Title	The efficient parallel implementation of the approximate inverse preconditioning for the shifted linear systems : focus on the Sherman-Morrison formula(Mathematical Sciences for Large Scale Numerical Simulations)
Author(s)	Moriya, Kentaro; Zhang, Linjie; Nodera, Takashi
Citation	数理解析研究所講究録 (2007), 1573: 139-147
Issue Date	2007-11
URL	http://hdl.handle.net/2433/81314
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第 1573 巻 2007 年 139-147

The efficient parallel implementation of the approximate inverse preconditioning for the shifted linear systems – focus on the Sherman-Morrison formula –

Kentaro Moriya, Aoyama Gakuin University
Linjie Zhang, Graduated School of Keio University
Takashi Nodera, Keio University

1 Introduction

We study the following linear systems of equations:

$$\begin{aligned} Ax &= b \qquad (1)\\ \tilde{A}_i\tilde{x}_i &= b, \quad \tilde{A}_i = (A + \xi_i I), \qquad i = 1, 2, \ldots, N-1. \qquad (2)\end{aligned}$$

where A, $\tilde{A}_i \in C^{n\times n}$ be nonsingular and nonhermitian matrices, and let $\xi_i \in C^n$ be such that the shifted matrices $\tilde{A}_i$ is nonsingular. Namely, the linear system (1) is called by seed system. The coefficient matrices of linear systems (1) and (2) have only different entries on their main diagonal.

In this paper, we propose a new technique that applies AISM (Approximate Inverse with the Sherman-Morrison formula) method to these linear systems of equations. Using the proposed technique, we also compare the performance of the preconditioned GMRES(m) algorithm with the Shifted-GMRES(m) algorithm. At last, numerical experiments are given.

2 Shifted-GMRES(m) algorithm

We define the following two Krylov subspaces

$$\begin{aligned} K_m(A,\ r_0) &= \mathrm{span}\{r_0,\ Ar_0,\ \ldots,\ A^{m-1}r_0\}\\ \tilde{K}_m(\tilde{A}_i,\ \tilde{r}_0) &= \mathrm{span}\{\tilde{r}_0,\ \tilde{A}_i\tilde{r}_0,\ \ldots,\ \tilde{A}_i^{m-1}\tilde{r}_0\}.\end{aligned}$$

If $r_0 = \beta_0\tilde{r}_0$, then $K_m(A,\ r_0) = \tilde{K}(\tilde{A}_i,\ \tilde{r}_0)$ is satisfied.
[Proof] As for $(\tilde{A}_i)^k\tilde{r}_0 \in \tilde{K}_m(\tilde{A}_i,\ \tilde{r}_0)$, where $k = 0,\ 1,\ \ldots,\ m-1$

$$(\tilde{A}_i)^k\tilde{r}_0 = \beta_0(A + \xi_i I)^k r_0 = \sum_{j=0}^{k}\beta_0\{{}_kC_j\xi_i^{(k-j)}\}(A)^j r_0 \in K_m(A,\ r_0) \qquad \square$$

Therefore, the approximate solutions of all the shifted linear systems can be solved by using only one Krylov subspace. However, if we use the preconditioner of the coefficient matrix A, $K(AM^{-1}, r_0)$ is not equivalent to $\tilde{K}(\tilde{A}M^{-1}\tilde{r_0})$ and the equality between two Krylov subspaces is no more satisfied. The disadvantage of this iterative solver is that it is not easy to apply the preconditioner to these linear systems of equations.

3 Some Preconditioners

In this section, we describe the brief introduction of MR and AISM preconditioner.

