A GMRES(m) method with two stage deflated preconditioners

J. Shiroishi[1] T. Nodera[2]

(Received 5 February 2011; revised 4 June 2011)

Abstract

The GMRES(m) method is often used to compute Krylov subspace solutions of large sparse linear systems of equations. Morgan developed a new procedure that deflates the smallest eigenvalues and improves the eigenvalue distribution. Several preconditioning techniques have been explored in numerous research papers. In particular, the deflated GMRES proposed by Erhel and others replaces the smallest eigenvalues of the original coefficient matrix of the linear system with the largest modulus of the eigenvalues. We explore a new deflated GMRES which uses a two stage deflation technique. Further, the results of the numerical experiments for test matrices are tabulated to illustrate that our approach is effective in solving a wide range of problems.

Contents

http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/3984
gives this article, © Austral. Mathematical Soc. 2011. Published June 16, 2011. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

1 Introduction

GMRES [10] is used to solve large sparse nonsymmetric linear systems of equations expressed as

$$Ax = b\,, \quad A \in \mathbb{R}^{n\times n}, \quad x, b \in \mathbb{R}^n. \tag{1}$$

It is natural to consider that the convergence of GMRES depends on the distribution of eigenvalues of the coefficient matrix A, and that the existence of small eigenvalues slows down the convergence of GMRES. The convergence of GMRES occurs as though the small eigenvalues were removed from matrix A. This phenomenon is called superlinear convergence [11]. The Restarted GMRES method, with restart frequency m, is usually known as GMRES(m). This method is commonly used, because the full GMRES significantly increases computation cost and memory requirements as the number of iterations proceed. Unfortunately, the convergence of GMRES(m) is inferior to that of a full GMRES, because the information of eigenpairs, especially small eigenpairs, is lost when it is restarted. The convergence of GMRES(m) behaves as though the information of the small Ritz values is not retained.

Recently, we succeeded in developing several procedures that improve the convergence of GMRES(m). One of our new techniques employs deflation. De-

flation is when an invariant subspace corresponding to the smallest eigenvalues is approximated, and the influence on the components of the residual vector is eliminated. There are two types of deflation techniques. One of them is Morgan's Deflated GMRES(m, k) [5, 9], which is mathematically equivalent to GMRES-E [6] and GMRES-IR [7]. Morgan's technique accelerates convergence by approximating small eigenvalues.

The technique explored in this article adds an invariant subspace corresponding to the smallest eigenvalues to the Krylov subspace of the next GMRES cycle. An alternative deflation technique is to accelerate convergence by removing small eigenvalues. This is executed by constructing a preconditioner as proposed by Erhel et al. [3], Baglama et al. [1], and Burrage et al. [2]. In this technique, the smallest eigenvalue of the coefficient matrix A in equation (1) replaces the largest modulus of the eigenvalues. In both aforementioned techniques, the information of small eigenpairs is retained when it is restarted, and the superlinear convergence is maintained.

Our proposal is a modification of techniques derived from Deflated GMRES(m, k) [5] and DEFLGMRES(m, l) [3]. Our preconditioning technique was derived from Deflated GMRES(m, k) and the technique for constructing a preconditioner was derived from DEFLGMRES(m, l). We refer to this new method as Deflated GMRES(m, k, l). This new method is compatible with Deflated GMRES(m, k) and the algorithm is very simple and easy to program.

Section 2 overviews Deflated GMRES(m, k) and its pros and cons. Section 3 details our new method, Deflated GMRES(m, k, l). Section 4 tabulates and analyzes the results of numerical experiments for a representative set of matrices. Section 5 concludes.

2 Deflated GMRES(m, k)

GMRES is normally used to solve a large sparse nonsymmetric system of linear equations (1). GMRES is one of the Krylov subspace methods based

on the Arnoldi process. An approximate solution is chosen from the Krylov subspace $K_m(A, r_0)$ to minimize the residual norm

$$x_m \in x_0 + K_m(A, r_0) \quad \text{such that} \quad \|r_m\| \to \min,$$

where

$$K_m(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{m-1} r_0\}.$$

The Arnoldi method creates an orthonormal basis $\{v_1, \ldots, v_m\}$ for the Krylov subspace, and an $(m+1) \times m$ upper Hessenberg matrix $\bar{H}_m$ such that

$$AV_m = V_{m+1} \bar{H}_m \quad \text{where } V_m = [v_1, \ldots, v_m].$$

An approximate solution, where the residual norm is minimized, is computed to solve the following least square problem

$$x_m = x_0 + V_m d_m, \quad \text{and} \quad d_m = \text{argmin}_d \, \|\beta e_1 - \bar{H}_m d\|,$$

where $\beta = \|r_0\|$. The computational cost and memory requirements of this process increases significantly with the number of iterations. The restarted version of GMRES is often used to address these issues and this method is called GMRES(m).

