In *Proceedings of the 16-th International Symposium on Mathematical Theory of Networks and Systems*, Leuven, Belgium, July 5-9, 2004.

Low Rank Iterative Methods for Projected Generalized Lyapunov Equations

Tatjana Stykel*

Abstract

We generalize an alternating direction implicit method and a cyclic Smith method for projected generalized Lyapunov equations. Low rank versions of these methods are also presented that can be used to compute low rank approximations to the solutions of generalized Lyapunov equations with symmetric, positive semidefinite right-hand side. Numerical example is given.

1 Introduction

Consider a projected generalized continuous-time algebraic Lyapunov equation (GCALE)

$$EXA^T + AXE^T = -P_l BB^T P_l^T, \qquad X = P_r X P_r^T, \tag{1.1}$$

where E, $A \in \mathbb{R}^{n,n}$ and $B \in \mathbb{R}^{n,m}$ are given matrices, P_l and P_r are the spectral projections onto the left and right deflating subspaces of a regular pencil $\lambda E - A$ corresponding to the finite eigenvalues and $X \in \mathbb{R}^{n,n}$ is an unknown matrix. Such an equation arises in stability analysis of differential-algebraic equations, inertia theory for matrix pencils and control problems for descriptor systems [12, 13].

For $E = I$, equation (1.1) is the standard Lyapunov equation. Numerical methods for such an equation has been the topic of numerous publications, see [1, 3, 5, 7, 10, 14].

Assume that a pencil $\lambda E - A$ is *stable*, i.e., it is regular and all the finite eigenvalues of $\lambda E - A$ have negative real part. Then the projected GCALE (1.1) has a unique symmetric, positive semidefinite solution X, see [12] for details. This solution can be computed by the generalized Schur-Bartels-Stewart method or the generalized Schur-Hammarling method [12] that are based on the preliminary reduction of the pencil $\lambda E - A$ to the generalized Schur form, solution of the generalized Sylvester and Lyapunov equations and back transformation. These methods cost $O(n^3)$ operations and they have the memory complexity $O(n^2)$. Thus, the generalized Schur-Bartels-Stewart method and the generalized Schur-Hammarling method can be used only for small or medium size problems. Moreover, they do not take into account the sparsity and/or structure of the matrix coefficients.

Due to the practical importance of the numerical solution of large-scale projected generalized Lyapunov equations that occur in balanced truncation model reduction for descriptor systems [12, 13], development of iterative methods for such equations is a challenge problem.

*Institut für Mathematik, MA 4-5, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany, e-mail: `stykel@math.tu-berlin.de`. This work was supported by the DFG Research Center "Mathematics for key technologies" in Berlin.

In this paper we generalize an alternating direction implicit method [7, 10, 14] and a cyclic Smith method [1, 10] for the projected GCALE (1.1) with large sparse matrix coefficients. Low rank versions of these methods are also presented that can be used to compute a low rank approximation to the solution of (1.1) with a low rank right-hand side.

A main problem in numerical solution of the projected GCALE (1.1) is that we need to compute the spectral projections P_l and P_r onto the left and right deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite eigenvalues. Note that the matrix coefficients E and A arising in applications, such as control of fluid flow, electrical circuits simulation and constrained multibody systems, have some special block structure. This structure can be used to construct the projections P_l and P_r in the explicit form, see [4, 8, 13].

2 Generalized alternating direction implicit method

An *alternating direction implicit* (ADI) *method* was originally proposed for linear systems [9] and then used in [7, 10, 14] to solve standard continuous-time Lyapunov equations. The case of nonsingular E has been considered in [6]. Here we present a generalization of the ADI method for the projected GCALE (1.1).

For some parameter τ, the first equation in (1.1) can be rewritten as

$$(E+\tau A)XA^T + AX(E-\tau A)^T = -P_lBB^TP_l^T. \tag{2.2}$$

Then the generalized ADI iteration for the projected GCALE (1.1) is given by

$$\begin{aligned} (E+\tau_k A)X_{k-\frac{1}{2}}A^T &= -P_lBB^TP_l^T - AX_{k-1}(E-\tau_k A)^T, \\ (E+\overline{\tau}_k A)X_k^TA^T &= -P_lBB^TP_l^T - AX_{k-\frac{1}{2}}^T(E-\overline{\tau}_k A)^T \end{aligned} \tag{2.3}$$

with an initial matrix $X_0 = 0$ and the shift parameters $\tau_1, \ldots, \tau_k \in \mathbb{C}^-$. If the pencil $\lambda E - A$ is stable, then the matrices A, $E+\tau_k A$ and $E+\overline{\tau}_k A$ are nonsingular. In this case we find from (2.3) that

$$\begin{aligned} X_k = {} & (E+\tau_k A)^{-1}(E-\overline{\tau}_k A)X_{k-1}(E-\tau_k A)^T(E+\overline{\tau}_k A)^{-T} \\ & -2\mathrm{Re}(\tau_k)(E+\tau_k A)^{-1}P_lBB^TP_l^T(E+\overline{\tau}_k A)^{-T}. \end{aligned} \tag{2.4}$$

Moreover, it follows from $P_l(E+\tau_k A) = (E+\tau_k A)P_r$ that $X_k = P_rX_kP_r^T$, i.e., the second equation in (1.1) is satisfied exactly.

