Applied Mathematics and Computation 206 (2008) 42–49

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A variant algorithm of the ORTHOMIN(m) method for solving linear systems

Kuniyoshi Abe [a,*], Shao-Liang Zhang [b]

[a] Faculty of Economics and Information, Gifu Shotoku University, Nakauzura, Gifu 500-8288, Japan
[b] Graduate School of Engineering, Nagoya University, Furocho, Nagoya 464-8603, Japan

ARTICLE INFO

Keywords:
Linear systems
Krylov subspace method
ORTHOMIN(m) method
Singular matrices

ABSTRACT

We propose a variant of ORTHOMIN(m) for solving linear systems $A\boldsymbol{x} = \boldsymbol{b}$. It is mathematically equivalent to the original ORTHOMIN(m) method, but uses recurrence formulas that are different from those of ORTHOMIN(m); they contain alternative expressions for the auxiliary vectors and the recurrence coefficients. Our implementation has the same computational costs as ORTHOMIN(m). As a result of numerical experiments on nonsingular linear systems, we have confirmed the equivalence of our proposed variant of ORTHOMIN(m) with the original ORTHOMIN(m) using finite precision arithmetic; numerical experiments on singular linear systems show that our proposed algorithm is more accurate and less affected by rounding errors than the original ORTHOMIN(m).

1. Introduction

We consider the solution of the following linear system of equations via Krylov subspace methods:

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

where A is a large sparse n-by-n matrix and $\boldsymbol{b}$ is a n-vector.

The Bi-Conjugate Gradient (BiCG) [1] method is a well-known Krylov subspace method derived from the Petrov–Galerkin approach [2]. It is used for solving a linear system (1) with a nonsingular and nonsymmetric coefficient matrix. Recently, a number of hybrid BiCG methods, such as the BiCG STABilized (BiCGSTAB) [3], BiCGStab2 [4], Generalized Product-type BiCG (GPBiCG) [5] and BiCGstab(l) [6] methods, have been developed to make the convergence faster and smoother than that of BiCG. Krylov subspace methods derived from the minimum residual approach [2], such as the Generalized Conjugate Residual (GCR) [7] method, have also been proposed. Since the storage requirements and computational costs can become expensive for GCR with the increase in the number of iterations, two implementations of GCR have been commonly employed. One is a restarted version in which the iteration is restarted every m steps by using the approximate solution vector obtained after the previously implemented m steps as the initial vector. The other is a truncated version in which the subsequent iteration proceeds by using long recurrence formulas, in which only a new set of m previous vectors and constants obtained after each iteration is stored and the other vectors and constants are discarded. The restarted and truncated versions of GCR are known as GCR(m) and ORTHOMIN(m) [8], respectively. In addition to GCR, the Conjugate Residual (CR) [7] method for solving a linear system with a symmetric matrix also falls within the class of minimum residual solvers [2]. In CR, the residual and the approximation are updated by using short recurrence formulas, in which the preceding two vectors and constants are stored.

Thus, the truncated minimum residual solvers can be classified either as, (i) the CR method, which uses short recurrence formulas to update the residual and the approximation, or (ii) the ORTHOMIN(m) method, which uses the long recurrence

* Corresponding author.
E-mail address: abe@gifu.shotoku.ac.jp (K. Abe).

0096-3003/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.amc.2008.08.035

formulas to update the residual and the approximation. The convergence behavior of CR (or equivalently denoted by ORTHOMIN(2)) is different from that of ORTHOMIN(m) ($m\neq 2$). A variant of ORTHOMIN(2) [9], which is mathematically equivalent to the original ORTHOMIN(2) but different in its implementation, has been recently reported. However, a variant of ORTHOMIN(m) ($m\neq 2$) has not been proposed so far, in spite of the fact that the convergence is different between ORTHOMIN(2) and ORTHOMIN(m) ($m\neq 2$), as mentioned above. In addition, it has not been clarified whether the merit described in [9] for $m = 2$ can be attained by using the variant of ORTHOMIN(m) for $m\neq 2$.

Therefore, a variant of ORTHOMIN(m) in which the residual and the approximation are updated by using long recurrence formulas has to be developed, and the convergence behavior of the variant of ORTHOMIN(m) on both nonsingular and singular systems has to be examined. In the present paper, we propose a variant of ORTHOMIN(m) for solving a linear system with a nonsymmetric coefficient matrix by using the same analogy as that discussed in [9]. Our proposed algorithm is mathematically equivalent to the original ORTHOMIN(m); however, the recurrence formulas, in which alternative expressions for the auxiliary vectors and the recurrence coefficients are introduced, are different from those of ORTHOMIN(m). Our implementation of ORTHOMIN(m) requires the same computational costs as the original ORTHOMIN(m) algorithm. By numerical experiments on nonsingular linear systems, we confirm that our implementation is equivalent to ORTHOMIN(m) in finite precision computations. Numerical experiments on singular systems demonstrate that our implementation is more effective than the standard ORTHOMIN(m).