Recently, a number of techniques to improve the convergence of GMRES(m) were developed. In 2002, Morgan [5] developed Deflated GMRES(m, k). This method computes the information about the small eigenpairs at each restart, and carries them forward to the next cycle. The property of superlinear convergence is retained by approximating the small eigenvalues rapidly. To be more precise, k harmonic Ritz vectors $z_1, \ldots, z_k$ [8] are computed at each restart and added to the subspace of the next cycle. This subspace is extended to dimension $m+1$ from the current residual vector r_0 with the Arnoldi method. This subspace is

$$\text{span}\{z_1, \ldots, z_k, r_0, Ar_0, \ldots, A^{m-k-1} r_0\}.$$

Morgan [5] showed that this subspace maintains the Krylov subspace.

3 Deflated GMRES(m, k, l)

This section explores a technique which applies an additional preconditioning stage to Deflated GMRES(m, k). Algorithm 1 lists the proposed algorithm, called Deflated GMRES(m, k, l). In the first stage [2, 3], the preconditioner M^{-1} is applied to the original system of linear equations $Ax = b$, where $M^{-1} = I_n + U(|\lambda_n|^{-1}T^{-1} - I_l)U^T$. In the second stage, Deflated GMRES(m, k) is applied to the equation $AM^{-1}y = b$, $x = M^{-1}y$ to expedite solving the system of linear equations (1). The proposed method provides a good approximation for the smallest eigenvalues rapidly and accelerates the convergence rate of Deflated GMRES. The procedure for constructing the preconditioner is based on a DEFLGMRES(m, l) technique, which was developed by Erhel et al. [3]. This technique replaces the largest module of the eigenvalues with l, the smallest eigenvalue of the system of linear equations (1). l harmonic Ritz vectors corresponding to the smallest harmonic Ritz values instead of l Schur vectors were used, and maximum harmonic Ritz values were used instead of maximum Ritz values. We used harmonic Ritz pairs in both the deflated procedure and the preconditioner. This method is compatible with Deflated GMRES(m, k) and it is easy for programmers to implement this method.

4 Numerical experiments

Numerical experiments illustrate the effectiveness of Deflated GMRES(m, k, l). A CELSIUS J360 computer with a 4 Gbyte main memory was used. Algorithms were implemented in the C programming language with double precision. The initial approximation was set to $x_0 = 0$, and the GMRES solvers were stopped when the relative residual norm was reduced at least by a factor of 10^{-12}, that is $\|r_m\|_2/\|r_0\|_2 < 10^{-12}$.

Algorithm 1: Deflated GMRES(m, k, l).

choose x_0 ;
set $M^{-1} = I_n$, itr $= 0$;
compute $r_0 = b - Ax_0$;
compute $x_m \in x_0 + M^{-1}K_m\left(AM^{-1}, r_0\right)$ s.t. $\|r_m\| \to \min$;
if $\|r_m\| < \epsilon$ **then**
    stop
end
set itr $=$ itr $+ 1$
compute eigenpairs (θ_i, g_i) of $H_m + h_{m+1,m}^2 H_m^{-T} e_m e_m^T$
if itr $= 1$ **then**
    orthonormalize $V_m g_1, \ldots, V_m g_l$ to form $u_1, \ldots, u_l$;
    set $T_l = U_l^T A U_l$;
    set $M^{-1} = I_n + U_l\left(|\theta_m| T_l^{-1} - I_l\right) U_l^T$;
    set $x_0 = x_m$ and goto 3 ;
else
    orthonormalize $V_m g_1, \ldots, V_m g_k$ to form $z_1, \ldots, z_k$;
    set $K_m = \text{span}\{z_1, \ldots, z_k, r_m, Ar_m, \ldots, A^{m-k-1} r_m\}$;
    set $x_0 = x_m$ and goto 4 ;
end

