TAIWANESE JOURNAL OF MATHEMATICS
Vol. 4, No. 3, pp. 385-396, September 2000

AN OVERVIEW OF MGMRES AND LAN/MGMRES METHODS FOR SOLVING NONSYMMETRIC LINEAR SYSTEMS

David R. Kincaid, David M. Young and Jen-Yuan Chen*

Abstract. We present an overview of the MGMRES and LAN/MGMRES iterative methods for solving large sparse linear systems.

1. Introduction

We begin with a brief discussion of background material on Idealized Generalized Conjugate Gradient (IGCG) methods and Krylov subspace methods. Following a review of the Generalized Minimum Residual (GMRES) method, we outline the MGMRES method, which is a modification of the GMRES method. Finally, we sketch a Lanczos-type procedure called the LAN/MGMRES method.

We consider linear systems of the form

$$Au = b,$$

with true solution $\overline{u} = A^{-1}b$. Here A is a large sparse nonsingular matrix of size $N \times N$. Recall that if we are given an arbitrary initial guess $u^{(0)}$ to be used in an iterative method, then the initial *residual vector* is $r^{(0)} = b - Au^{(0)}$. Iterative methods involve iterates $u^{(1)}, u^{(1)}, \ldots, u^{(n)}$ that hopefully converge to an approximation to the true solution; that is, the nth *residual vector* $r^{(n)} = b - Au^{(n)}$ is approximately the zero vector.

[0] Received March 1, 2000.
Communicated by P. Y. Wu.
2000 *Mathematics Subject Classification*: 65F10, 65H10, 65N20, 65N30.
Key words and phrases: Iterative method, generalized minimum residual (GM-RES) method, idealized generalized conjugate gradient method, Krylov subspace method, Lanczos-type procedure.
*The author was partially supported by NSC under grant 89-2115-005, Taiwan.

2. Krylov Subspace and IGCG(Z) Methods

Let Z be an *auxiliary matrix* for an iterative method such as $Z = I$, $Z = Y$, $Z = A^T$, or $Z = A^TY$, for example. If A is symmetric positive definite (SPD), then it can be shown that $Z = I$ for the conjugate gradient method and $Z = A^T$ for the conjugate residual method.

We state several important conditions for Krylov subspace methods and Idealized Generalized Conjugate Gradient IGCG(Z) methods.

Condition I:

$$u^{(n)} - u^{(0)} \in \mathcal{K}_n(r^{(0)}, A) = \text{Span}\left\{r^{(0)}, Ar^{(0)}, \dots, A^{n-1}r^{(0)}\right\}.$$

Here $\mathcal{K}_n(r^{(0)}, A)$ is the *Krylov space* associated with the initial residual vector $r^{(0)}$ and the matrix A.

Condition II (a) (*Minimization condition*): If ZA is SPD, then

$$\langle (u^{(n)} - \overline{u}), (u^{(n)} - \overline{u})\rangle_{ZA} = \|u^{(n)} - \overline{u}\|^2_{ZA^{1/2}} \quad \text{is minimized.}$$

Condition II (b) (*Galerkin condition*):

$$\langle r^{(n)}, v\rangle_Z = 0 \qquad \text{for all } v \in \mathcal{K}_n(r^{(0)}, A).$$

Here the *Z-inner product* is defined as $\langle x, y\rangle_Z = \langle Zx, y\rangle = y^TZx$.

The minimization condition (Condition II (a)) can also be written as:

$$\frac{1}{2}\langle u^{(n)}, u^{(n)}\rangle_{ZA} - \langle b, u^{(n)}\rangle_{ZA} \quad \text{is minimized.}$$

Notice that if $Z = A^TY$, where Y is SPD, then $ZA = A^TYA$ is SPD. It follows that Condition II (a) becomes $\langle r^{(n)}, r^{(n)}_Y = \|r^{(n)}\|^2_{Y^{\frac{1}{2}}}$ is minimized.