Let X be a solution of the projected GCALE (1.1). Then the error matrices $X - X_k$ can be computed from (2.2) and (2.4) recursively as $X - X_k = \mathcal{A}_k X \mathcal{A}_k^*$, where

$$\mathcal{A}_k = (E+\tau_k A)^{-1}(E-\overline{\tau}_k A)\cdot\ldots\cdot(E+\tau_1 A)^{-1}(E-\overline{\tau}_1 A)P_r \tag{2.5}$$

and $\mathcal{A}_k^*$ is the complex conjugate and transpose of $\mathcal{A}_k$. If all the finite eigenvalues of the pencil $\lambda E - A$ lie in the open left half-plane, then X_k converges to the solution of the projected GCALE (1.1). The rate of convergence is determined by the spectral radius of the matrix $\mathcal{A}_k$. The minimization of this spectral radius with respect to the parameters $\tau_1, \ldots, \tau_k$ leads to the generalized ADI minimax problem

$$\{\tau_1, \ldots, \tau_k\} = \operatorname*{arg\,min}_{\{\tau_1,\ldots,\tau_k\}\in\mathbb{C}^-} \max_{t\in\mathrm{Sp}_f(E,A)} \frac{|(1-\overline{\tau}_1 t)\cdot\ldots\cdot(1-\overline{\tau}_k t)|}{|(1+\tau_1 t)\cdot\ldots\cdot(1+\tau_k t)|}, \tag{2.6}$$

where $\mathrm{Sp}_f(E, A)$ denotes the finite spectrum of the pencil $\lambda E - A$. To compute the suboptimal ADI shift parameters for the standard problem a heuristic algorithm has been proposed in [10]. Using this algorithm the suboptimal shift parameters $\tau_1, \ldots, \tau_k$ for the generalized problem (2.6) can be determined from a set of largest and smallest approximate finite eigenvalues of $\lambda E - A$ computed by an Arnoldi process.

Recently, an efficient modification of the ADI method has been proposed to compute low rank approximations to the solution of standard Lyapunov equations with large-scale matrix coefficients [7, 10]. This is the *low rank alternating direction implicit* (LR-ADI) *method*. It was observed that the eigenvalues of the symmetric solutions of Lyapunov equations with low rank right-hand side generally decay very rapidly, see [2, 11], and, hence, such solutions may be well approximated by low rank matrices.

A low rank version of the generalized ADI iteration (2.3) is derived analogously to the standard case [7, 10]. First of all note that the matrix X_k in (2.4) is symmetric, positive semidefinite and the Cholesky factor Z_k of $X_k = Z_k Z_k^*$ has the form

$$\begin{aligned} Z_k &= [\sqrt{-2\mathrm{Re}(\tau_k)}(E + \tau_k A)^{-1} P_l B,\ \ (E + \tau_k A)^{-1}(E - \overline{\tau}_k A) Z_{k-1}\,] \\ &= [\sqrt{-2\mathrm{Re}(\tau_k)} S_k P_l B,\ \ \sqrt{-2\mathrm{Re}(\tau_{k-1})} S_k R_k S_{k-1} P_l B,\ \ldots,\ \sqrt{-2\mathrm{Re}(\tau_1)} S_k R_k \cdots S_1 P_l B\,], \end{aligned}$$

where $S_k = (E + \tau_k A)^{-1}$ and $R_k = E - \overline{\tau}_k A$. Taking into account that

$$S_k A S_j = S_j A S_k, \qquad R_k A^{-1} R_j = R_j A^{-1} R_k, \qquad S_k R_j = A^{-1} R_j S_k A$$

for all integer k and j, the matrix Z_k can be rewritten as

$$Z_k = [\, Z_0, F_{k-1} Z_0,\ \ F_{k-2} F_{k-1} Z_0,\ \ \ldots,\ \ F_1 F_2 \cdots F_{k-1} Z_0\,],$$

where $Z_0 = \sqrt{-2\mathrm{Re}(\tau_k)}(E + \tau_k A)^{-1} P_l B = \sqrt{-2\mathrm{Re}(\tau_k)} P_r (E + \tau_k A)^{-1} B$ and

$$F_j = \sqrt{\frac{\mathrm{Re}(\tau_j)}{\mathrm{Re}(\tau_{j+1})}} S_j R_{j+1} = \sqrt{\frac{\mathrm{Re}(\tau_j)}{\mathrm{Re}(\tau_{j+1})}} \left(I - (\tau_j + \overline{\tau}_{j+1})(E + \tau_j A)^{-1} A\right).$$

If we reenumerate the shift parameters in reverse order, then we obtain the following generalization of the LR-ADI method for the projected GCALE (1.1).

Algorithm 1 *The generalized LR-ADI method for the projected GCALE* (1.1).

