ADI iterative solution of Lyapunov equations

Article *in* Applied Mathematics Letters · December 1988

DOI: 10.1016/0893-9659(88)90183-8

1 author:

E.L. Wachspress

105 PUBLICATIONS **1,768** CITATIONS

SEE PROFILE

Appl. Math. Lett. Vol. 1, No. 1, pp. 87-90, 1988
Printed in the U.S.A.

0893-9659/88 $3.00 + 0.00

Iterative Solution of the Lyapunov Matrix Equation

EUGENE L. WACHSPRESS
University of Tennessee at Knoxville
April, 1987

Abstract. Iterative solution of the Lyapunov matrix equation $AX + XB = C$ using ADI theory described in [1] is reviewed here. A procedure for implementing this technique when A and B are sparse is introduced in this paper.

1. Background.

A recurring problem in linear algebra is to find an $m \times n$ matrix X which satisfies the Lyapunov matrix equation

$$AX + XB = C \tag{1}$$

for specified real matrices $A(m \times m)$, $B(n \times n)$ and $C(m \times n)$. When A and B have only eigenvalues with positive real part (N-stable as defined by Young in [7]), R. A. Smith's iteration [5] converges quadratically to the unique solution. It was shown in [1] that Smith's algorithm is equivalent to Alternating-Direction-Implicit (ADI) iteration with a single parameter and that one may define an alternative procedure with variable parameters. This more general ADI iteration will be explored in greater detail here for the case where A and B are sparse. References to other direct and iterative methods for solving (1) are given in [2], [3]. The direct method of Golub, Nash, and vanLoan [2] is currently the method of choice when $m = n$. However, for large systems iterative methods should prevail. Direct methods require computation of $0(n^4)$ whereas iterative methods of $0(n^3)$ are known.

2. Relevance of ADI Iteration Theory.

The ADI model problem which arises in solving PDEs of elliptic type is to find the vector u which satisfies $(H + V)u = s$ when given the real n-vector s and the real symmetric and *commuting* $n \times n$ matrices H and V whose sum is positive definite. If one replaces these matrices by the linear operators $H(X) = AX$ and $V(X) = XB$ with A and B real and symmetric and $A + B$ SPD, then the operators H and V *commute for any A and B*. Thus, Eq. 1 is a model ADI problem! The theory for ADI iteration is definitive here [6]. One readily determines parameters which solve the minimax problem

$$h(t) = \underset{\{p,q\}}{\text{minimum}} \; \underset{\substack{a \le x \le b \\ c \le y \le d \\ a+c>0}}{\text{maximum}} = \prod_{r=1}^{t} \left| \frac{(x - q_r)(y - p_r)}{(x + p_r)(y + q_r)} \right| . \tag{2}$$

Analytic bounds on error reduction are known for the general case. Characteristic behavior is illustrated for the case where x and y have the same range with $b/a = \sigma$ representative of the conditioning of the problem.

A bound on error reduction

$$e(t) = \frac{\|X^{(t)} - X\|}{\|X\|} \tag{3}$$

after t iterations is given approximately by

$$e(t) < h(t) = \exp\left[-\frac{\pi^2 t.}{\ln(4\sigma)} \right] . \tag{4}$$

The ADI iteration equations for solution of (1) with $x^{(0)}$ chosen as 0 are:

$$(A + p_r E)x^{(r-\frac{1}{2})} = C - X^{(r-1)}(B - p_r I) \tag{5.1}$$

$$X^{(r)}(B + q_r I) = C - (A - q_r E)X^{(r-\frac{1}{2})} \qquad r = 1, 2, \ldots, t \tag{5.2}$$

with E and I the identity matrices of order n and m, respectively.

Numerical studies reported in [1, 4] verified the analysis. Theoretical convergence rates were achieved for matrices with real spectra. When A and B have complex spectra, existing theory is less definitive and current research [4] addresses this problem. Applicability of the real theory when the imaginary components are small compared to real components has been demonstrated. Realistic error bounds have been determined for some complex spectral regions. Optimum parameters are chosen from estimated spectra. Efficient procedures for computing spectral bounds are an important component of ADI iteration programs. This question will not be addressed here.

3. Review of Smith's Algorithm.

Smith recognized that if one defines the matrices

$$K = (p + q)(pE + A)^{-1}C(qI + B)^{-1} \tag{6.1}$$

$$G = (pE + A)^{-1}(qE - A) \tag{6.2}$$

$$H = (qI + B)^{-1}(pI - B) \tag{6.3}$$

that X satisfies the equation

$$X - GXH = K \tag{7}$$

and the iteration

$$\begin{aligned} X^{(0)} &= K \\ X^{(r)} &= X^{(r-1)} + GX^{(r)}H, \qquad r = 1, 2, \ldots \end{aligned} \tag{7}$$

converges to X.

(Smith actually considered only the equations with $p = q$. The more general iteration considered here is more efficient when A and B have significantly different spectra.)

Smith then observed that this linearly convergent sequence could be converted to a quadratically convergent sequence by squaring G and H step:

$$G(r) = G(r-1)^2, \qquad H(r) = H(r-1)^2, \tag{9.1}$$

$$X^{(r)} = X^{(r-1)} + G(r)X^{(r-1)}H(r). \tag{9.2}$$

It is easily shown that $X^{(r)}$ from (9) is equal to the matrix obtained after $2^r - 1$ steps with (8). The error reduction $e(r)$ is bounded by the product of the spectral norms of $G(r)$ and $H(r)$. Optimum values for p and q minimize this bound and are chosen according to ADI theory for repeated use of one p, q doublet. The bound on $e(r)$ is then

$$e(r) < E(r) = \left[\left(\sqrt{\sigma} - 1\right) / \left(\sqrt{\sigma} + 1\right)\right]^{2^{r+1}} \tag{10}$$

where σ is as defined for Eq. 4.