1: **for** $k = 1$ **to** n **do**
2:   **select** $m_k^{(0)}$
3:   **for** $j = 0$ **to** IMAX **do**
4:     $r_k^{(j)} = e_k - Am_k^{(j)}$
5:     $\tilde{r}_k^{(j)} = e_k - \tilde{A}\tilde{m}_k^{(j)}$
6:     $\alpha = (r_k^{(j)},\ Ar_k^{(j)}) \ / \ (Ar_k^{(j)},\ Ar_k^{(j)})$
7:     $\tilde{\alpha} = (\tilde{r}_k^{(j)},\ \tilde{A}\tilde{r}_k^{(j)}) \ / \ (\tilde{A}\tilde{r}_k^{(j)},\ \tilde{A}\tilde{r}_k^{(j)})$
8:     $m_k^{(j)} = m_k^{(j)} + \alpha r_k^{(j)}$
9:     $\tilde{m}_k^{(j)} = \tilde{m}_k^{(j)} + \tilde{\alpha}\tilde{r}_k^{(j)}$
10:   **endfor**
11: **endfor**

Figure 1. MR method

3.1 MR Method

The preconditioner M^{-1} is computed by the following recurrences

$$\begin{aligned} r_k^{(j)} &= e_k - Am_k^{(j)} \\ m_k^{(j)} &= m_k^{(j)} + \alpha r_k^{(j)}, \end{aligned}$$

where $m_k^{(j)}$ is the k-th column vector of M^{-1} in the j-th step of MR iteration. The scalar α is determined so that the residual norm $\|r_k^{(j)}\|_2$ is minimized. It is usually set as

$$\alpha = (r_k^{(j)},\ Ar_k^{(j)}) \ / \ (Ar_k^{(j)},\ Ar_k^{(j)}).$$

We present the MR method in Figure 1. The notation "IMAX" means the iterations of MR method. While the line number 4, 6 and 8 present the computation of preconditioner of the linear system (1), the line number 5, 7 and 9 present the computation of preconditioner of the linear systems (2). As the number of the shifted linear systems (2) is more increased, the computation of this preconditioner becomes more expensive. Therefore, it is not so appropriate to apply this preconditioner to the shifted linear systems.

3.2 AISM method

We define $p_k = e_k$ and $q_k = (a_k - se_k)^T$, where a_k and e_k are the k-th column vector of A, and the identity vector, respectively. Using the following three recurrence formula

$$\begin{aligned} u_k &= p_k - \sum_{i=1}^{k-1} \frac{(v_i)_k}{sr_i} u_i, \\ v_k &= q_k - \sum_{i=1}^{k-1} \frac{(q_k, u_i)}{sr_i} v_i, \end{aligned}$$

and

$$r_k = 1 + (v_k)_k / s.$$

```
1:  for k = 1 to n do
2:    p_k = e_k
3:    q_k = a^k - s e_k
4:    u_k = p_k
5:    v_k = q_k
6:    for i = 1 to k - 1 do
7:      u_k = u_k - {(v_i)_k / (s r_i)} u_i
8:      v_k = v_k - {(q_k, u_i) / (s r_i)} v_i
9:    endfor
10:   for i = 1 to n do
11:     if |(u_k)_i| < tolU   set (u_k)_i = 0
12:     if |(v_k)_i| < tolV   set (v_k)_i = 0
13:   endfor
14:   r_k = 1 + (v_k)_k / s
15: endfor
```

Figure 2. The AISM method

The AISM preconditioner is described as follows.

$$M^{-1} = sI - A^{-1} = s^{-2} U \Omega^{-1} V^T \qquad (3)$$

where

$$\begin{aligned} U &= \{u_1,\ u_2,\ \ldots, u_n\}, \\ V &= \{v_1,\ v_2,\ \ldots, v_n\}, \end{aligned}$$

and

$$\Omega = \operatorname{diag}\{r_1,\ r_2,\ \ldots, r_n\}.$$

In Figure 2, we present the AISM method. The computation of u_k and v_k, $(k = 1, 2, \ldots, n)$ in line number 5 and 6 can be parallelized partially based on Moriya et al. [5]. Therefore, AISM method is parallelized in the numerical example. Just like in MR method, the dropping off process is used in the statement of line number 9 and 10. If the k-th entries of u_k and v_k are less than the thresholds tolU and tolV, respectively. About more detail of the AISM preconditioner, see Bru et al.[4].