The *index* $m = m(r^{(0)}, A)$ of $u^{(0)}$, with respect to A, is the largest integer m such that the vectors $v^{(0)}, v^{(1)}, \dots, v^{(m)}$ are linearly independent. For example, letting $v^{(0)} = r^{(0)}, v^{(1)} = Ar^{(0)}, \dots, v^{(m)} = A^mr^{(0)}$, it can be shown that

$$\begin{aligned} u^{(0)} - \overline{u} \in \mathcal{K}_{m+1}(r^{(0)}, A) &= \text{Span}\{r^{(0)}, Ar^{(0)}, \dots, A^mr^{(0)}\} \\ &= \text{Span}\{v^{(0)}, V^{(1)}, \dots, v^{(m)}\} \end{aligned}$$

if $m \le N - 1$. Then $u^{(m+l)} = \overline{u}$ hopefully.

The IGCG(Z) method is $(n^*, u^{(0)})$-*computable* if $n^* \le m + 1$ and if for all $n \le n^*$ there exists a unique $u^{(n)}$ satisfying $u^{(n)} - u^{(0)} \in \mathcal{K}_n(r^{(0)}, A)$ and

$\langle Zr^{(n)}, v\rangle = 0$ for all $v \in \mathcal{K}_n(r^{(0)}, A)$. Moreover, the IGCG(Z) method is $(n^*, u^{(0)})$-*computable* if and only if the moment matrix $\Delta_{n^*}(ZA, r^{(0)})$ is *strongly regular*. Here the *moment matrix* is given by

$$\Delta_{n^*}(ZA, r^{(0)}) = \begin{bmatrix} \langle v^{(0)}, v^{(0)}\rangle_{ZA} & \cdots & \langle v^{(n^*-1)}, v^{(0)}\rangle_{ZA} \\ \vdots & & \vdots \\ \langle v^{(0)}, v^{(n^*-1)}\rangle_{ZA} & \cdots & \langle v^{(n^*-1)}, v^{(n^*-1)}\rangle_{ZA} \end{bmatrix}.$$

This matrix is strongly regular if all the principal submatrices are nonsingular, which means that for a matrix of order n, the n submatrices of sizes $1 \times 1, 2 \times 2, \cdots, n \times n$ in the top-left-hand corner are nonsingular.

In orthogonal implementations, there are two phases.

Phase I. Construct basis vectors $w^{(0)}, w^{(1)}, \ldots, w^{(n-1)}$ by orthogonalizing Krylov vectors with respect to C :

$$\langle w^{(i)}, w^{(j)}\rangle_C = 0 \quad \text{for } i \neq j.$$

Here C is usually SPD.

Phase II. Choose $c_0^{(n)}, c_1^{(n)}, \ldots, c_{n-1}^{(n)}$ so that the Galerkin condition $\langle Zr^{(n)}, w^{(i)}\rangle = 0$ for $0 \leq i \leq n-1$ is satisfied. We have

$$\begin{aligned} u^{(n)} &= u^{(0)} + c_0^{(n)} w^{(0)} + \cdots + c_{n-1}^{(n)} w^{(n-1)} \\ &= u^{(0)} + W_{n-1} c^{(n)}, \end{aligned}$$

where

$$W_{n-1} = \left[W^{(0)}\ w^{(1)}\ \cdots w^{(n-1)}\right], \quad c^{(n)} = \left[c_0^{(n)}\ c_1^{(n)}\ \cdots c_{n-1}^{(n)}\right]^T.$$

In Phase I, we have

$$w^{(n)} = Aw^{(n-1)} + \beta_{n,0} w^{(0)} + \cdots + \beta_{n,n-1} w^{(n-1)}.$$

Examples are as follows: $C = A^T Z$ corresponds to the ORTHODIR(Z) method, $C = A$ corresponds to the ORTHORES(Z) method, and $C = Y$ together with $Z = A^T Y$ corresponds to the GMRES($A^T Y$) method when Y is SPD. The latter method is really the GGMRES method. For the GMRES method, we have $C = I$.

In Phase II, we have

$$w^{(n)} = r^{(n)} + \alpha_{n,0} w^{(0)} + \cdots + \alpha_{n,n-1} w^{(n-1)}.$$

Examples are as follows: $C = ZA$ corresponds to the ORTHOMIN(Z) method while $Z = A^T$ implies the conjugate residual method and $Z = I$ implies the conjugate gradient method.