The work per iteration for the Smith and ADI methods is comparable when A and B are not sparse. For example, when $m = n$ the major arithmetic for Smith is four multiplications of matrices of order n and for ADI is the solution of $2n$ linear systems of order n. The work per iteration in general is of order N^3 where N is the larger of n and m. A comparison of ADI error reduction of Eq. 4 with the Smith error reduction of Eq.

10 reveals that ADI can compete only when error greater than about .001 is acceptable. Smith's algorithm compares even more favorably with complex spectra. For example, when the eigenvalues are bounded by a circle in the right half plane with center on the real axis the optimum ADI parameters are all equal to the single value used in the Smith algorithm.

4. ADI Iteration When A and B Are Sparse.

If A and B are sparse, the relative efficiency of the ADI method improves. The matrices G and H in Smith's algorithm are generally full so little benefit accrues from sparseness of A and B. On the other hand, ADI iteration applied directly to the Lyapunov equation makes use of the sparseness and is well suited for parallel and vector computation. In solving elliptic PDEs with ADI one iterates over a 2D grid, solving tridiagonal systems alternately by rows and columns. In solving the Lyapunov equation with ADI one may update alternately all rows and columns of the approximation to matrix X. Element ij of matrix C is analogous to the component of a source term at grid point ij of a discrete elliptic system. One possible implementation is the following:

Columns of matrix U are u_i and of U^T are u_j. In alternating direction iteration, one often generates a matrix from its columns or rows and then uses its rows or columns in the next step.

The initial "half-sweep" iteration is performed with right-hand side

$$s_i^{(1/2)} = c_i$$

The iteration requires solution of n linear systems, which can be done in parallel or with vectors of length n:

$$(A + p_1 E)x_i^{(1/2)} = s_i^{(1/2)}$$

One cycles through the following four steps for $r = 1, 2, \ldots, t-1$ and then the first three steps to compute $x_j^{(t)}$ which yields $X^{(t)}$:

(1) The right-hand side for "full-sweep" r is

$$s_j^{(r)} = c_j + (p_r + q_r)x_j^{(r-\frac{1}{2})} - s_j^{(r-\frac{1}{2})} \qquad \text{for } j = 1, 2, \ldots, m.$$

(2) The m linear systems may be solved in parallel or with vectors of length m for the full-sweep:

$$(B^T + q_r I)x_j^{(r)} = s_j^{(r)} \qquad \text{for } j = 1, 2, \ldots, m.$$

(3) The right-hand side for the next half-sweep is

$$s_i^{(r+\frac{1}{2})} = c_i + (p_{r+1} + q_r)x_i^{(r)} - s_i^{(r)} \qquad \text{for } i = 1, 2, \ldots, n.$$

(4) The half-sweep iteration with vectors of length n or parallel computation with n processors is

$$(A + p_{r+1}E)x_i^{(r+\frac{1}{2})} = s_i^{(r+\frac{1}{2})}.$$

One now returns to step 1 with r incremented.

It is seen that the bulk of the arithmetic is solution of m sparse linear systems of order n and n sparse linear systems of order m each iteration. The computation time for this algorithm is comparable to that used for solving the model ADI Dirichlet problem over a grid of $n \times m$ points. One might even be able to adapt a Poisson solver to this problem.

Use of the optimum cycle of parameters rather than a fixed doublet is most beneficial for real spectra, in which case the bounds in Eqs. 4 and 10 apply. If sparseness reduces the

work per iteration of the ADI method to half that of Smith's iteration, then ADI is more efficient when a relative error greater than around 10^{-8} is acceptable. For given spectral bounds Jordan's theory in [**6**] yields an effective σ for use in the error bounds of Eqs. 4 and 10. For a given work ratio one may then ascertain in advance relative merits of the Smith and the ADI iterations.

5. References.

1. Ellner, N. S. and Wachspress, E. L., *New ADI Model Problem Applications*, Proc. FJCC, Dallas, Texas (1968), 528–34.
2. Golub, G. H., Nash, S. and vanLoan, C., *A Hessenberg-Shur method for solution of the problem* $AX + XB = C$, IEEE Trans. Automat. Control **AC24** (1979), 909–913.
3. Rothschild, D. and Jameson A., *Comparison of four numerical algorithms for solving the Lyapunov matrix equation*, Int. J. Control **11, No. 2** (1970), 181–198.
4. Saltzman, N., *ADI Parameters for Some Complex Spectra*, Univ. of Tennessee Master's thesis (1987).
5. Smith, R. A., *Matrix equation* $XA + BX = C$, SIAM J. Appl. Math. **16, No. 1** (1968), 198–201.
6. Wachspress, E. L., "Iterative Solution of Elliptic Systems," Prentice Hall, 1966.
7. Young, D. M., "Iterative Solution of Large Elliptic Systems," Academic Press, 1971.

Mathematics Department, 121 Ayres Hall, University of Tennessee, Knoxville, TN 37996-1300, USA