4 The technique applying AISM method to the shifted linear systems

While the preconditioner of seed system (1) is given in the equation (3), the preconditioner of shifted linear systems (2) is described as

$$\begin{aligned} \tilde{M}^{-1} &= s^{-1} I - \tilde{A}^{-1} \\ &= (s^{-1} - \xi^{-1}) I - A^{-1} \\ &= \tilde{s}^{-1} I - A^{-1}. \end{aligned}$$

Therefore, if s and $\tilde{s}$ are the same values, the same preconditioner can be used for the linear systems (1) and (2). We propose the technique that applies only one common preconditioner to all the linear systems. In the proposed technique, we set $s = \tilde{s}$ and select the appropriate values for both of preconditioners of linear systems (1) and (2).

According to Bru et al. [4], it is known that the preconditioner M^{-1} performs well, when $s > \rho(A)$ is satisfied in system (1), where $\rho(A)$ is the spectral radius of A. Then all the eigenvalues near zero point can be moved to the left side of complex plain, and the convergence of the residual norm is improved. Based on the theorem in Bru et al. [4], the conditions

$$s > \rho(A), \quad s > \rho(\tilde{A}_i), \qquad \text{for } i = 1,\ 2,\ \ldots, N-1 \tag{4}$$

are satisfied, the AISM method is expected to compute an effective preconditioner for all the shifted linear systems. One of the appropriate selections that achieve $s > \rho(A)$ is

$$s = 1.5\ \|A\|_\infty, \tag{5}$$

and just like the same reason, if

$$s = 1.5\ \|\tilde{A}_i\|_\infty, \qquad \text{for} \quad i = 1,\ 2,\ \ldots,\ N-1. \tag{6}$$

is set, $s > \rho(\tilde{A}_i)$ is also satisfied. However, it is impossible to satisfy both conditions (5) and (6). Instead of this two conditions, we propose the selection of s so that

$$s > 1.5\ \|A\|_\infty \tag{7}$$

and

$$s > 1.5\ \|\tilde{A}_i\|_\infty, \qquad \text{for } i = 1,\ 2,\ \ldots,\ N-1 \tag{8}$$

are satisfied. If conditions (7) and (8) are satisfied, conditions (4) are also satisfied. We select

$$s = 1.5(\|A\|_\infty + \max_i |\xi_i|) \tag{9}$$

as the appropriate scholar s for all the shifted linear systems. If the equation (9) is selected, both conditions (7) and (8) are satisfied.

[Proof]

$$\begin{aligned} s &= 1.5(\|A\|_\infty + \max_i |\xi_i|) \\ &> 1.5\ \|A\|_\infty > \rho(A) \end{aligned}$$

and

$$\begin{aligned} s &= 1.5(\|A\|_\infty + \max_i |\xi_i|) \\ &= 1.5(\max_i\ \|A\|_\infty + \max_i |\xi_i|) \\ &> 1.5 \max_i \{\|A + \xi_i I\|_\infty\} \\ &= 1.5 \max_i \{\|\tilde{A}_i\|_\infty\} > 1.5\{\|\tilde{A}_i\|_\infty\} \geq \rho(\tilde{A}_i) \qquad \square \end{aligned}$$

Therefore, if the equation (9) is employed as the diagonal shifted value s, we can obtain one common appropriate preconditioner for all the shifted linear systems of equations.

5 Numerical results

In this section, we present results of the two numerical experiments. Our computations were done in the following PC cluster system with 8 CPUs.

cluster Node: IBM Xseries346 (× 4)

CPU: Pentium4 3.6GHz (2 per one node)

OS: Fedora Core 4 Linux

Local memory: 1GB per one node

Communication library: MPI[7]

The main experiments are measuring the speedup ratio of the AISM preconditioner and comparing the AISM preconditioned GMRES(m) algorithm with the Shifted-GMRES(m) algorithm. The preconditioning parameters are as follows.