4.1 First test problem

Consider the system of linear equations (1), where the coefficient matrix A and the right-hand vector b are

$$A = \begin{pmatrix} 1 & 1 & & & & \\ -1 & 2 & 1 & & & \\ & -1 & 3 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & n-1 & 1 \\ & & & & -1 & n \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{pmatrix}.$$

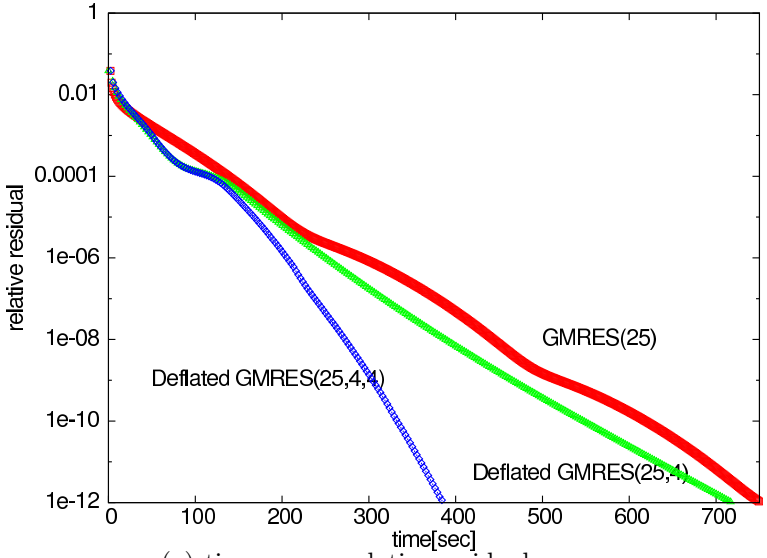
TABLE 1: Results of Deflated GMRES(m, k, l) for Test problem 4.1: the left side values (itr) are the numbers of iterations; the right-hand side values (sec) are the computation times.

k	$l = 0$		$l = 1$		$l = 2$		$l = 3$		$l = 4$	
	itr	sec	itr	sec	itr	sec	itr	sec	itr	sec
0	14800	751	-	-	-	-	-	-	-	-
1	11833	1187	10370	1079	5594	602	3314	355	4946	525
2	9708	1009	8399	921	5340	586	5110	560	5018	566
3	7219	802	6518	739	5264	587	5066	584	5616	655
4	6304	719	5951	701	4985	596	4397	528	3137	386
5	5645	686	5170	619	5070	627	4610	555	4590	557
6	5155	641	4990	634	4724	599	4135	531	4021	520
7	4831	633	4406	582	4586	606	4334	579	3886	505
8	4581	625	4164	570	3977	549	4045	563	3994	546
9	4409	619	3762	532	3586	507	3938	561	3890	556
10	4300	630	4100	598	4025	567	4130	607	3755	550

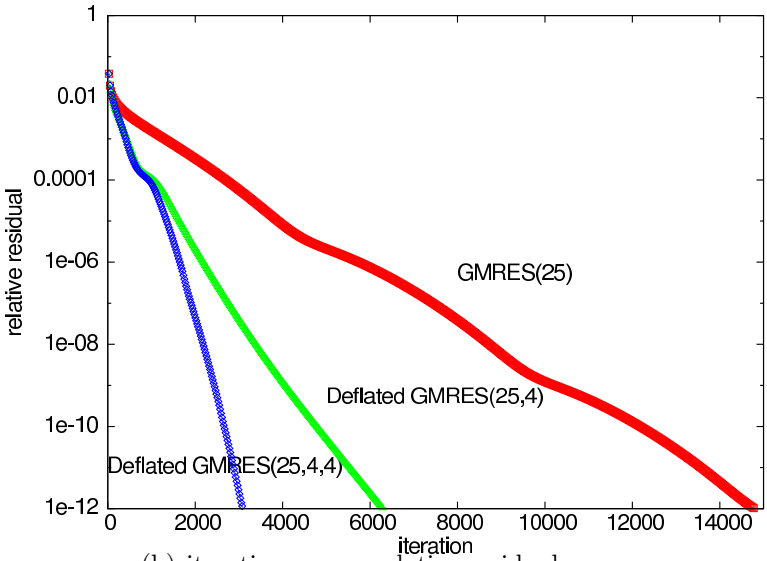
The size of matrix A is $n = 65536$ and the restart frequency is set to $m = 25$. We show the results of numerical experiments for Test problem 4.1 in Table 1. The left-hand side values (itr) are the numbers of iterations and the right-hand side values (sec) are the computation times. Deflated GMRES(m, k, l) is run as Deflated GMRES(m, k) when $l = 0$, and as GMRES(m) when $k = 0$ and $l = 0$. The convergence of our proposed method was significantly faster than that of conventional methods. Markedly better results were obtained when $(k, l) = (1, 3)$, and $(4, 4)$.

