

Computing an Approximate Inverse Preconditioner by Successive Approximations Method

Amer Kaabi

Department of Basic Science
Khoramshahr Marine Science and Technology University
Khoramshahr, Iran
kaabi_amer@yahoo.com

Abstract

A method for computing an approximate inverse preconditioner of a matrix A which all eigenvalues have negative real parts is proposed. The approximate solution of the special Sylvester matrix equation $AX + XA = I$, which is an approximate inverse preconditioner of matrix A , is obtained by the successive approximations method. Some numerical experiments on test matrices from Harwell-Boing collection for comparing the numerical performance of the new method with an available well-known algorithm are presented.

Keywords: Successive approximations method, Sylvester equation, Approximate inverse preconditioner

1 Introduction.

In this paper, we consider the solution of large and sparse systems of linear equations

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

by means of iterative techniques. It is widely known that the preconditioning techniques have an important role in improving the performance of the Krylov subspace methods. Of various preconditioners, sparse approximate inverse preconditioners have recently received much attention since they are much suited for parallel implementation. There are some kinds of sparse approximate inverse preconditioners such as Frobenius norm minimization [7], factorized sparse approximate inverse [2], and rank-one update methods [3]. The main purpose of this paper is to obtain an approximate inverse preconditioner P of A such that $PA \approx I$ by computing an approximate solution of special Sylvester equation

$$AX + XA = I \tag{2}$$

where $A \in R^{n \times n}$ is a given matrix, $I \in R^{n \times n}$ is identity matrix and $X \in R^{n \times n}$ is unknown solution matrix. we shall henceforth that all eigenvalues of matrix A have negative real parts (that is A is stable) and large and sparse. Then the unique solution X of equation (1) is given by [1, 6]

$$X = - \int_0^{\infty} e^{At} I e^{At} dt = - \int_0^{\infty} (e^{At})^2 dt \quad (3)$$

and iterative methods are more appropriate for determining X . In this paper, we use method of successive approximations to obtain the approximate solution of equation (2). We know that the Sylvester equation(2) has a unique solution for X if and only if the matrices A and $-A$ have no eigenvalues in common,i.e.

$$\lambda_i(A) - \lambda_j(-A) \neq 0 \quad (4)$$

where $\lambda_i(A)$, $\lambda_i(-A)$, $i = 1, \dots, n$ are the eigenvalues of A and $-A$ respectively , see [5]. Therefore, we can obtain an approximate inverse preconditioner for the matrix A which has no symmetry eigenvalues.

This paper is organized as follows.In section 2 we explain the method of successive approximations for solving the Sylvester matrix equations. The parallel implementation of this method is demonstrated in section 3. In section 4, we give AINV preconditioner. In section 5 some numerical examples are tested. The conclusion remarks is given in section 6.

2 Method of Successive Approximations.

Let the matrices A , X and I in equation (2) be partitioned into the form

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1\mu} \\ A_{21} & A_{22} & \cdots & A_{2\mu} \\ \vdots & \vdots & & \vdots \\ A_{\mu 1} & A_{\mu 2} & \cdots & A_{\mu\mu} \end{bmatrix} \quad (5)$$

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1\mu} \\ X_{21} & X_{22} & \cdots & X_{2\mu} \\ \vdots & \vdots & & \vdots \\ X_{\mu 1} & X_{\mu 2} & \cdots & X_{\mu\mu} \end{bmatrix}, I = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1\mu} \\ I_{21} & I_{22} & \cdots & I_{2\mu} \\ \vdots & \vdots & & \vdots \\ I_{\mu 1} & I_{\mu 2} & \cdots & I_{\mu\mu} \end{bmatrix} \quad (6)$$

where the blocks $A_{i,j}$, $X_{i,j}$ and $I_{i,j}$ are square matrices of order l : thus $n = \mu l$. We consider the following splitting of matrix A :

$$A = M + N. \quad (7)$$

We assume that the symmetric matrix $M + M^T$ is negative definite. Thus, M is stable [4, 8]. Let $X^{(k)}$ be an approximation to the unique solution X of