3. Gmres Method

We now sketch the GMRES method of Saad and Schultz [6]. Let $Z = A^T Y$, where Y is a SPD matrix. Note that $ZA = A^T YA$ is a SPD matrix. As mentioned above, Condition II (a) becomes $\langle Yr^{(n)}, r^{(n)}\rangle = \|r^{(n)}\|_{Y^{1/2}}$ is minimized.

In Phase I, we have

$$\begin{cases} \widehat{w}^{(0)} = r^{(0)} \\ w^{(0)} = {\sigma_0}^{-1}\widehat{w}^{(0)}, \quad \text{where } \sigma_0 = \langle Y\widehat{w}^{(0)}, \widehat{w}^{(0)}\rangle^{\frac{1}{2}}, \end{cases}$$
$$\vdots$$
$$\begin{cases} \widehat{w}^{(n)} = Aw^{(n-1)} + \beta_{n,0}w^{(0)} + \cdots + \beta_{n,n-1}w^{(n-1)} \\ w^{(n)} = {\sigma_n}^{-1}\widehat{w}^{(n)}, \quad \text{where } \sigma_n = \langle Y\widehat{w}^{(n)}, \widehat{w}^{(n)}\rangle^{\frac{1}{2}}. \end{cases}$$

Here

$$\langle Yw^{(i)}, w^{(j)}\rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

We have the basic relation

$$A[w^{(0)}\ w^{(1)}\ \cdots w^{(n-1)}] = [w^{(0)}\ w^{(1)}\ \cdots w^{(n)}]H_n,$$

or

$$AW_{n-1} = W_n H_n.$$

Here H_n is an upper Hessenberg matrix of order n.

Example $(n = 2)$:

$$A[w^{(0)}\ w^{(1)}] = [w^{(0)}\ w^{(1)}\ w^{(2)}] \begin{bmatrix} -\beta_{1,0} & -\beta_{2,0} \\ \sigma_1 & -\beta_{2,1} \\ 0 & \sigma_2 \end{bmatrix}.$$

Hence, we have

$$AW_1 = W_2 H_2.$$

■

In Phase II of the GMRES method, we have

$$\begin{aligned} u^{(n)} &= u^{(0)} + c_0^{(n)}w^{(0)} + \cdots + c_{n-1}^{(n)}w^{(n-1)} \\ &= u^{(0)} + W_{n-1}c^{(n)}. \end{aligned}$$

Consequently, from this equation we obtain

$$\begin{aligned} r^{(n)} &= b - Au^{(n)} \\ &= r^{(0)} - AW_{n-1}c^{(n)} \\ &= r^{(0)} - W_n H_n c^{(n)} \\ &= W_n\left(e^{(n+1)} - H_n c^{(n)}\right), \end{aligned}$$

using $AW_{n-1} = W_n H_n$ and $r^{(O)} = W_n e^{(n+1)}$, where $e^{(n+1)} = [\sigma_n, 0, \ldots, 0]^T_{n+1}$. Thus, we find

$$\begin{aligned} \langle Yr^{(n)}, r^{(n)}\rangle &= \langle YW_n(e^{(n+1)} - H_n c^{(n)}, W_n(e^{(n+1)} - H_n c^{(n)})\rangle \\ &= \|e^{(n+1)} - H_n c^{(n)}\|_2^2, \end{aligned}$$

since $W_n^T Y W_n = I_n$ and Y is SPD.

Example ($n = 2$): Determination of $c^{(2)}$. The system

$$H_2 c^{(2)} = e^{(3)}$$

has the form

$$\begin{bmatrix} -\beta_{1,0} & -\beta_{2,0} \\ \sigma_1 & -\beta_{2,1} \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} c_0^{(2)} \\ c_1^{(2)} \end{bmatrix} = \begin{bmatrix} \sigma_2 \\ 0 \\ 0 \end{bmatrix}.$$

Using Givens rotations $Q = Q_1 Q_2$ with $QQ^T = I$, we have

$$QH_2 c^{(2)} = Qe^{(3)},$$

which is of the form

$$\begin{bmatrix} \boxed{\times} & \times \\ 0 & \boxed{\times} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_0^{(2)} \\ c_1^{(2)} \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ \boxed{\times} \end{bmatrix}.$$