In the following section, the standard ORTHOMIN(m) algorithm is described, and we derive a variant implementation of ORTHOMIN(m), providing the details for the variant algorithm. We shall compare the computational costs of our variant implementation of ORTHOMIN(m) with the original. Through numerical experiments on nonsingular systems, we demonstrate that our algorithm is equivalent to the original in finite precision computations. In Section 3, the convergence of ORTHOMIN(m) as well as an error bound for ORTHOMIN(m) applied to singular linear systems are described. Numerical experiments on singular linear systems show that our proposed variant of ORTHOMIN(m) is less affected by rounding errors than the original.

2. ORTHOMIN(m) algorithm

In this section, we review the ORTHOMIN(m) algorithm and we propose a modification of the recurrence in ORTHOMIN(m) by using the same analogy as that used to derive ORTHOMIN(2) in [9]. Furthermore, we take model problems with nonsingular matrices and examine the convergence behavior of our variant with that of the original.

2.1. A variant of ORTHOMIN(m)

According to [8], the ORTHOMIN(m) algorithm is defined as follows.

Algorithm 1 (*ORTHOMIN(m)*). Let $\boldsymbol{x}_0$ be an initial guess, and put $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$. Set $\beta_{-1,j} = 0$.
For $k = 0, 1, \ldots$ repeat the following steps until the condition $\|\boldsymbol{r}_k\|_2 \leqslant \varepsilon_{TOL}\|\boldsymbol{r}_0\|_2$ holds:

$$\boldsymbol{q}_k = A\boldsymbol{r}_k + \sum_{j=k-m}^{k-1} \beta_{k-1,j}\boldsymbol{q}_j, \tag{2}$$

$$\boldsymbol{p}_k = \boldsymbol{r}_k + \sum_{j=k-m}^{k-1} \beta_{k-1,j}\boldsymbol{p}_j,$$

$$\alpha_k = \frac{(\boldsymbol{r}_k, \boldsymbol{q}_k)}{(\boldsymbol{q}_k, \boldsymbol{q}_k)},$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k\boldsymbol{p}_k,$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k\boldsymbol{q}_k, \tag{3}$$

$$\beta_{k,j} = -\frac{(A\boldsymbol{r}_{k+1}, \boldsymbol{q}_j)}{(\boldsymbol{q}_j, \boldsymbol{q}_j)} \quad (k - m + 1 \leqslant j \leqslant k).$$

The recurrence coefficients α_k and $\beta_{k-1,j}$ of ORTHOMIN(m) are replaced by the parameters ζ_k and $\frac{\zeta_{k-1}}{\zeta_k}\eta_{k,j+1}$, respectively. The formulas (2) and (3) for updating the residual become

$$\zeta_k\boldsymbol{q}_k = \zeta_k A\boldsymbol{r}_k + \zeta_{k-1}\sum_{j=k-m+1}^{k} \eta_{k,j}\boldsymbol{q}_{k-1}, \tag{4}$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \zeta_k\boldsymbol{q}_k. \tag{5}$$

By introducing a new auxiliary vector as

$$\boldsymbol{y}_k := \zeta_{k-1}\boldsymbol{q}_{k-1},$$

the formulas (4) and (5) are converted into

$$\boldsymbol{y}_{k+1} = \zeta_k A\boldsymbol{r}_k + \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{y}_j, \tag{6}$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \boldsymbol{y}_{k+1}. \tag{7}$$

Moreover, when introducing a new auxiliary vector as

$$\boldsymbol{z}_k = \zeta_{k-1}\boldsymbol{p}_{k-1}, \tag{8}$$

the approximation $\boldsymbol{x}_k$ and the direction $\boldsymbol{p}_k$ are updated by

$$\boldsymbol{z}_{k+1} = \zeta_k \boldsymbol{r}_k + \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{z}_j, \tag{9}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{z}_{k+1}. \tag{10}$$

In Orthomin(m) the recurrence coefficients α_k and $\beta_{k-1,j}$ are determined by minimizing the residual norm in the space spanned by $A\boldsymbol{r}_k$ and $\boldsymbol{q}_{j-1}$ ($j = k-m+1,\ldots,k$) [10]. This is equivalent to minimizing the residual norm in the space spanned by $A\boldsymbol{r}_k$ and $\boldsymbol{y}_j$ ($j = k-m+1,\ldots,k$) in our case since the vector $\zeta_{k-1}\boldsymbol{q}_{k-1}$ is replaced by $\boldsymbol{y}_k$. That is, the recurrence coefficients ζ_k and η_{kj} are the solution of the minimization problem

$$\|\boldsymbol{r}_{k+1}\|_2 = \min_{\zeta_k,\eta_{kj}} \left\| \boldsymbol{r}_k - \zeta_k A\boldsymbol{r}_k - \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{y}_k \right\|_2. \tag{11}$$