MR method

- **Dropping off tolerance:** tol=0.1, 0.01
- **Iterations:** IMAX = 1, 2

AISM method

- **Dropping off tolerance:** tolU = 0.1, 0.01
- **Dropping off tolerance:** tolV = 0.1, 0.01
- **Diagonal shifted value:** $s = 1.5(\|A\|_\infty + \max_i |\xi_i|)$

[**Example 1**] In the square region $\Omega = [0,\ 1]^2$, we now consider the boundary value problem of PDE

$$-[\{\exp(-xy)\}u_x]_x - [\{\exp(xy)\}u_y]_y + 10.0(u_x + u_y) - 60.0u = f(x, y)$$
$$u(x,\ y)|_{\partial\Omega} = 1 + xy$$

We discretize this problem by using five points differential scheme with 192^2 grid points to obtain the coefficient matrix of order 36,864. We study the eigenvalue problem of the coefficient matrix based on the Figure 3. We choose the central point $c = (0.15,\ 0)$ and the radius $R = 0.14$. The number of shifted linear systems N is 8. The right hand side b is determined so that all of its entries are 1.0. The shifted linear systems in line 3 of this figure are solved by the preconditioned GMRES(m) algorithm and the Shifted-GMRES(m) algorithm to compare these iterative solvers. We start the iterations with the initial approximation of zero vector. Table 1 presents the computation time and iterations needed for satisfying the stopping criterion

$$\|r_i\|_2 \ /\ \|b\|_2 < 1.0 \times 10^{-12} \tag{10}$$

about all the residual norms, where $\|r_i\|_2$ is the i-th residual norm of GMRES iterations. The value in bracket "()" means the number of the residual norms that can not converge within one hour. Only in the case of using AISM method, the residual norms of all the linear systems can

```
1:  select c, R, N, m and vectors b, d
2:  set ω_j = c + R exp(2πi j / N), j = 0, 1, ..., N − 1
3:  solve (A − ω_j I) x_j = b, j = 0, 1, ..., N − 1
4:  set f(ω_j) = d^H x_j, j = 0, 1, ..., N − 1
5:  compute μ̂_j = (1/N) Σ_{k=1}^{N−1} (ω_k − c)^{j+1} f(ω_k), j = 0, 1, ..., 2m − 1
6:  compute eigenvalues θ_0, ..., θ_m of Ĥ_m − λ H̄_m
7:  compute λ_j = θ_j + c
```

Figure 3. The algorithm to solve the eigenvalue problem with using the shifted linear systems

Algorithm	preconditioner	Restart cycle					
		20		30		40	
	time	time	iter	time	iter	time	iter
shifted-GMRES(m)	0.0	(5)	-	(5)	-	(5)	-
GMRES(m)+AISM(tolU, tolV = 0.1)	89.0	135.0	74	144.0	73	156.0	73
GMRES(m)+AISM(tolU, tolV = 0.01)	92.0	138.0	72	151.0	71	160.0	71
GMRES(m)+MR(tol = 0.1, imax = 1)	897.0	(6)	-	(6)	-	(6)	-
GMRES(m)+MR(tol = 0.1, imax = 2)	2399.0	(4)	-	(4)	-	(4)	-
GMRES(m)+MR(tol = 0.01, imax = 1)	901.0	(6)	-	(6)	-	(6)	-
GMRES(m)+MR(tol = 0.01, imax = 2)	2396.0	(4)	-	(4)	-	(4)	-

(-): The number of the converged residual norms when one hour has passed.

Table 1. Example 1: Computation time and iterations of shifted linear systems (time: computation time (s), iter: iterations)

converge. Some of the residual norms can not converge in cases of MR method and the Shifted-GMRES(m) algorithm. The computation time of MR preconditioner is much more expensive than AISM preconditioner, and its cost is not practical. On the other hand, with using AISM preconditioner, the iterations are terminated at most three minutes. Therefore, we find that it is effective to apply one common preconditioner to all the linear systems.

Figure 4 presents the number of the converged residual norms as for the computation time. In the AISM method, all of the residual norms converge almost simultaneously. In case of the Shifted-GMRES(m) algorithm, the convergence of the 5-th residual norm is about 1,000 seconds slower than the last converged residual norm. Also, it takes about 1,000 seconds for the first residual norm to converge. In MR preconditioner, six residual norms converge almost the same time. However, the run time cost is about 1,000 seconds to converge, and the last two residual norms do not converge.