To get the least squares solution, we solve the first two equations for $c_0^{(2)}$ and $c_1^{(2)}$. ∎

Note that the sum of the squares of the residuals are preserved:

$$\begin{aligned} \langle Q(b - Au), Q(b - Au)\rangle &= \langle (b - Au), Q^T Q(b - Au)\rangle \\ &= \langle b - Au, b - Au\rangle. \end{aligned}$$

Some comparisons for orthogonal implementations. If the matrix ZA is SPD (that is, if $Z = A^T Y$ for some SPD matrix Y), then the ORTHODIR method converges but the ORTHOMIN and ORTHORES methods may breakdown. The ORTHODIR method is often numerically unstable and requires more work per iteration than the GMRES method. The GMRES(A^TY) method, where Y is a SPD matrix, is mathematically equivalent to the ORTHODIR(A^TY) method, but requires less work per iteration and is more stable. The GMRES(A^TY) method is widely used but the work per iteration increases as n increases.

4. Mgmres Method

We now sketch the MGMRES method, which is a modification of the GMRES method. We assume Y is symmetric and nonsingular (not necessarily SPD). Also, we suppose that YA is symmetric,

In Phase I of the MGMRES method, we have

$$\widehat{w}^{(n)} = AW^{(n-1)} + \beta_{n,n-1}w^{(n-1)} + \beta_{n,n-2}w^{(n-2)},$$

and $\langle w^{(n)}, Yw^{(i)}\rangle = 0$ for $0 \le i \le n-1$. Then we obtain

$$w^{(n)} = {\sigma_n}^{-1}\widehat{w}^{(n)},$$

where $\sigma_n = |\langle Y\widehat{w}^{(n)}, \widehat{w}^{(n)}\rangle|^{1/2}$. Here the absolute value signs are used since the expression within them may be negative. Moreover, the process fails if $\sigma_n = 0$. Next, we have

$$W_n^T Y W_n = \mathrm{diag}(\pm 1, \pm 1, \dots, \pm 1) \equiv D_n.$$

Here D_n is a diagonal matrix with ± 1 as diagonal entries. For the GMRES method, if Y is a SPD matrix, then $D_n = \mathrm{diag}(1, 1, \dots, 1)$.

In Phase II of the MGMRES method, we use the Galerkin condition

$$W_{n-1}^T(Zr^{(n)}) = 0.$$

Also, we have $Z = A^T Y$, $r^{(n)} = W_n(e^{(n+1)} - H_n c^{(n)})$, and $\langle Zr^{(n)}, w^{(i)}\rangle = 0$ for $0 \le i \le n-1$. So we obtain

$$H_n^T W_n^T Y W_n H_n c^{(n)} = H_n^T W_n^T Y W_n e^{(n+1)},$$

which implies that

$$H_n^T D_n H_n c^{(n)} = H_n^T D_n e^{(n+1)}.$$

If $D_n = I$, we get the *normal equations*

$$H_n^T H_n c^{(n)} = H_n^T e^{(n+1)}.$$

Applying a sequence of Givens rotations, we form an upper triangular system

$$QH_n = \tilde{H}_n,$$

where $Q^T Q = I$.

Example ($n = 2$):

$$H_2 = \begin{bmatrix} -\beta_{1,0} & -\beta_{2,0} \\ \sigma_1 & -\beta_{2,1} \\ 0 & \sigma_2 \end{bmatrix} \quad \Longrightarrow \quad QH_2 = \tilde{H}_2 = \begin{bmatrix} \boxed{\times} & \times \\ 0 & \boxed{\times} \\ 0 & 0 \end{bmatrix},$$

where $Q = Q_1 Q_2$. ■

Using $H_n = Q^{-1}\tilde{H}_n = Q^T \tilde{H}_n$ and $Q^{-1} = Q^T$, we obtain

$$\tilde{H}_n^T Q D_n Q^T \tilde{H}_n c^{(n)} = \tilde{H}_n^T Q D_n e^{(n+1)}.$$

Letting

$$\begin{aligned} z &= Q D_n Q^T \tilde{H}_n c^{(n)}, \\ y &= \tilde{H}_n^T Q D_n e^{(n+1)}, \end{aligned}$$

we obtain

$$\tilde{H}_n^T z = y.$$

So our strategy is to first solve this system to get z and then solve

$$\tilde{H}_n c^{(n)} = Q D_n^{-1} Q^T z.$$

Example $(n = 2)$: The system

$$\tilde{H}_2^T\, z = y$$

has the form

$$\begin{bmatrix} \boxed{\times} & 0 & 0 \\ x & \boxed{\times} & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