equation (1), $k = 0, 1, 2, \dots$. The method of successive approximations can be written as

$$MX^{(k+1)} + X^{(k+1)}M = I - NX^{(k)} - X^{(k)}N \tag{8}$$

where $X^{(0)}$ is an arbitrary initial matrix approximation to the exact solution. In order to study the convergence of this method, we note that equation (2) can also be written in the form

$$MX + XM = I - NX - XN. \tag{9}$$

Since the matrix M is stable, the solution of (2) can also be written in the form

$$X = - \int_0^\infty e^{Mt}(I - NX - XN)e^{Mt}dt \tag{10}$$

Since the matrix M is stable, equation (8) has a unique solution given by the formula

$$X^{(k+1)} = - \int_0^\infty e^{Mt}(I - NX^{(k)} - X^{(k)}N)e^{Mt}dt \tag{11}$$

Then from (10) and (11) we obtain

$$E^{(k+1)} = \int_0^\infty e^{Mt}(NE^{(k)} + E^{(k)}N)e^{Mt}dt \tag{12}$$

where $E^{(k)} = X - X^{(k)}$ is the error associated with the k -th iterate of (8). Thus, about the convergence of the method (8), the following theorem holds.

Theorem 1. Let A and M are stable matrices. Thus the method (8) converges to the solution X of (1) for all $X^{(0)}$ if

$$\frac{\|N\|_2}{|\mu(M)|} < 1, \tag{13}$$

where,

$$\mu(M) = \frac{1}{2}\lambda_{max}(M + M^T).$$

Proof. Since the matrix A is stable, the unique solution of the equation (2) can be written as follows (see [5])

$$X = - \int_0^\infty (e^{At})^2 dt.$$

Since the matrix M is also stable, the unique solution of the equation (9) is given by the following formula

$$X^{(k+1)} = - \int_0^\infty e^{Mt}(I - NX^{(k)} - X^{(k)}N)e^{Mt}dt$$

So, we have

$$E^{(k+1)} = \int_0^\infty e^{Mt}(NE^{(k)} + E^{(k)}N)e^{Mt}dt \tag{14}$$

where $E^{(k)} = X - X^{(k)}$ is the error associated with the k -th iterate of (8). Let u and v denote the left and the right singular vectors of unit length of $E^{(k+1)}$ such that (see [5])

$$u^T E^{(k+1)} v = \| E^{(k+1)} \|_2 . \tag{15}$$

Combining (14) and (15), we obtain

$$\begin{aligned} \| E^{(k+1)} \|_2 &= | u^T \int_0^\infty e^{Mt}(I - NX^{(k)} - X^{(k)}N)e^{Mt}dv | \\ &= | \int_0^\infty (e^{Mt}u)^T(I - NX^{(k)} - X^{(k)}N)(e^{Mt}v)dt | . \end{aligned}$$

By using Cauchy-Schwartz inequality, we have

$$\| E^{(k+1)} \|_2 \leq \| NE^{(k)} + E^{(k)}N \|_2 \left(\int_0^\infty \| e^{M^T t} u \|_2 \| e^{Mt} v \|_2 dt \right).$$

Using again Cauchy-Schwartz inequality, we obtain

$$\int_0^\infty \| e^{M^T t} u \|_2 \| e^{Mt} v \|_2 dt \leq \left(\int_0^\infty \| e^{M^T t} u \|_2^2 dt \right)^{\frac{1}{2}} \left(\int_0^\infty \| e^{Mt} v \|_2^2 dt \right)^{\frac{1}{2}}$$

By noting that $\| e^{M^T t} u \|_2 \leq \| e^{M^T t} \|_2$ and $\| e^{Mt} v \|_2 \leq \| e^{Mt} \|_2$, we have

$$\| E^{(k+1)} \|_2 \leq \| E^{(k)} \|_2 (\| N \|_2 + \| N \|_2) \left(\int_0^\infty \| e^{M^T t} u \|_2^2 dt \right)^{\frac{1}{2}} \left(\int_0^\infty \| e^{Mt} v \|_2^2 dt \right)^{\frac{1}{2}} \tag{16}$$