We measure the parallel performance of AISM method. In Figure 5, up to 4 PEs, the speedup ratio is almost linear, and it is decreased in the case of using 8 PEs, and the speedup ratio is about 4.5 times.

[Example 2] We consider the matrix, named "ECL32", in the Florida Sparse Matrix Collection [6]. The order and non zeros of the matrix are 51,993, 347,097, respectively. The right hand side is determined so that all the entries are 1.0. Just like in Example 1, the number of shifted linear systems N are 8, and we solve the shifted linear systems described in the line 3 of Figure

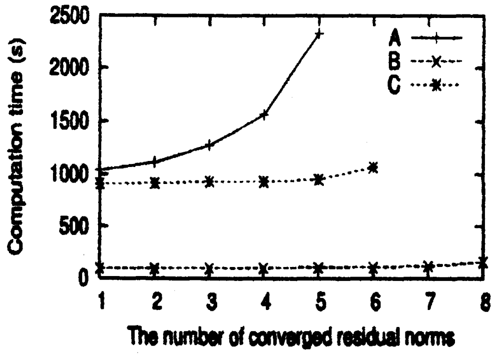

Figure 4. Example 1: The relation of the number of converged residual norms and computation time (A: Shifted-GMRES(40), B: MR+GMRES(40), tol=0.1, IMAX=1, C: AISM+GMRES(40), tolU, tolV=0.1)

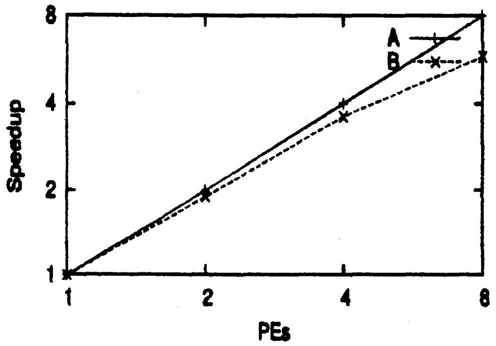

Figure 5. Example 1: Performance analysis of AISM method, (A: ideal, B: AISM method, tolU, tolV=0.1)

3. In this example, the central point of $c = (1.0,\ 0)$ and the radius of $R = 0.99$ are selected. Table 2 presents the computation time and iterations needed for stopping criterion (10). According to this table, AISM method enable all the residual norms to converge about four or five times faster than MR method. Also the preconditioning cost of AISM method is not so expensive as MR method. Even if the iterations of MR method "IMAX" is increased, the computation cost can not be reduced, and rather expensive. The cost of MR method is more than 10 times as expensive as AISM method.

In the Shifted-GMRES(m) algorithm, only the last residual norm can not converge. Therefore, we analyze the relation between the converged residual norms and computation time. From Figure 6, the shifted-GMRES(m) algorithm enables seven residual norms to converge much faster than the other preconditioned GMRES(m) algorithm. However, the last one can not converge. The Shifted-GMRES(m) algorithm is expensive for not all the linear systems, and the convergence is rather quick than AISM method. Only one residual norm does not converge. On the other hand, the preconditioned GMRES(m) algorithm enables all the residual norm to converge.

Algorithm	precon-ditioner	Restart cycle					
		20		30		40	
	time	time	iter	time	iter	time	iter
Shifted-GMRES(m)	0.0	(7)	-	(7)	-	(7)	-
GMRES(m)+AISM(tolU, tolV = 0.1)	181.0	446.0	311	479.0	272	551.0	276
GMRES(m)+AISM(tolU, tolV = 0.01)	280.0	517.0	209	559.0	202	639.0	198
GMRES(m)+MR(tol = 0.1, imax = 1)	1141.0	2078.0	558	1790.0	476	1695.0	446
GMRES(m)+MR(tol = 0.1, imax = 2)	2572.0	2903.0	295	2870.0	270	2944.0	267
GMRES(m)+MR(tol = 0.01, imax = 1)	1144.0	2080.0	558	1793.0	476	1696.0	446
GMRES(m)+MR(tol = 0.01, imax = 2)	2571.0	2892.0	295	2880.0	270	2932.0	267

(-): The number of the converged residual norms when one hour has passed.