The condition (11) reveals the orthogonalities for $\boldsymbol{r}_k$ and $\boldsymbol{y}_i$ as

$$(\boldsymbol{r}_{k+1}, A\boldsymbol{r}_k) = 0, \tag{12}$$

$$(\boldsymbol{r}_{k+1}, \boldsymbol{y}_i) = 0 \quad (j = k-m+1,\ldots,k). \tag{13}$$

The orthogonalities (12) and (13) and the expressions (6) and (7) imply

$$(\boldsymbol{r}_k, \boldsymbol{y}_i) = (\boldsymbol{y}_{k+1}, \boldsymbol{y}_i) = 0 \quad (j = k-m+1,\ldots,k). \tag{14}$$

In addition, by the formulas (6) and (7), the inner product $(\boldsymbol{y}_k, \boldsymbol{y}_k)$ is rewritten as

$$(\boldsymbol{y}_k, \boldsymbol{y}_k) = (\boldsymbol{r}_{k-1} - \boldsymbol{r}_k, \zeta_{k-1}A\boldsymbol{r}_{k-1} + \sum_{j=k-m}^{k-1} \eta_{k-1,j}\boldsymbol{y}_j),$$

and the properties (12)–(14) lead to

$$(\boldsymbol{y}_k, \boldsymbol{y}_k) = \zeta_{k-1}(\boldsymbol{r}_{k-1}, A\boldsymbol{r}_{k-1}). \tag{15}$$

The recurrence coefficients ζ_k and η_{kj} are the solution of the dimensional minimization problem (11), namely, the recurrence coefficients can be computed by

$$\frac{\partial}{\partial \zeta_k} \left\| \boldsymbol{r}_k - \zeta_k A\boldsymbol{r}_k - \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{y}_j \right\|_2^2 = 0,$$

$$\frac{\partial}{\partial \eta_{kj}} \left\| \boldsymbol{r}_k - \zeta_k A\boldsymbol{r}_k - \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{y}_j \right\|_2^2 = 0.$$

By combining the above conditions together with properties (12)–(15), the calculation for the recurrence coefficients ζ_k and η_{kj} can be carried out through

$$\zeta_k = \frac{(A\boldsymbol{r}_k, \boldsymbol{r}_k)}{(A\boldsymbol{r}_k, A\boldsymbol{r}_k) - \sum_{j=k-m+1}^{k} \frac{1}{v_j}(A\boldsymbol{r}_k, \boldsymbol{y}_j)(A\boldsymbol{r}_k, \boldsymbol{y}_j)}, \tag{16}$$

$$\eta_{kj} = -\frac{\zeta_k}{v_j}(\boldsymbol{y}_j, A\boldsymbol{r}_k) \quad (j = k-m+1,\ldots,k). \tag{17}$$

Here v_k is defined by $v_k := \zeta_{k-1}(A\boldsymbol{r}_{k-1}, \boldsymbol{r}_{k-1})$.

Summing up the expressions (6), (7), (9), (10), (16) and (17), we propose the following variant of Orthomin(m).

Algorithm 2 (*Proposed implementation*). Let $\boldsymbol{x}_0$ be an initial guess, and put $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$. Set $v_0 = (\boldsymbol{y}_0, \boldsymbol{y}_0)$.

Table 1
Computational costs per iteration

Method	Matrix-vector products	Inner products	`Add or scaling`
Variant	1	$3 + m$	$4 + 4m$
Orthomin(m)	1	$3 + m$	$4 + 4m$

`Add or scaling`: Adding a vector to another vector or scalar multiplication of a vector.

For $k = 0, 1, \ldots$ until the condition $\|\boldsymbol{r}_k\|_2 \leqslant \varepsilon_{TOL}\|\boldsymbol{r}_0\|_2$ holds, iterate:

$$\zeta_k = \frac{(A\boldsymbol{r}_k, \boldsymbol{r}_k)}{(A\boldsymbol{r}_k, A\boldsymbol{r}_k) - \sum_{j=k-m+1}^{k} \frac{1}{\nu_j}(A\boldsymbol{r}_k, \boldsymbol{y}_j)(A\boldsymbol{r}_k, \boldsymbol{y}_j)},$$

$$\eta_{kj} = -\frac{\zeta_k}{\nu_j}(\boldsymbol{y}_j, A\boldsymbol{r}_k) \quad (k - m + 1 \leqslant j \leqslant k),$$

$$\left(\text{for } k = 0, \quad \zeta_k = \frac{(A\boldsymbol{r}_k, \boldsymbol{r}_k)}{(A\boldsymbol{r}_k, A\boldsymbol{r}_k)}, \quad \eta_k = 0\right)$$

$$\nu_{k+1} = \zeta_k(A\boldsymbol{r}_k, \boldsymbol{r}_k),$$

$$\boldsymbol{z}_{k+1} = \zeta_k\boldsymbol{r}_k + \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{z}_j,$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{z}_{k+1},$$

$$\boldsymbol{y}_{k+1} = \zeta_k A\boldsymbol{r}_k + \sum_{j=k-m+1}^{k} \eta_{kj}\boldsymbol{y}_j,$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \boldsymbol{y}_{k+1}.$$

We summarize the computational costs per iteration in Table 1, which shows that our variant implementation of Orthomin(m) costs the same as the original Orthomin(m).

