.04	1.0166	$1.5D-7$
30	20	4	96.92	24.23	1.1176	$4.3D-7$
30	25	3	77.21	25.73	0.8903	$1.5D-7$
30	30	3	86.72	28.91	1	$2.1D-6$

Table 2: *The results obtained by Algorithm 2 for* $N = n^2 = 484$ *and* $\alpha = 25$, $\|A(25) - A(25)^H\|_1 = 2.1739$

m	q	it	$CPU(s)$	CPU_a	$ratio$	Residual norm
10	2	41	268.96	6.56	0.9368	$7.5D-6$
10	5	40	289.20	7.23	1.0073	$5.7D-6$
10	8	35	279.30	7.98	0.9728	$9.3D-6$
10	10	34	287.10	8.44	1	$9.D-6$
20	2	18	255.96	14.22	0.9813	$6.3D-6$
20	5	17	245.31	14.43	0.9405	$3.1D-6$
20	10	16	240.16	15.01	0.9207	$3.D-6$
20	15	15	241.95	16.13	0.9276	$4.9D-6$
20	20	15	260.84	17.39	1	$5.6D-6$
30	2	12	154.48	12.87	0.5575	$2.D-6$
30	5	10	125.39	12.54	0.4525	$4.1D-6$
30	10	9	119.61	13.29	0.4317	$6.5D-6$
30	15	10	185.60	18.56	0.6699	$5.2D-6$
30	20	9	207.42	23.05	0.7486	$6.7D-6$
30	25	9	254.28	28.25	0.9177	$8.8D-6$
30	30	9	277.07	30.79	1	$8.D-6$

Table 3: *The results obtained by Algorithm 2 for* $N = n^2 = 484$ *and* $\alpha = 1000$, $\|A(1000) - A(1000)^H\|_1 = 86.9565$

Several notes are made in order: First, it can be seen from Table 4 that V_{m+1} lost orthonormality completely and became almost column linearly dependent rapidly with small q for nearly symmetric matrices, but V_{m+1} kept strong linear independence for very unsymmetric matrices even for small q; second, $\|r_{IG}^{(k)}\|$ converged smoothly as k increases, and no oscillating phenomena occurred; third, for the very unsymmetric $A(1000)$, both GMRES and IGMRES(q) converged very slowly, as can be expected; fourth, the upper bound of (11) was realistic and almost obtainable.

All the above experiments were run on $A(\alpha)$ with positive definite real parts. Now we report some results on some shifted $A(\alpha) := A(\alpha) - 2I$ with indefinite real parts using the non-restarted IGMRES(q) for various q. As is well known [33], for this class of matrices, compared with the problems earlier, solving (1) may become much more difficult. Table 5 shows the final results computed with $m = 35$.

Obviously, Table 5 shows that solving such (1) became much more difficult using either GMRES or IGMRES(q). In contrast to rather irregular convergence behaviors of IOMs(q), IGMRES(q) exhibited quite smoothly decreasing tendency, this is due to the incomplete minimization property of IGMRES(q) over $\mathcal{K}_m(r^{(0)}, A)$. Also, IGMRES($q$) cannot behave like GMRES for some q even though $A(\alpha)$ is nearly symmetric.

Example 2. Let A be a 300×300 bidiagonal matrix with $1, 2, \ldots, 300$ on the main

m	q	α	$\|A - A^H\|_1$	$\|V_{m+1}^+\|$	$\|r_{IG}^{(m)}\|/\|r_G^{(m)}\|$	Residual norm
35	2	0.01	0.002	$1.4D+9$	$1.9D+5$	$4.9D-11$
35	10	0.01	0.002	$1.7D+8$	$2.5D+4$	$6.4D-12$
35	18	0.01	0.002	$1.7D+5$	$1.D+3$	$2.6D-13$
35	25	0.01	0.002	37.2383	25.0965	$6.5D-15$
35	35	0.01	0.002	1		$2.6D-16$
35	2	1	0.2	$3.D+10$	$3.D+8$	$3.9D-5$
35	10	1	0.2	$6.5D+8$	$4.2D+5$	$5.4D-8$
35	18	1	0.2	$2.9D+6$	$1.1D+5$	$1.4D-8$
35	25	1	0.2	$4.1D+3$	$1.9D+3$	$2.5D-10$
35	35	1	0.2	1		$1.3D-13$
35	2	10	2	$2.4D+12$	$2.7D+9$	$5.4D-4$
35	10	10	2	$4.8D+10$	$5.D+8$	$1.D-4$
35	18	10	2	$7.3D+6$	$1.8D+6$	$3.6D-8$
35	25	10	2	$5.9D+4$	$3.7D+4$	$7.5D-9$
35	35	10	2	1		$2.D-13$
35	2	1000	200	$3.1D+2$	5.2266	1.868810098
35	10	1000	200	3.8500	2.4105	0.861908402
35	18	1000	200	1.1782	1.0343	0.369817160
35	25	1000	200	1.0064	1.0001	0.357579715
35	35	1000	200	1		0.357557027

Table 4: *Final results obtained by IGMRES(q) with* $m = 35$ *for* 81×81 $A(\alpha)$

diagonal and 1 in the first superdiagonal, and let the right hand side $b = (1, 1, \ldots, 1)^H$ and take all initial guess $x^{(0)} = 0$. Again, we required that Algorithm 2 be stopped as soon as $\|r_{IG}^{(m)}\| \leq 10^{-5}$. The results computed are shown in Table 6.