A comparison of the convergence of GMRES(25), Deflated GMRES$(25, 4)$ and Deflated GMRES$(25, 4, 4)$ was made. Figure 1 displays the convergence behaviour of the relative residual norm of the above methods.

The distribution of eigenvalues at 1000 iterations is displayed in Figure 2(b). The convergence of our proposed method can accelerate rapidly because the smallest harmonic Ritz values provide good approximations for the smallest

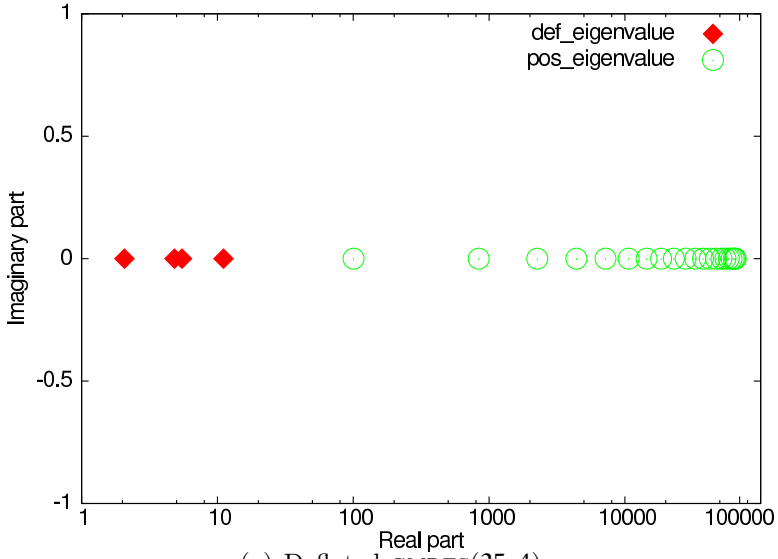


(a) time versus relative residual norm

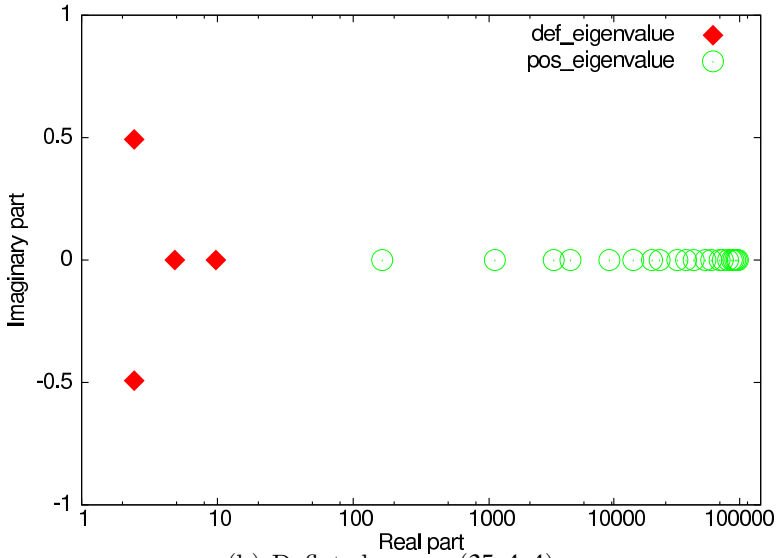


(b) iteration versus relative residual norm

FIGURE 1: The convergence behaviour of the relative residual norm for Test problem 4.1.



(a) Deflated GMRES$(25, 4)$



(b) Deflated GMRES$(25, 4, 4)$

FIGURE 2: The distribution of eigenvalues at 1000 iterations for Test problem 4.1: ◊, deflated eigenvalues; ○, positive eigenvalues.