2.2. Numerical experiments on nonsingular linear systems

In this subsection, we present numerical experiments on model problems with nonsingular matrices and compare the convergence behavior of the variant with that of the original.

Throughout this paper numerical calculations were carried out in double-precision floating-point arithmetic on a PC with an Intel Xeon 3.06 GHz processor equipped with a Fujitsu Fortran compiler.

2.2.1. Model problem

We solve a linear system with a nonsingular and nonsymmetric coefficient matrix derived from the 5-point central difference approximation of the two-dimensional convection-diffusion equation:

$$-\Delta u + \gamma\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right) + \beta\pi^2 u = f(x, y). \tag{18}$$

(18) is solved on the unit square $\Omega = (0, 1) \times (0, 1)$ with zero Dirichlet boundary conditions [11]. The right-hand side vector is determined by substituting a vector $\hat{\boldsymbol{x}} = (1, \ldots, 1)^{\mathrm{T}}$ into the equation $\boldsymbol{b} = A\hat{\boldsymbol{x}}$. The mesh size is chosen as $h = 1/101$ $(= 1/(M + 1))$ for Ω, so that the resulting linear system has a $M^2 \times M^2$ coefficient matrix. The numerical experiments were done for 24 combinations of $\gamma = 1, 3, 5, 10, 20, 30$ and $\beta = -1, -3, -5, -10$. The iteration was started with $\boldsymbol{x}_0 = (1, 2, \ldots, n)^{\mathrm{T}}$. Orthomin($m$) and the variant implementation were restarted every 15 iterations.

2.2.2. Numerical results

In Fig. 1, we show representative results for the combinations of γ and β given by: $(\gamma, \beta) = (3, -10)$ and $(20, -5)$. The convergence plots show the iteration counts (on the horizontal axis) versus the relative residual 2-norms $(\log_{10}\|\boldsymbol{r}_k\|_2/\|\boldsymbol{r}_0\|_2)$. Here the residual $\boldsymbol{r}_k$ is recursively updated in these algorithms. The symbols Variant and Orthomin(m) stand for the convergence behavior of the variant of Orthomin(m) and that of the original Orthomin(m), respectively.

The convergence behavior of our algorithm coincides with that of the original for all the combinations of the above mentioned parameters γ and β. Therefore, we assert that our variant implementation is equivalent to the original Orthomin(m) algorithm, when using finite precision computations, on nonsingular linear systems.

3. Application to singular linear systems

In this section, convergence properties and error bound properties of Orthomin(m), when applied to a singular linear system, are described. Next the original Orthomin(m) algorithm and our variant implementation are applied to a model problem with a singular matrix. It is shown that the variant method is less affected by rounding errors than the original Orthomin(m) method.

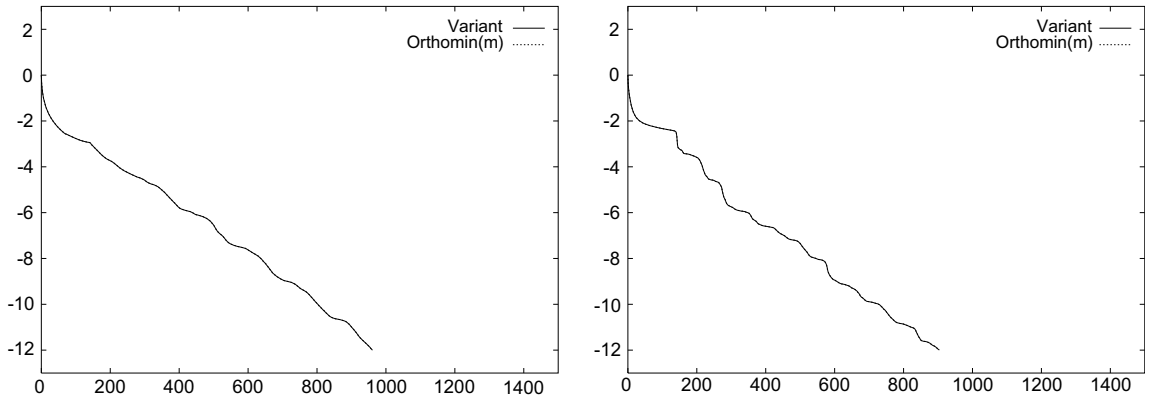

Fig. 1. Convergence history of the variant implementation (restarted every 15 iterations) and the original ORTHOMIN(15) for the linear system with $(\gamma,\beta)=(3,-10)$ (on the left) and $(\gamma,\beta)=(20,-5)$ (on the right).