And we obtain

$$z = \begin{bmatrix} z_1 \\ z_2 \\ 0 \end{bmatrix} + k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

where k is arbitrary.

$$\tilde{H}_2 c^{(2)} = Q D_2^{-1} Q^T \left\{ \begin{bmatrix} z_1 \\ z_2 \\ 0 \end{bmatrix} + k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \begin{bmatrix} z_1' \\ z_2' \\ 0 \end{bmatrix},$$

for suitable k. (If $D_2 = I_2$, let $k = 0$.) Failure occurs if the third component of $Q D_2^{-1} Q^T [z_1\ z_2\ 0]^T$ is not zero and the third component of $Q D_2^{-1} Q^T [0\ 0\ 1]^T$ is zero. For GMRES, $Q D_2 Q^T = I$ and we let $k = 0$.

$$\tilde{H}_2 c^{(2)} = \begin{bmatrix} \boxed{\times} & \times \\ 0 & \boxed{\times} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_0^{(2)} \\ c_1^{(2)} \end{bmatrix} = \begin{bmatrix} z_1' \\ z_2' \\ 0 \end{bmatrix}.$$

Finally, we solve for $c_0^{(2)}$ and $c_1^{(2)}$. Note this process might fail (if $z_3' \neq 0$).

$$QD_2^{-1}Q^T \begin{bmatrix} z_1 \\ z_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ \boxed{\times} \end{bmatrix},$$

and

$$QD_2^{-1}Q^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ 0 \end{bmatrix}.$$

(Here $\boxed{\times} \neq 0$, which will not happen for the GMRES method since $D_2 = I$.)
■

In the computation of the MGMRES methods, we assume A is nonsingular, Y is symmetric and nonsingular, $Z = A^TY$, and $n^* \leq m$, which is the *index* of the $r^{(0)}$ vector. In Phase I, the MGMRES method is $(n^*, u^{(0)})$-computable if and only if $\Delta_n(Y, r^{(0)})$ is strongly regular. (This condition is not required for the ORTHODIR(A^TY) method.) In Phase II, if $\Delta_n(Y, r^{(0)})$ is strongly regular then the MGMRES method is $(n^*, u^{(0)})$-computable if and only if $\Delta_n\left(A^TYA, r^{(0)}\right)$ is strongly regular (that is, if the direct implementation of the IGCG(A^TY) method is $(n^*, u^{(0)})$-computable). (The IGCG(A^TY) method is $(n^*, u^{(0)})$-computable if and only if the ORTHODIR(A^TY) method is $(n^*, u^{(0)})$-computable.)

In a practical implication, if Phase I of the MGMRES method does not breakdown, and if the IGCG(A^TY) method is $(n^*, r^{(0)})$-computable, then SO is the MGMRES method.

5. Lan/mgmres Method

We now sketch a Lanczos-type method based on the MGMRES procedure. Consider the *double system*

$$\left\{ \begin{array}{rcl} A_u & = & b \\ A^T\tilde{u} & = & \tilde{b}. \end{array} \right.$$

Here the second equation is called the shadow system for some $\tilde{b}$. We write the double system as

$$\mathcal{A}\mathcal{U} = \mathcal{B},$$

where

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}, \quad \mathcal{U} = \begin{bmatrix} u \\ \tilde{u} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} b \\ \tilde{b} \end{bmatrix}.$$

We can select Z as either of the following symmetric matrices

$$\mathcal{Y} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \qquad \mathcal{Y}\mathcal{A} = \begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}.$$

To apply the MGMRES method, let $Z = A^T Y$ where $A = \mathcal{A}$ and $Y = \mathcal{Y}$.