TABLE 2: Results of Deflated GMRES(m, k, l) for Test problem 4.2: The left-hand side values (itr) are the numbers of iterations; the right-hand side values (sec) are the computation times.

k	$l = 0$		$l = 1$		$l = 2$		$l = 3$		$l = 4$	
	itr	sec	itr	sec	itr	sec	itr	sec	itr	sec
0	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	19374	318	19144	319
5	20705	324	-	-	19000	314	18145	305	18685	318
6	14746	238	11232	188	-	-	-	-	-	-
7	-	-	11796	202	14720	257	-	-	11108	199
8	12272	209	14296	253	-	-	-	-	-	-
9	-	-	-	-	11662	215	17894	334	15393	292
10	21890	395	16940	317	13780	261	12620	242	25420	494

complex eigenvalues of the matrix A.

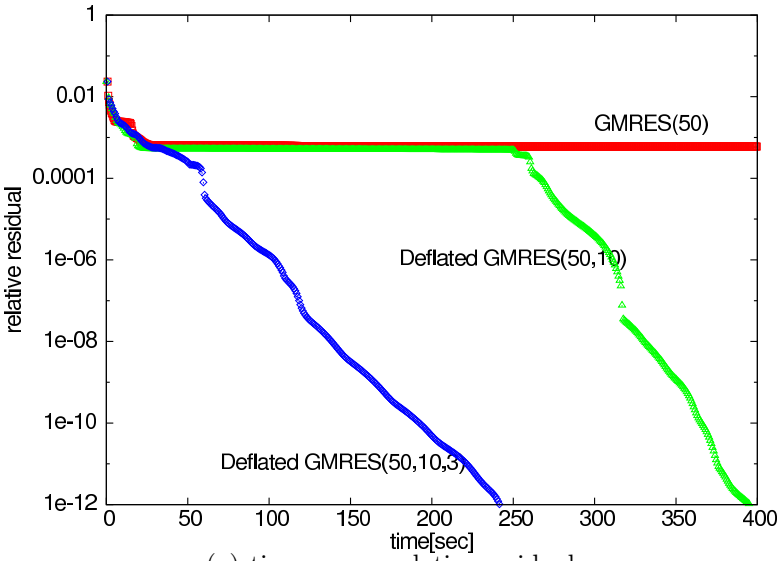
4.2 Second test problem

For our second example we consider a Dirichlet boundary value problem defined by a partial differential equation discussed by Joubert [4]:

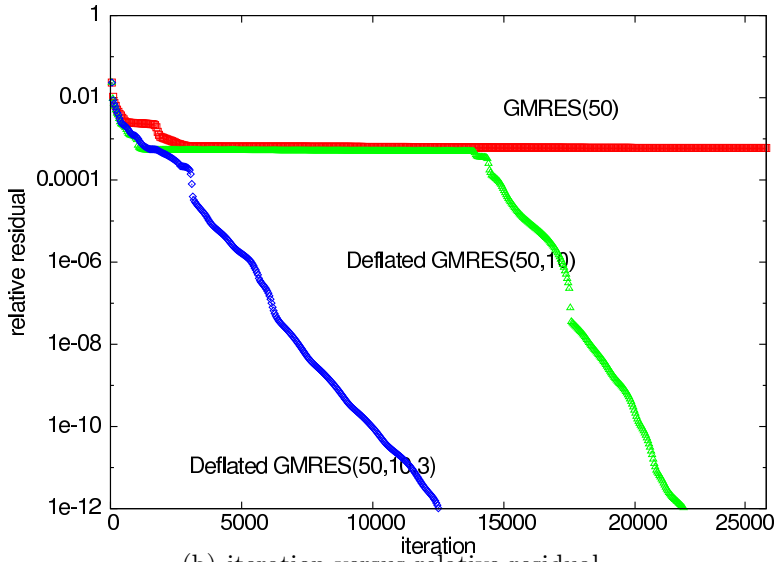
$$-\Delta u + D\left[(y - \tfrac{1}{2})u_x + (x - \tfrac{1}{3})(x - \tfrac{2}{3})u_y\right] - 43\pi^2 u = G \quad \text{on } \Omega = [0,1]^2,$$
$$u(x,y) = 1 + xy \quad \text{on } \partial\Omega\,,$$

where $G(x,y)$ is chosen so that $u(x,y) = 1 + xy$. We discretised the above equation with mesh size $1/(2^7 + 1)$, using a five point, central difference scheme and this resulted in a linear system of equations (1) with a coefficient matrix A of size $n = 16384$. Table 2 shows the results of some numerical experiments for Test problem 4.2.