Throughout this section the following notation is used for any n-by-n matrix X.

$\mathrm{Range}(X) := \{\boldsymbol{y}\in\mathbb{R}^n \mid \boldsymbol{y}=X\boldsymbol{z} \quad \text{for} \quad \boldsymbol{z}\in\mathbb{R}^n\}$: The rangespace of X.
$\mathrm{Null}(X) := \{\boldsymbol{z}\in\mathbb{R}^n \mid X\boldsymbol{z}=\boldsymbol{0}\}$: The nullspace of X.
$\mathrm{rank}X := \dim(\mathrm{Range}(X))$: The rank of X.
$M(X) := \frac{X+X^{\mathrm{T}}}{2}$: The symmetric part of X.
$R(X) := \frac{X^{\mathrm{T}}-X}{2}$: The skew-symmetric part of X.
$\lambda_{\max}(X)$: The non-zero eigenvalue of the largest magnitude.
$\lambda_{\min}(X)$: The non-zero eigenvalue of the smallest magnitude.
$\rho(X) := |\lambda_{\max}(X)|$: The spectral radius of X.

3.1. Convergence on singular linear systems

In this subsection, we treat the singular case of (1), where a unique solution may not exist. For the case of applying ORTHOMIN(m) to a singular linear system, the least squares problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b}-A\boldsymbol{x}\|_2 \tag{19}$$

is solved instead of (1). When $\bar{\boldsymbol{x}}$ denotes a solution of the system (19), the corresponding residual $\bar{\boldsymbol{r}}=\boldsymbol{b}-A\bar{\boldsymbol{x}}$ is always unique, and the identity $A^{\mathrm{T}}\bar{\boldsymbol{r}}=0$ holds. Thus the problem can be called a minimum residual problem, because $\bar{\boldsymbol{r}}$ gives the minimum residual. ORTHOMIN(m) mathematically converges to a least squares solution under certain conditions described in [12–15] when applied to the system (19). Several theoretical results which have already been known are described below.

First, μ ($n>\mu$) denotes the rank of the coefficient matrix, and Q_1 is an $n\times\mu$ matrix having orthonormal basis $\boldsymbol{q}_1,\boldsymbol{q}_2,\ldots,\boldsymbol{q}_\mu$ of Range(A) as its columns. Then the following necessary and sufficient condition for ORTHOMIN(m) to converge to the least squares solution without breakdown has already been clarified [12,13,15].

Theorem 1 ([12,13,15]). *CI is equivalent to CII.*

(CI) *ORTHOMIN(m) converges to the least squares solution without breakdown for an arbitrary right hand-side vector $\boldsymbol{b}$ and an initial vector $\boldsymbol{x}_0$.*
(CII) *$M(Q_1^{\mathrm{T}}AQ_1)$ is definite, and $\mathrm{Null}(A)=\mathrm{Null}(A^{\mathrm{T}})$ holds.*

Remark 1 ([13,15]). "$M(Q_1^{\mathrm{T}}AQ_1)$ is definite" is equivalent to "$M(A)$ is semi-definite and $\mathrm{rank}A=\mathrm{rank}M(A)$ holds" under the condition $\mathrm{Null}(A)=\mathrm{Null}(A^{\mathrm{T}})$.

Next, the convergence as well as an error bound of ORTHOMIN(m) have been established in [15], which contains the following result.

Theorem 2. *Let $\boldsymbol{r}_k$ and $\bar{\boldsymbol{r}}$ be the k-th residual and the minimum residual, respectively, generated by ORTHOMIN(m). If $\mathrm{Null}(A)=\mathrm{Null}(A^{\mathrm{T}})$ and if $M(A)$ is positive semi-definite, then the following inequality is valid:*

$$\frac{\|\boldsymbol{r}_{k+1}-\bar{\boldsymbol{r}}\|_2}{\|\boldsymbol{r}_k-\bar{\boldsymbol{r}}\|_2} \leqslant \sqrt{1-\frac{\lambda_{\min}^2(M(A))}{\lambda_{\max}(M(A))\lambda_{\min}(M(A))+\rho(R(A))^2}}.$$

Note that when $\bar{\boldsymbol{r}}=0$, that is, in the nonsingular case, the above result reduces to that of [7].

Since the mathematical equivalence between our proposed algorithm and the original ORTHOMIN(m) algorithm has already been established, then it follows that our variant algorithm and the original ORTHOMIN(m) algorithm share the same convergence and error bound properties for singular systems.

3.2. Numerical experiments on singular linear systems

In this section, we present numerical experiments on a model problem with a singular matrix, and reveal the merit of our proposed variant of ORTHOMIN(m).