Related Lanczos methods are the LANDIR($\mathcal{Y}$) method, the LANDIR($\mathcal{A}^T\mathcal{Y}$) method (equivalently, the LAN/MGMRES method), the LANMIN(Y) method (equivalently, the BCG method), the LANMIN($\mathcal{A}^T\mathcal{Y}$) method, the LANRES($\mathcal{Y}$) method, and the LANRES($\mathcal{A}^T\mathcal{Y}$) method.

We discuss the motivation for the LAN/MGMRES method. Let $Z = A^T Y$ and Y is a SPD matrix. The methods ORTHODIR(Y) and ORTHODIR(A^TY) are more robust than the methods ORTHOMIN(Y) and ORTHOMIN(A^TY), respectively, but they are often numerically unstable. The GMRES(A^TY) method is mathematically equivalent to the ORTHODIR(A^TY) method, but is more stable and requires less work per iteration.

Let

$$\mathcal{Z} = \mathcal{A}^T\mathcal{Y}, \qquad \mathcal{Y} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

In theory, the methods LANDIR($\mathcal{Y}$) and LANDIR($\mathcal{A}^T\mathcal{Y}$) are more robust than the ORTHOMIN($\mathcal{Y}$) method (equivalently, the BCG method) and the method ORTHOMIN ($\mathcal{A}^T\mathcal{Y}$), respectively, but they are often numerically unstable. The method LAN/MGMRES (equivalently, the MGMRES($\mathcal{A}^T\mathcal{Y}$) method) is *almost* equivalent to the LANDIR($\mathcal{A}^TY$) method and is hopefully more stable. (However, an additional condition is needed so that Phase I of the LAN/MGMRES method can be carried out.)

We now outline Phase I of the LAN/MGMRES method. Let $u^{(0)}$ be arbitrary and compute $r^{(0)} = b - Au^{(0)}$. Let $\tilde{u}^{(0)}$ be arbitrary or set $\tilde{u}^{(0)} = u^{(0)}$ for the shadow system and compute $\tilde{r}^{(0)} = \tilde{b} - A^T\tilde{u}^{(0)}$. Then let

$$\begin{cases} \widehat{w}^{(0)} & = & r^{(0)} \\ \widehat{\tilde{w}}^{(0)} & = & \tilde{r}^{(0)}. \end{cases}$$

Next set

$$s_0 = 2\langle \widehat{w}^{(0)}, \widehat{\tilde{w}}^{(0)} \rangle.$$

(The process fails if $s_0 = 0$.) Then set $\sigma_0 = \sqrt{|s_0|}$ and compute

$$\begin{cases} w^{(0)} & = & {\sigma_0}^{-1}\widehat{w}^{(0)} \\ \tilde{w}^{(0)} & = & {\sigma_0}^{-1}\widehat{\tilde{w}}^{(0)} \end{cases}$$

and

$$\begin{cases} \widehat{w}^{(n)} &= Aw^{(n-1)} + \beta_{n,n-1}w^{(n-1)} + \beta_{n,n-2}w^{(n-2)} \\ \widehat{\tilde{w}}^{(n)} &= A^T\tilde{w}^{(n-1)} + \beta_{n,n-1}\tilde{w}^{(n-1)} + \beta_{n,n-2}\tilde{w}^{(n-2)}. \end{cases}$$

Now we have

$$\langle w^{(n)}, \tilde{w}^{(n-1)}\rangle = \langle w^{(n)}, \tilde{w}^{(n-2)}\rangle = 0.$$

And set $s_n = 2\langle \widehat{w}^{(n)}, \widehat{\tilde{w}}^{(n)}\rangle$. (Process fails if $s_n = 0$.) Set $\sigma_n = \sqrt{|s_n|}$. Finally, we have

$$\begin{cases} w^{(n)} &= {\sigma_n}^{-1}\widehat{w}^{(n)} \\ \tilde{w}^{(n)} &= {\sigma_n}^{-1}\widehat{\tilde{w}}^{(n)}. \end{cases}$$