The convergence of our proposed method was often faster than that of conventional methods, and succeeds in cases that the conventional methods

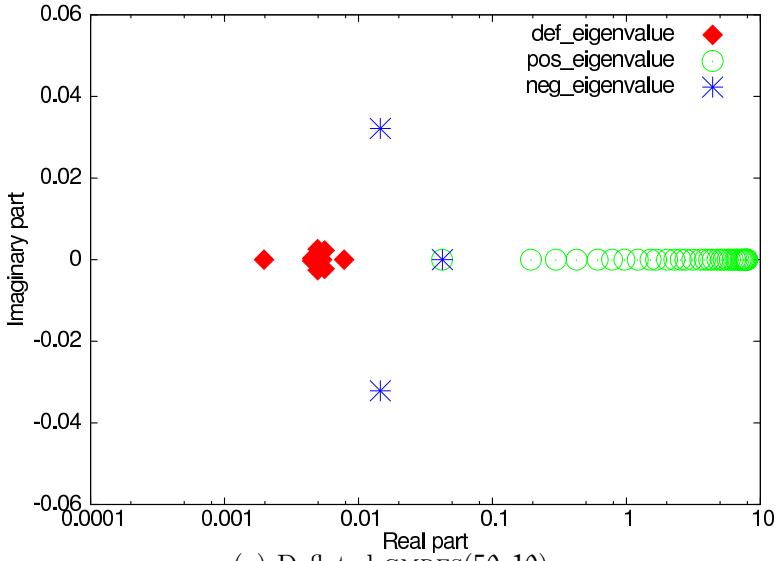


(a) time versus relative residual

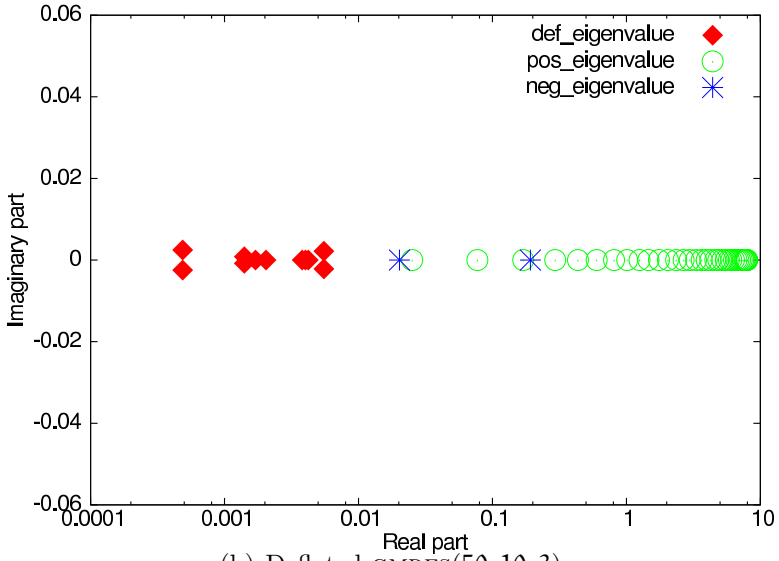


(b) iteration versus relative residual

FIGURE 3: The convergence behaviour of the relative residual norm for Test problem 4.2.



(a) Deflated GMRES(50, 10)



(b) Deflated GMRES(50, 10, 3)

FIGURE 4: The distribution of eigenvalues at 4000 iterations for Test problem 4.2: ◊, deflated eigenvalues; ○, positive eigenvalues; ∗, negative eigenvalue.

were unable to solve, that is, $k = 4, 7$ and 9. However, the convergence of our proposed method was sometimes slower than that of conventional methods, that is, when $k = 8$. Figure 3 compares the convergence behaviour of residual norms for GMRES(50), Deflated GMRES$(50, 10)$ and Deflated GMRES$(50, 10, 3)$.