It can be seen from Table 6 that Algorithm 2 solved Example 2 successfully and was competitive with Algorithm 1.

Some other experiments have been run. The results have shown that the restarted IGMRES(q) are often competitive with and even more efficient than the restarted GMRES, in particular when A is nearly symmetric. In addition, IGMRES(q) often exhibit the superiority over IOMs(q). Compared with IOMs(q) for solving large unsymmetric eigenproblems, it seems that both IGMRES(q) and IOMs(q) are much less sensitive to choice of q [22].

In summary, both theoretical analysis and numerical examples have shown that IGMRES(q) behave much like GMRES once the basis vectors $\{v_i\}_1^{m+1}$ are well conditioned and Algorithm 2 is often competitive with and even more efficient than Algorithm 1. Unfortunately, $\{v_i\}_1^{m+1}$ generated by the incomplete orthogonalization process

m	q	α	$\|A - A_1^H\|$	$\|V_{m+1}^+\|$	$\|r_{IG}^{(m)}\|/\|r_G^{(m)}\|$	Residual norm
35	2	0.01	0.002	$6.6D + 10$	$1.6D + 3$	$1.3D - 4$
35	10	0.01	0.002	$2.5D + 10$	$1.4D + 3$	$1.1D - 4$
35	18	0.01	0.002	$3.7D + 6$	70.8862	$5.6D - 6$
35	25	0.01	0.002	$1.8D + 2$	9.7468	$7.7D - 7$
35	35	0.01	0.002	1		$7.9D - 8$
35	2	1	0.2	$4.2D + 11$	$1.9D + 2$	0.098141237
35	10	1	0.2	$8.5D + 9$	$1.D + 2$	0.053797047
35	18	1	0.2	$1.3D + 7$	40.1717	0.021291016
35	25	1	0.2	$9.3D + 6$	4.5115	0.002391121
35	35	1	0.2	1		$5.3D - 4$
35	2	10	2	$9.5D + 13$	$4.3D + 2$	3.052570624
35	10	10	2	$1.1D + 11$	$4.1D + 2$	2.949279109
35	18	10	2	$4.8D + 7$	$1.D + 2$	0.729324925
35	25	10	2	$1.3D + 5$	36.9249	0.261192335
35	35	10	2	1		0.007073604

Table 5: *The results obtained by IGMRES(q) for some shifted* 81×81 $A(\alpha)$ *with indefinite real parts*

m	q	it	$CPU(s)$	CPU_a	$ratio$	Residual norm
10	5	13	82.75	6.36	1.1133	$3.6D - 6$
10	8	11	72.77	6.61	0.9791	$5.4D - 6$
10	10	11	74.33	6.76	1	$5.5D - 6$
20	5	5	63.36	12.67	0.7877	$6.1D - 6$
20	10	7	86.45	12.35	1.0747	$8.D - 7$
20	15	6	78.44	13.07	0.9752	$1.5D - 6$
20	20	6	80.44	13.41	1	$3.5D - 6$
30	5	4	69.28	17.32	0.7989	$5.3D - 7$
30	10	5	99.45	19.89	1.1468	$7.7D - 8$
30	15	4	88.16	22.04	1.0166	$1.5D - 7$
30	20	4	96.92	24.23	1.1176	$4.3D - 7$
30	25	3	77.21	25.73	0.8903	$1.5D - 7$
30	30	3	86.72	28.91	1	$2.1D - 6$

Table 6: *The results obtained by IGMRES(q) applied to Example 2,* $\|A - A^H\|_1 = 2$

cannot be guaranteed to be well conditioned for any unsymmetric A and the parameter q in theory. Therefore, it seems that we should seek other strategies that can replace the incomplete orthogonalization process in Algorithm 2 and generate a well conditioned basis $\{v_i\}_1^{m+1}$ but at the same time keep arithmetic operations and storage as low as possible, compared to GMRES. In such a way, the restarted versions of the modified IGMRES(q) can be expected to be more efficient than the restarted GMRES. I believe that this direction of research is very promising and attractive.

Acknowlegement. I thank Professor L. Elsner for his reading of this manuscript and offering valuable comments and suggestions.

References

[1] A. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 11–29.

[2] O. Axelsson, *On preconditioned conjugate gradient methods*, CNA–143, Center for Numerical Analysis, Univ. of Texas, Austin, 1978.