Since, by hypothesis, the symmetric matrix $M + M^T$ is negative definite the following bounds

$$\begin{aligned} \| e^{M^T t} \|_2 &\leq e^{\mu(M^T)t} \\ \| e^{Mt} \|_2 &\leq e^{\mu(M)t} \end{aligned}$$

can be used to estimate

$$\int_0^\infty \| e^{M^T t} \|_2^2 dt$$

and

$$\int_0^\infty \| e^{Mt} \|_2^2 dt$$

respectively. Thus

$$\int_0^\infty \| e^{M^T t} \|_2^2 dt \leq \frac{1}{2 | \mu(M) |}$$

and

$$\int_0^\infty \| e^{Mt} \|_2^2 dt \leq \frac{1}{2 |\mu(M)|}.$$

We conclude from (16) that

$$\| E^{(k+1)} \|_2 \leq \frac{\| N \|_2}{|\mu(M)|} \| E^{(k)} \|_2 \tag{17}$$

Thus the method (8) is convergent to the solution X of (2) for all $X^{(0)}$ if

$$\frac{\| N \|_2}{|\mu(M)|} < 1.$$

□

3 Parallel Implementation.

When we implement the iterative method (8) on parallel computer a suitable choice of M in the splitting (7) of A is

$$M = \text{diag}\{A_{11}, A_{22}, \dots, A_{\mu\mu}\} \tag{18}$$

In this case, at each iteration k of the method (8) we must solved a set of Sylvester equations of the form

$$A_{ij}X_{ij} + X_{ij}A_{ij} = I_{ij} - (NX^k + X^kN)_{ij} \tag{19}$$

for $i, j = 1, \dots, \mu$, where $(NX^k + X^kN)_{ij}$ is the block (i, j) of the matrix $(NX^k + X^kN)_{ij}$ partitioned commensurately with the block structure of I. Generally, equation (19) is small, dense Sylvester equation that can be solved directly for X_{ij} using direct method [1].

When μ is even, another interesting splitting of the matrices A is

$$M = \text{diag}\left\{ \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}, \begin{bmatrix} A_{3,3} & A_{3,4} \\ A_{4,3} & A_{4,4} \end{bmatrix}, \dots, \begin{bmatrix} A_{\mu-1,\mu-1} & A_{\mu-1,\mu} \\ A_{\mu,\mu-1} & A_{\mu,\mu} \end{bmatrix} \right\} \tag{20}$$

it reduce the method (8) to solving a set of Sylvester equations of the form

$$\begin{aligned} & \begin{bmatrix} A_{l,l} & A_{l,l+1} \\ A_{l+1,l} & A_{l+1,l+1} \end{bmatrix} \begin{bmatrix} X_{l,s} & X_{l,s+1} \\ X_{l+1,s} & X_{l+1,s+1} \end{bmatrix} + \\ & \begin{bmatrix} X_{l,s} & X_{l,s+1} \\ X_{l+1,s} & X_{l+1,s+1} \end{bmatrix} \begin{bmatrix} A_{s,s} & A_{s,s+1} \\ A_{s+1,s} & A_{s+1,s+1} \end{bmatrix} = \\ & \begin{bmatrix} F_{l,s} & F_{l,s+1} \\ F_{l+1,s} & F_{l+1,s+1} \end{bmatrix} \tag{21} \end{aligned}$$

$l, s = 1, 3, \dots, \mu - 1$ and F_{ij} is the block (i, j) of the matrix $I - (NX^k + X^kN)$. These equations can be solved simultaneously by a direct method on

a parallel computer. An analogous set of independent Sylvester equations can be obtained if we choose

$$M = \text{diag}\{A_{11}, \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix}, \dots, \begin{bmatrix} A_{\mu-2\mu-2} & A_{\mu-2\mu-1} \\ A_{\mu-1\mu-2} & A_{\mu-1\mu-1} \end{bmatrix}, A_{\mu\mu}\} \quad (22)$$

With splitting (18), (20) and (21), method (8) has a high inherent parallelism. Now, by computing an approximate solution of the special Sylvester equation (2), we can produce a new inverse preconditioner for a linear system with coefficient matrix A having eigenvalues with negative (or positive) real parts. Thus, we can propose the following algorithm.