The distribution of eigenvalues at 4000 iterations is shown in Figure 4. Note that the modulus of negative eigenvalues are plotted in the positive domain. Figure 4(b) shows the convergence of our proposed method can accelerate rapidly, because the smallest harmonic Ritz values approximates the smallest complex eigenvalues in matrix A. On the other hand, GMRES(m) does not converge because it cannot approximate the small complex eigenvalues and the Deflated GMRES(m, k) method converges by approximating the small complex eigenvalues at around 15000 iterations.

5 Conclusions

The new technique we explore in this article is the Deflated GMRES(m, k, l) method, which is a modification of the Deflated GMRES(m, k) method proposed by Morgan [5]. The procedure of constructing the preconditioner is based on a DEFLGMRES(m, l) technique proposed by Erhel et al. [3]. In our proposed method, the approximation is better than conventional methods, because harmonic Ritz vectors are used in the deflating procedure and constructing a preconditioner. Our proposed method is compatible with Deflated GMRES(m, k) and its computation scheme is easily implemented.

Our numerical experiments show the effectiveness of Deflated GMRES(m, k, l). In both Test problems 4.1 and 4.2, the convergence of Deflated GMRES(m, k, l) is usually faster than that of conventional methods. Test problem 4.2 shows our proposed method was able to solve problems that conventional methods could not solve, and it was able to accelerate the convergence of the residual norm, because it quickly formulated a better distribution of eigenvalues through our proposed two stage preconditioning.

References

[1] J. Baglama, D. Calvetti, G. H. Golub and L. Reichel, Adaptively pre-conditioned GMRES algorithms, *SIAM J. Sci. Comput.*, **20** (1998), pp. 243–269. doi:10.1137/S1064827596305258 C224

[2] K. Burrage and J. Erhel, On the performance of various adaptive preconditioned GMRES strategies, *Numer. Linear Algebra Appl.*, **5**(1998), pp. 101–121. doi:10.1002/(SICI)1099-1506(199803/04)5:2¡101::AID-NLA127¿3.0.CO;2-1 C224, C226

[3] J. Erhel, K. Burrage and B. Pohl, Restarted GMRES preconditioned by deflation, *J. Comput. Appl. Math.*, **69** (1996), pp. 303–318. doi:10.1016/0377-0427(95)00047-X C224, C226, C234

[4] W. Joubert. Lanczos methods for the solution of nonsymmetric systems of linear equations, *SIAM J. Matrix. Anal. Appl.*, **13** (1992), pp. 926–943. doi:10.1137/0613056 C231

[5] R. B. Morgan, GMRES with deflated restarting, *SIAM J. Sci. Comput.*, **24** (2002), pp. 20–37. doi:10.1137/S1064827599364659 C224, C225, C234

[6] R. B. Morgan, A restarted GMRES method augmented with eigenvectors, *SIAM J. Matrix Anal. Appl.*, **16** (1995), pp. 1154–1171. doi:10.1137/S0895479893253975 C224

[7] R. B. Morgan, Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations, *SIAM J. Matrix Anal. Appl.*, **21** (2000), pp. 1112–1135. doi:10.1137/S0895479897321362 C224

[8] R. B. Morgan, Harmonic projection methods for large non-symmetric eigenvalue problems, *Numer. Linear Algebra Appl.*, **5**(1998), pp. 33–55. doi:10.1002/(SICI)1099-1506(199801/02)5:1¡33::AID-NLA125¿3.0.CO;2-1 C225

[9] S. Rollin and W. Fichtner, Improving the accuracy of GMRES with deflated restarting, *SIAM J. Sci. Comput.*, **30**(2007), pp. 232–245. doi:10.1137/060656127 C224

[10] Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, **7** (1986), pp. 856–869. doi:10.1137/0907058 C223

[11] H. A. van der Vorst and C. Vuik, The superlinear convergence behavior of GMRES, *J. Comput. Appl. Math.*, **48** (1993), pp. 327–341. doi:10.1016/0377-0427(93)90028-A C223

Author addresses

1. **J. Shiroishi**, School of Fundamental Science and Technology, Graduate School of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama 223-8522, Japan.
mailto:j_shiroishi@math.keio.ac.jp

2. **T. Nodera**, Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama 223-8522, Japan.
mailto:nodera@math.keio.ac.jp