3.2.1. Model problem

As shown in [16], applying 5-point central differences to the partial differential equation

$$\Delta u + d\frac{\partial u}{\partial x} = f(x,y), \quad 0 < x,y < 1 \tag{20}$$

over the unit square $\Omega=(0,1)\times(0,1)$ with the periodic boundary conditions $u(x,0)=u(x,1)$, $u(0,y)=u(1,y)$ yields a singular linear system with a nonsymmetric coefficient matrix. The mesh size is chosen as $h=1/100$ $(=1/M)$ for Ω, so that the resulting linear system has the following $M^2\times M^2$ coefficient matrix:

$$A := \frac{1}{h^2}\begin{bmatrix} T_M & I_M & & I_M \\ I_M & \ddots & \ddots & \\ & \ddots & \ddots & I_M \\ I_M & & I_M & T_M \end{bmatrix}. \tag{21}$$

Here I_M is the $M\times M$ unit matrix and T_M the $M\times M$ matrix given by

$$T_M := \begin{bmatrix} -4 & \alpha_+ & & \alpha_- \\ \alpha_- & \ddots & \ddots & \\ & \ddots & \ddots & \alpha_+ \\ \alpha_+ & & \alpha_- & -4 \end{bmatrix},$$

where $\alpha_\pm = 1 \pm \frac{dh}{2}$. The numerical experiment was done for $d=0$ and 0.3.

For the matrix A the identity $A\boldsymbol{e}=A^{\mathrm{T}}\boldsymbol{e}=\boldsymbol{0}$ holds, so that $\mathrm{Null}(A)=\mathrm{Null}(A^{\mathrm{T}})=\mathrm{Span}\{\boldsymbol{e}\}$, where $\boldsymbol{e}\equiv(1,1,\ldots,1)^{\mathrm{T}}$. The symmetric part $M(A)$ of (21) is readily verified to be positive semi-definite, and $\mathrm{rank}A=\mathrm{rank}M(A)=n-1$ holds. Since this matrix A satisfies the hypotheses of Theorem 1 (CII) and Theorem 2, our variant implementation and the original ORTHOMIN(m) can be applied to this singular system. When the coefficient matrix is singular, we can choose $\boldsymbol{b}_R\in\mathrm{Range}(A)$ and $\boldsymbol{b}_N\in\mathrm{Null}(A^{\mathrm{T}})$ so that the right-hand side vector is expressed as $\boldsymbol{b}=\boldsymbol{b}_R+\boldsymbol{b}_N$. Also, the direct sum decomposition $\mathbb{R}^n=\mathrm{Range}(A)\oplus\mathrm{Null}(A)$ holds under the condition $\mathrm{Null}(A)=\mathrm{Null}(A^{\mathrm{T}})$. In solving the singular linear system, the selection $\boldsymbol{b}_N\neq\boldsymbol{0}$ yields

$$\|\boldsymbol{r}_k\|_2^2=\|\boldsymbol{b}-A\boldsymbol{x}_k\|_2^2=\|\boldsymbol{b}_N\|_2^2+\|\boldsymbol{b}_R-A\boldsymbol{x}_k\|_2^2.$$

Hence, the residual norm $\|\boldsymbol{r}_k\|_2$ is expected to converge to the minimum residual norm $\|\boldsymbol{b}_N\|_2$ $(=\|\bar{\boldsymbol{r}}\|_2)$ and should keep the level.

We now consider a perturbed system using our choice for a singular matrix A: First the unperturbed right-hand side vector is taken from the rangespace of A by substituting a random vector $\tilde{\boldsymbol{x}}$ into the equation $\boldsymbol{b}_R=A\tilde{\boldsymbol{x}}$. Next, a perturbation is made to yield the right-hand side

$$\boldsymbol{b}=\boldsymbol{b}_R+\delta\frac{\boldsymbol{e}}{\|\boldsymbol{e}\|_2}.$$

Consequently the system

$$A\boldsymbol{x}=\boldsymbol{b}_R+\delta\frac{\boldsymbol{e}}{\|\boldsymbol{e}\|_2} \tag{22}$$

is solved for $\boldsymbol{x}$. The iteration was started with $\boldsymbol{x}_0=\boldsymbol{0}$. Our variant implementation and the original ORTHOMIN(m) are restarted every 30 iterations. The perturbation parameter δ is selected as 10^{-6} in our experiments. Although the iterations are normally truncated when the residual norm $\|\boldsymbol{r}_k\|_2$ comes to 10^{-6}, we continued the iterations until 1500 and 3000 steps for

the singular system (22) with $d = 0$ and 0.3, respectively. Since this is a minimal residual problem with $\bar{\boldsymbol{r}} \neq \boldsymbol{0}$, a proper stopping criterion is unknown. So, we execute a sufficiently many number of iterations.