Table 2. Example 2: Computation time and iterations of shifted linear systems (time: computation time (s), iter: iterations)

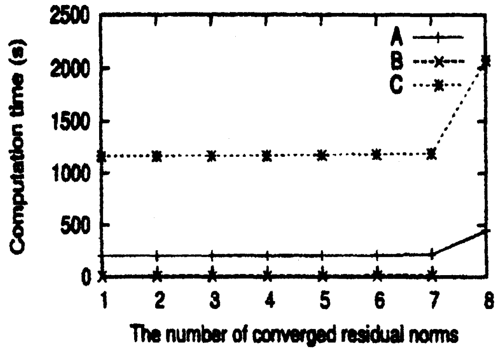

Figure 6. Example 2: The relation of the number of Converged residual norms and computation time (A: Shifted-GMRES(20), B: MR+GMRES(20), tol=0.1, IMAX=1, C: AISM+GMRES(20), tolU, tolV=0.1)

Figure 7 shows the speedup ratio of AISM method. In this experiment, the parallel performance is not so effective as Example 1, since the sparse structure of the matrix is more irregular. In case of 8 PEs, the speedup of about 4 times is obtained.

6 Concluding remarks

We have proposed a new technique of AISM method for applying the shifted linear systems. In the original scheme, either the Shifted-GMRES(m) algorithm without preconditioning or the preconditioned GMRES(m) algorithm with expensive computation cost, like MR method, is usually used. On the other hand, the proposed technique can compute one common preconditioner of all the systems. it does not depend on the number of linear systems. From two numerical examples, it is effective to apply the AISM preconditioner to the shifted linear systems with using the proposed technique. We can also obtain the speedup ratio of about 4 times by using 8 PEs. Therefore, this technique can be effective for applying the shifted linear systems.

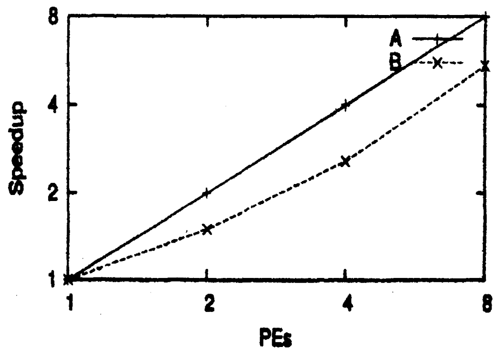

Figure 7. Example 2: Performance analysis of AISM method, (A: ideal, B: AISM, tolU, tolV=0.1)

In the future work, we plan to study the detailed numerical performance of our algorithm to allocating each of shifted systems (2) to seed system (1).

References

[1] GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems, *SIAM J. Sci. Stat. Comput.*, No. 7, pp. 856–869 (1986).

[2] Frommer, A, and Glassner, U.: Restarted GMRES for Shifted Linear Systems, *SIAM J. Sci. Comput.*, Vol. 19, pp. 15–26 (1998).

[3] Huckel. T.: Approximate Sparsity Patterns for the Inverse of a Matrix and Preconditioning, *Appl. Numer. Math.*, No. 30, pp. 291–303 (1999).

[4] Bru, R., Credán, J., and Marín, J., Mas, J.: Preconditioning Sparse Nonsymmetric Linear systems with the Sherman-Morrison Formula, *SIAM J. Sci. Comput.*, Vol. 25, No. 2, pp. 701–715 (2003).

[5] K. Moriya, L. Zhang, and T. Nodera: The computation of the approximate inverse by parallelizing the Sherman-Morrison formula (in Japanese), *J. of IPSJ*, Vol. 48, No. 3 (2007) to appear.

[6] University of Florida Sparse Matrix Collection. [Online] http://www.cise.ufl.edu/research/sparse/matrices.

[7] LAM/MPI Parallel Computing. [Online] http://www.lam-mpi.org.