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

[4] O. Axelsson, *A generalized conjugate gradient direction method and its application on a singular perturbation problem*, Proceedings of 8th Biennial Numerical Analysis Conference, G. A. Watson, ed., Lecture Notes in Mathematics, Springer–Verlag, No. 773, 1980.

[5] P. N. Brown and A. C. Hindmarsh, *Reduced storage methods in stiff ODE systems*, Appl. Math. Comput., 31 (1989), pp. 40–91.

[6] P. N. Brown and Y. Saad, *Hybrid Krylov methods for nonlinear systems of equations*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 450–481.

[7] P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 58–78.

[8] S. C. Eisenstat, H. C. Elman and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[9] H. C. Elman, *Iterative methods for large sparse nonsymmetric systems of linear equations*, Ph.D. Thesis, Computer Science Dept., Yale University, 1982.

[10] V. Faber and T. Manteuffel, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.

[11] V. Faber and T. Manteuffel, *Orthogonal error methods*, SIAM J. Numer. Anal., 24 (1987), pp. 170–187.

[12] R. Fletcher, *Conjugate gradient methods for indefinite systems*, Proceedings, Numerical Analysis Conference, Dundee, G. A. Watson, ed., Lecture Notes in Mathematics, Springer–Verlag, No. 506, 1976, pp. 73–89.

[13] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 291–313.

[14] R. W. Freund, G. H. Golub and N. M. Nachtigal, *Iterative solution of linear systems*, Acta. Numerica, (1992), pp. 1–44.

[15] G. W. Gear and Y. Saad, *Iterative solution of linear equations in ODE codes*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 583–601.

[16] G. H. Golub and D. P. O'Leary, *Some history of the conjugate gradient and Lanczos algorithms: 1946–1976*, SIAM Review, 31 (1989), pp. 50–102.

[17] J. F. Grcar, *Operator coefficient methods for linear equations*, Technical Report Sand89-8691, Sandia National Laboratories, New Mexico 87185 and Livermore, California 94551, 1989.

[18] M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms, Part I*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 594–639.

[19] M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms, Part II*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 15–58.

[20] Z. Jia, *The convergence of Krylov subspace methods for large unsymmetric linear systems*, Preprint, University of Bielefeld, 1994.

[21] Z. Jia, *On IOMs(q): incomplete orthogonalization methods for large unsymmetric systems*, Presented at the 15th Biennial Conference on Numerical Analysis, 29 June–2 July, 1993, Dundee, UK. Preprint, University of Bielefeld, 1994.

[22] Z. Jia, *Some numerical methods for large unsymmetric eigenproblems*, Ph.D. Thesis, University of Bielefeld, 1994.

[23] W. D. Joubert and D. M. Young, *Necessary and sufficient conditions for simplifications of generalized conjugate-gradient algorithms*, Linear Algebra Appl., 88/89 (1987), pp. 449–485.

[24] W. D. Joubert, *On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems*, Submitted to J. Numer. linear Algebra Appl., 1992.

[25] W. D. Joubert and G. F. Carey, *Parallelizable restarted iterative methods for nonsmmetric linear systems, Part I: Theory and Part II: Parallel implementation*, Submitted to Inter. J. Computer Math., 1992.

[26] C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. Res. Natur. Bur. Stand., 49 (1952), pp. 33–53.

[27] R. B. Morgan, *Restarting GMRES for large nonsymmetric systems of linear equations*, Preprint, 1992.

[28] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 717–629.

[29] B. N. Parlett, *Reduction to tridiagnoal form and minimal realizations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 567–593.

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

[31] Y. Saad, *Krylov subspace methods for solving large unsymmetric linear systems*, Math. Comput., 37 (1981), pp. 105–126.

[32] Y. Saad, *The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems*, SIAM J. Numer. Anal., 19 (1982), pp. 485–506.

[33] Y. Saad, *Practical use of some Krylov subspace methods for solving indefinite and nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 856–869.

[34] Y. Saad and M. H. Schultz, *Conjugate gradient-like algorithms for solving nonsymmetric linear systems*, Math. Comput., 44 (1985), pp. 417–424.

[35] Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[36] Y. Saad, *Krylov subspace methods on supercomputers*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 1200–1232.

[37] J. N. Shadid and R. S. Tuminaro, *A comparison of preconditioned nonsymmetric Krylov methods on a large-scale MIND machine*, SIAM J. Sci. Comput., 15 (1994), pp. 440–459.

[38] C. H. Tong, *A comparative study of preconditioned Lanczos methods for nonsymmetric linear systems*, Technical Report Sand91-8240, Sandia National Laboratories, New Mexico 87185 and Livermore, California 94551, 1991.

[39] P. K. W. Vinsome, *ORTHOMIN, an iterative method for solving sparse sets of simultaneous linear systems*, 4th Symposium of Numerical Simulation of Reservior Performance, Society of Petroleum Engineers of the AIME, 1976, pp. 149–160.

[40] H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

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