3.2.2. Numerical results

The convergence behaviors when applying the variant implementation and when applying the original ORTHOMIN(m) to (22) are examined. Since the matrix of (21) satisfies the hypotheses of Theorem 1 and Theorem 2, the residual norms of the two implementations should theoretically converge to the minimum residual norm.

The numerical results when applying our variant implementation and when applying the original ORTHOMIN(m) to (22) with $d = 0$ and 0.3 are displayed in Fig. 2. The convergence plots show the iteration counts (on the horizontal axis) versus the residual 2-norms ($\log_{10}\|\boldsymbol{r}_k\|_2$). Here the residual $\boldsymbol{r}_k$ is recursively updated in these algorithms. The symbols Variant and ORTHOMIN(m) in Fig. 2 stand for the convergence behavior of the variant of ORTHOMIN(m) and that of the original ORTHOMIN(m), respectively.

Table 2 shows the iteration counts at which the residual norm of ORTHOMIN(m) becomes 10^{-6} or less.

Also, we compute the explicitly computed residual norm $\|\boldsymbol{b} - A\boldsymbol{x}_k\|_2$ in order to examine the accuracy of the approximate solution. Table 3 shows the explicitly computed residual norms of the variant and that of the original for a sufficiently large number of iterations.

From the figures and the tables we observe the following: the original implementation of ORTHOMIN(m) cannot keep the residual at the level of its theoretical minimum; the residual norm is reduced even further. This behavior is inconsistent with the analytical theory. We remark that in practical simulations, the residual norm should not fluctuate as this means that the solver cannot determine an appropriate stopping criterion. In other words the original implementation, when considered for a singular linear system, due to floating point arithmetic computations, cannot converge in some cases. On the other hand, our variant implementation keeps the residual at the expected level.

We have also observed, when applied to our singular test problem, that the explicitly computed residual norm using the original implementation of ORTHOMIN(m) becomes about 10^{-3} and 10^{-4} when the iterations are continued for a sufficiently long time. There is a wide gap between the convergence behaviors of the explicitly computed residuals vs. the recursively updated residuals for the original ORTHOMIN(m). On the other hand, our proposed variant implementation keeps the explicitly computed residual norm at 10^{-6}.

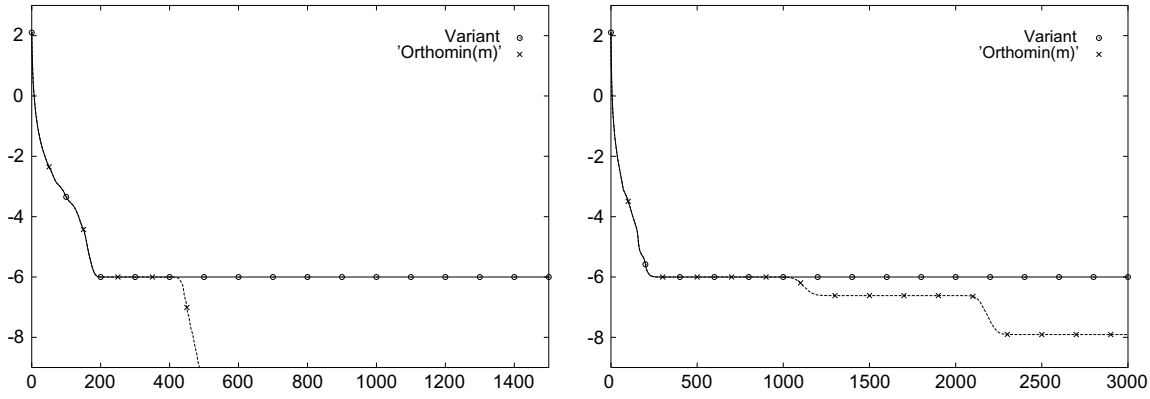

Fig. 2. Convergence history of the variant implementation (restarted every 30 iterations) and that of the original ORTHOMIN(30) for the linear system with $d = 0$ (on the left) and $d = 0.3$ (on the right).

Table 2
The iteration counts at which $\|\boldsymbol{r}_k\|_2$ of ORTHOMIN(m) becomes 10^{-6} or less in Fig. 2

	$\|\boldsymbol{r}_k\|_2 = 10^{-6}$	$\|\boldsymbol{r}_k\|_2 < 10^{-6}$
$d = 0$	220	332
$d = 0.3$	361	859

Table 3
The explicitly computed residual norms of the variant implementation (restarted every 30 iterations) and that of the original ORTHOMIN(30)

	Variant	ORTHOMIN(30)
$d = 0$	10^{-6}	$10^{-2.45}$
$d = 0.3$	10^{-6}	$10^{-3.80}$

4. Concluding remarks

We have derived a variant implementation of ORTHOMIN(m). This implementation is mathematically equivalent to the original ORTHOMIN(m), but it uses different recurrence formulas and a different auxiliary vector. Our implementation has the same computational costs as the original ORTHOMIN(m). Our numerical experiments on a nonsingular system show that our implementation is equivalent to ORTHOMIN(m) when implemented using finite precision arithmetic. On the other hand, our numerical experiments for a singular system show that our variant implementation is more accurate and less affected by rounding errors than ORTHOMIN(m).