Algorithm 3.1. Successive approximations method for computing an approximate inverse preconditioner SAINV.

1. Choose an integer number μ , a tolerance ε and matrix A of order n .
2. Set $l = n/\mu$. Partition the matrices A , X and I into the form (5) and (6).
3. Split matrix A of the form (7). Split matrix M of the form (18).
4. By using the successive approximations method, obtain an approximate solution X_{app} of the equation (19), such that $\|I - AX_{\text{app}} - X_{\text{app}}A\|_F < \varepsilon$.
5. Set $P = X_{\text{app}}$. Use a dropping strategy for the columns of P .
6. EndDo.

Note that, in the step 3 of above algorithm, if we split matrix M of the form (20), then we must solve the equation (21) in step 4.

4 AINV Preconditioner.

The construction of the AINV preconditioner is based on an algorithm which computes two sets of vectors $\{z_i\}_{i=1}^n$ and $\{w_i\}_{i=1}^n$, which are A -biconjugate, i.e., such that $w_i^T A z_j = 0$ if and only if $i \neq j$. Given a nonsingular matrix $A \in R^{n \times n}$, there is a close relationship between the problem of inverting A and computing two sets of A -biconjugate vectors $\{z_i\}_{i=1}^n$ and $\{w_i\}_{i=1}^n$. If we introduce the matrices

$$Z = [z_1, z_2, \dots, z_n] \text{ and } W = [w_1, w_2, \dots, w_n]$$

then

$$W^T A Z = \text{diag}\{p_1, p_2, \dots, p_n\},$$

where

$$p_i = w_i^T A z_i \neq 0.$$

It follows that W and Z are necessarily nonsingular and

$$A^{-1} = Z D^{-1} W^T = \sum_{i=1}^n \frac{z_i w_i^T}{p_i}.$$

Hence, the inverse of A is known if two complete sets of A -biconjugate vectors are known. Assuming that A has an LU factorization, matrices W and

Z whose columns are A -biconjugate can be explicitly computed by means of a biconjugation process applied to the standard basis vectors e_1, \dots, e_n . It is easy to see that $Z = U^{-1}$ and $W = L^{-T}$ where $A = LDU$ with L and U unit lower and upper triangular and D diagonal. If a_i^T denotes the i th row of A , the biconjugation procedure to compute Z can be written as follows.

Algorithm 4.1 Left-looking AINV algorithm

1. Let $z_1^{(0)} = e_1$, $p_1^{(0)} = a_{11}$.
2. For $i=2, \dots, n$ Do:
3. $z_i^{(0)} = e_i$
4. For $j=i, \dots, i-1$ Do:
5. $p_j^{(j-1)} = a_j^T z_i^{(j-1)}$
6. $z_i^{(j)} = z_i^{(j-1)} - \left(\frac{p_i^{(j-1)}}{p_j^{(j-1)}}\right) z_j^{(j-1)}$
7. EndDo
8. $p_i^{(i-1)} = a_i^T z_i^{(i-1)}$
9. EndDo

The computation of W is identical, except that a_i^T is replaced by c_i^T , where c_i is the i th column of A . Note that Z and W are computed incompletely, by removing elements less than prescribed drop tolerance. This leads to incomplete factors $\bar{Z} \approx Z$, $\bar{W} \approx W$ and $\bar{D} \approx D$, and the factorized approximate inverse takes the form $M = \bar{Z}\bar{D}^{-1}\bar{W}^T$. For more details see [2]. There exist several other approximate inverse techniques, but in this paper we limit ourselves to a comparison between AINV and the new preconditioner that is the approximate solution of special Sylvester matrix equations which computed by successive approximate method.