We now outline Phase II of the LAN/MGMRES method just for the non-shadow system. We have

$$\begin{aligned} u^{(n)} &= u^{(0)} + c_0 w^{(0)} + c_1 w^{(1)} + \cdots + c_{n-1}^{(n)} w^{(n-1)} \\ &= u^{(0)} + W_{n-1}c^{(n)} \\ &= \tilde{u}^{(n)} + \widetilde{W}_{n-1}c^{(n)}. \end{aligned}$$

The last equation is for the shadow system. Here

$$\begin{aligned} W_{n-1} &= \begin{bmatrix} w^{(0)} & w^{(1)} & \cdots & w^{(n-1)} \end{bmatrix}, \\ \widetilde{W}_{n-1} &= \begin{bmatrix} \tilde{w}^{(0)} & \tilde{w}^{(1)} & \cdots & \tilde{w}^{(n-1)} \end{bmatrix}, \\ c^{(n)} &= \begin{bmatrix} c_0^{(n)} & c_1^{(n)} & \cdots & c_{n-1}^{(n)} \end{bmatrix}^T. \end{aligned}$$

So

$$H_n^H D_n H_n c^{(n)} = H_n^T D_n e^{(n+1)}.$$

Example $(n = 2)$:

$$H_2 = \begin{bmatrix} -\beta_{1,0} & -\beta_{2,0} \\ \sigma_1 & -\beta_{2,1} \\ 0 & \sigma_2 \end{bmatrix}, \quad D_2 = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}, \quad (d_i = \pm 1),$$

$$c^{(2)} = \begin{bmatrix} c_0^{(2)} \\ c_1^{()2} \end{bmatrix}, \quad e^{(3)} = \begin{bmatrix} \sigma_2 \\ 0 \\ 0 \end{bmatrix}.$$

Use Givens rotations to find Q with $QQ^T = I$ and apply it to

$$QH_2 = \tilde{H}_2 = \begin{bmatrix} \boxed{\times} & \times \\ 0 & \boxed{\times} \\ 0 & 0 \end{bmatrix}.$$

Solve for $c^{(2)}$ in

$$\tilde{H}_2^T Q D_2 Q^T \tilde{H}_2 c^{(2)} = \tilde{H}_2^T Q D_2 e^{(3)}.$$

This process may fail. However, if Phase I is computable, then Phase II is computable if and only if LAN/IGCG ($\mathcal{A}^T\mathcal{Y}$) is computable. ∎

Additional details on the methods sketched in this paper can be found in [1, 2].

Acknowledements

This work was supported, in part, by grant TARP-197 from the State of Texas through the Texas Advanced Research Project at The University of Texas at Austin. This paper is based on notes of talks given by Professor David M. Young.

References

1. J.-Y. Chen, Iterative solution of large nonsymmetric linear systems, Report CNA-285, Center for Numerical Analysis, University of Texas at Austin, 1997.
2. J.-Y. Chen, D. R. Kincaid and D. M. Young, Generalizations and modifications of the GMRES iterative method, *Numer. Algorithms* **21** (1999), 119-146.
3. K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, Report CNA-176, Center for Numerical Analysis, University of Texas at Austin, 1982.
4. K. C. Jea and D. M. Young, On the simplification of generalized conjugate gradient methods for nonsymmetrizable linear systems, *Linear Algebra Appl.* **52/53** (1983), 399-417.
5. Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publisher, Boston, MA, 1996.
6. Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SlAM J. Sci. Comput.* **7** (1986), 856-869.
7. D. M. Young, L. J. Hayes and K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, Part I: The nonsymmetrizable case, Report CNA-162, Center for Numerical Analysis, University of Texas at Austin, 1981.
8. D. M. Young and K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, *Linear Algebra Appl.* **34** (1980), 159-194.
9. D. M. Young and K. C. Jea, Generalized conjugate gradient acceleration of iterative methods, Part 11: The nonsymmetrizable case, Report CNA-163, Center for Numerical Analysis, University of Texas at Austin, 1981.

David R. Kincaid and David M. Young
Center for Numerical Analysis, University of Texas at Austin
Austin, TX 78713-8510 USA
E-mail: kincaid@cs.utexas.edu

Jen-Yuan Chen
Department of Applied Mathematics, I-Shou University
Ta-Shi, Kaohsiung, Taiwan 840, R.O.C.
E-mail: jchen@isu.edu.tw