It is important to establish a stopping criterion on our proposed variant of ORTHOMIN(m) so that the variant algorithm can be practical in solving inconsistent linear systems with a singular matrix as a referee suggested. Stopping criteria for conjugate gradient method [17] and for solving a singular linear system have been proposed in [18] and [19,20], respectively. Moreover, $\|A^{\mathrm{T}}\boldsymbol{r}_k\|$ can be considered to be a stopping criterion because it converges to zero when solving a singular linear system, although the computational costs become expensive. However, it is not easy to give an efficient and feasible stopping criterion for any Krylov subspace methods, including our proposed variant of ORTHOMIN(m), without introducing extra computational costs. Therefore, the stopping criterion needs to be designed on a future work.

Acknowledgements

We would like to express sincere thanks to Professor Martin H. Gutknecht for his insightful and fruitful suggestions. We are also thankful to Dr. Mark Sussman for his valuable comments. We would like to thank reviewers for their helpful and useful suggestions. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research(C), 18560064, 2008.

References

[1] R. Fletcher, Conjugate gradient methods for indefinite systems, in: G. Watson (Ed.), Numerical Analysis Dundee 1975, Lecture Notes in Mathematics, vol. 506, Springer-Verlag, 1976.
[2] H.G. Golub, H.A. van der Vorst, Closer to the solution: Iterative linear solvers, in: I.S. Duff, G.A. Watson (Eds.), The State of the Art in Numerical Analysis, IMA Conference Series, Clarendon Press, 1997.
[3] H.A. van der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 13 (1992) 631–644.
[4] M.H. Gutknecht, Variants of BiCGStab for matrices with complex spectrum, SIAM J. Sci. Comput. 14 (1993) 1020–1033.
[5] S.-L. Zhang, GPBi-CG: Generalized product-type methods based on Bi-CG for solving nonsymmetric linear systems, SIAM J. Sci. Comput. 18 (1997) 537–551.
[6] G.L.G. Sleijpen, D.R. Fokkema, BiCGstab(l) and other hybrid Bi-CG methods, Numer. Algor. 7 (1994) 75–109.
[7] S.C. Eisenstat, H.C. Elman, M.H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations, SIAM J. Numer. Anal. 20 (1983) 345–357.
[8] P.K.W. Vinsome, Orthomin, An iterative method for solving sparse sets of simultaneous linear equations, in: Proceedings of the Fourth Symposium on Reservoir Simulation, Society of Petroleum Engineers of AIME, SPE 5729, 1976.
[9] K. Abe, S.-L. Zhang, T. Mitsui, C.-H. Jin, A variant of the ORTHOMIN(2) method for singular linear systems, Numer. Algor. 36 (2004) 189–202.
[10] A. Greenbaum, Iterative Methods for Solving Linear Systems, Frontiers in Applied Mathematics, vol. 17, SIAM, Philadelphia, 1997. p. 220.
[11] Y. Saad, A flexible inner-outer preconditioned GMRES algorithm, SIAM J. Sci. Stat. Comput. 14 (1993) 461–469.
[12] K. Abe, H. Ogata, M. Sugihara, S.-L. Zhang, T. Mitsui, Convergence theory of the CR method for linear singular systems, Trans. Jpn. Soc. Indus. Appl. Math. 9 (1999) 1–13 (in Japanese).
[13] K. Hayami, On the convergence of the conjugate residual method for singular systems, Trans. Jpn. Soc. Indus. Appl. Math. 13 (2003) 1–33. in Japanese.
[14] K. Hayami, M. Sugihara, On the convergence of the GCR(k) method for singular systems, NII Technical Reports, National Institute of Informatics, NII(2004-009E), Tokyo, 2004.
[15] S.-L. Zhang, Y. Oyanagi, M. Sugihara, Necessary and sufficient conditions for the convergence of ORTHOMIN(k) on singular and inconsistent linear systems, Numer. Math. 87 (2000) 391–405.
[16] P.N. Brown, H.F. Walker, GMRES on (nearly) singular systems, SIAM J. Matrix Anal. Appl. 18 (1997) 37–51.
[17] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards 49 (1952) 409–435.
[18] H.A. van der Vorst, Iterative Krylov Methods for Large Linear Systems, Cambridge University Press, Cambridge, 2003. p.240.
[19] N. Zhang, Y. Wei, A Note on Solving EP Inconsistent linear systems, Appl. Math. Comput. 169 (2005) 8–15.
[20] A. Shiode, S. Fujino, K. Abe, Preconditioning for symmetric positive definite matrices for MRTR method, Trans. Jpn. Soc. Comput. Eng. Sci. (2006). No. 20060007 (in Japanese).