5 Numerical Experiments.

The numerical experiments were performed on a set of sparse matrices drawn from the Harwell-Boeing collection. These matrices originate from a variety of applications such as oil reservoir simulation, circuit design, semiconductor device modeling, etc. For all examples, we used the stopping criterion

$$\|b - Ax\|_2 < 10^{-12}$$

and the maximum number of iterations allowed set to 2500. The right hand side of $Ax = b$ is taken such that the exact solution is $x = [1, 1, \dots, 1]^T$. The solution of above sparse linear systems, by restarted generalized minimum residual method (GMRES(m)) [9], is considered.

Table. Test problems (nnz(A)=nonzeros in matrix, and results for GMRES(m), GMRES(m) with AINV and SAINV (we choose $\varepsilon = 0.1$ in Algorithm 3.1) preconditioners). Timings are in seconds.

matrix	n	$nmz(A)$	l	m	T1	Iter1	T2	Iter2	T3	Iter3
bcsstk01	48	224	2	4	1.09	2500	1.01	808	1.02	36
nos1	237	627	3	2	-	†	6.15	2455	5.06	1180
bus494	494	1080	2	10	-	†	195	1748	214	1421
nos6	675	1965	5	10	76.5	2500	65.5	441	62.5	142
bus685	685	1967	5	4	-	†	-	†	142	1321
nos7	729	2673	9	20	128	1943	196	2455	107	1272
nos2	957	2547	3	8	145	2500	143	1857	141	15
bus1138	1138	4054	2	8	-	†	657	2035	577	1528

In above Table we present the results of test runs with GMRES(m) and preconditioned GMRES(m) with the approximate inverse preconditioners AINV and SAINV. For these three methods we give the number of iterations for convergence by Iter1, Iter2 and Iter3 respectively. Also, we refer the set-up times for the GMRES(m) iterations, the preconditioned GMRES(m) with AINV and SAINV iterations by T1, T2 and T3 respectively. A † means failure to attain convergence within 2500 iterations.

6 Conclusions.

We have proposed a technique (which refer by SAINV)for constructing a preconditioner for general linear systems with coefficient matrix A which has the eigenvalues with negative (or positive) real parts. Also, in this paper we have presented the results of experiments, with two implementation of the sparse approximate invers preconditioners AINV and SAINV. Based on our experiments, we conclude that these techniques are comparable from the point of view of robustness and rate of convergence, with SAINV being somewhat better on average. Finally, it was found that the computation of the preconditioner is much more expensive for SAINV than for AINV, but it was observed that the situation could be different in a parallel implementation. We conclude that both techniques offer excellent potential for use on high-performance computers.

References

- [1] R. H. Bartels and G. W. Stewart, Algorithm 432: Solution of matrix equation $AX+XB=C$, Circ. Syst and Signal Prod., 13(1994) pp. 820-828.
- [2] M. Benzi, M. Tuma, A sparse approximate inverse preconditioner for nonsymmetric linear systems, SIAM J. Sci. Comput., 19(1998), 968-994.
- [3] R. Bru, J. Cerdan, J. Marin, and J. Mas, preconditioning sparse nonsymmetric linear systems with sherman-morrison formula, SIAM J. Sci.Comput., 25(2003), 701-715.

- [4] J. N. Franklin, *Matrix Theory*, Prentice-Hall Inc., Englewood Cliffs, 1968.
- [5] E. Galligani, Solution of the Sylvester equations on Parallel Computer, *Parallel Computing* (1996)189-195.
- [6] G. H. Golub and G. F. Van Loan, *Matrix computations*, 2nd ed., Johns Hopkins U.P., Baltimore, 1989.
- [7] M. J. Grote and T. Huckle, Parallel preconditioning with sparse approximate inverses, *SIAM J. Sci. Stat. Comput.*, 18(1997), 838-853.
- [8] J. Ortega, *Matrix Theory, A second course*, Plenum Press, New York, 1987.
- [9] Y. Saad and M. H. Schultz, GMRES, A generalized minimal residual algorithm for nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, 7, pp.856-869, 1986.

Received: February, 2009