INPUT*: E, A, $P_l \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, shift parameters $\tau_1, \ldots, \tau_{k_{\max}} \in \mathbb{C}^-$.*
OUTPUT*: A low rank Cholesky factor Z_k of the solution $X \approx Z_k Z_k^*$ of* (1.1).

1. $Z^{(1)} = \sqrt{-2\mathrm{Re}(\tau_1)}\,(E + \tau_1 A)^{-1} P_l B, \quad Z_1 = Z^{(1)}$;
2. FOR $\ k = 2, 3, \ldots$

$$Z^{(k)} = \sqrt{\frac{\mathrm{Re}(\tau_k)}{\mathrm{Re}(\tau_{k-1})}} \left(I - (\overline{\tau}_{k-1} + \tau_k)(E + \tau_k A)^{-1} A\right) Z^{(k-1)}, \quad Z_k = [\, Z_{k-1},\ \ Z^{(k)}\,].$$

END FOR

As a stopping criterion we can use the condition $\|Z^{(k)}\|_F / \|Z_k\|_F \le tol$, where $\|\cdot\|_F$ denotes the Frobenius matrix norm and *tol* is a user-defined tolerance. The iteration can also be stopped as soon as the normalized residual norm

$$\eta(Z_k) = \frac{\|E Z_k Z_k^* A^T + A Z_k Z_k^* E^T + P_l B B^T P_l^T\|_F}{\|P_l B B^T P_l^T\|_F} \tag{2.7}$$

satisfies the condition $\eta(Z_k) \le tol$ or a stagnation of normalized residual norms is observed, see [10] for an efficient computation of the Frobenius norm of the residuals.

It should be noted that the matrices $(E+\tau_k A)^{-1}$ do not have to be computed explicitly. Instead, we solve linear systems $(E+\tau_k A)P_r x = P_l b$, $x = P_r x$ either by computing the (sparse) LU factorizations and forward/backward substitutions or by using iterative Krylov subspace methods. In the latter case the generalized LR-ADI method has the memory complexity $O(k_{ADI}mn)$ and costs $O(k_{ADI}k_{\mathrm{ls}}mn)$ flops, where k_{ADI} is the number of ADI iterations and k_{ls} is the number of linear solver iterations. This method becomes efficient for large-scale sparse Lyapunov equations only if $k_{ADI}k_{\mathrm{ls}}m$ is much smaller than n.

3 Generalized cyclic Smith method

For any parameter $\tau \in \mathbb{C}^-$, the projected GCALE (1.1) is equivalent to the projected discrete-time algebraic Lyapunov equation (DALE)

$$\mathcal{A}X\mathcal{A}^* - X = -P_r\,\mathcal{B}\mathcal{B}^*P_r^T, \qquad X = P_rXP_r^T, \tag{3.8}$$

where $\mathcal{A} = (E+\tau A)^{-1}(E-\overline{\tau}A)$ and $\mathcal{B} = \sqrt{-2\mathrm{Re}(\tau)}\,(E+\tau A)^{-1}B$. It can be shown that if the pencil $\lambda E - A$ is stable, then P_r is the spectral projection onto the invariant subspace of the matrix $\mathcal{A}$ corresponding to the eigenvalues inside the unit circle. In this case the *Smith iteration*

$$X_0 = P_r\,\mathcal{B}\mathcal{B}^*P_r^T, \qquad X_k = P_r\,\mathcal{B}\mathcal{B}^*P_r^T + \mathcal{A}X_{k-1}\mathcal{A}^* \tag{3.9}$$

converges linearly to the solution X of (3.8). Note that the Smith iteration (3.9) is, in fact, the generalized ADI iteration with a single parameter.

A modification of the Smith method has been proposed in [10] to compute a low rank Cholesky factor of the solution of standard Lyapunov equations with a low rank right-hand side. This version of the Smith method is based on the LR-ADI iteration with ℓ shift parameters applied in a cyclic manner and referred to as the *low rank cyclic Smith* (LR-Smith(ℓ)) *method*. This method can be generalized for the projected GCALE (1.1) as follows: first one computes the ℓ-th iterate Z_ℓ of the generalized LR-ADI method with the shift parameters $\tau_1,\ldots,\tau_\ell$ and then solve the DALE $\mathcal{A}_\ell X\mathcal{A}_\ell^* - X = -Z_\ell Z_\ell^*$, where $\mathcal{A}_\ell$ is as in (2.5) with $k=\ell$. We have the following algorithm to compute the low rank Cholesky factor of the solution of the projected GCALE (1.1).

Algorithm 2 *The generalized LR-Smith(ℓ) method for the projected GCALE* (1.1).

INPUT*: E, A, $P_l \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, shift parameters $\tau_1,\ldots,\tau_\ell \in \mathbb{C}^-$.*
OUTPUT*: A low rank Cholesky factor $Z_{k\ell}$ of the solution $X \approx Z_{k\ell}Z_{k\ell}^*$ of* (1.1).

1. *Compute Z_ℓ using Algorithm* 1*;*
2. FOR $k = 2,3,\ldots$

   a. $$Z^{(k\ell)} = \mathcal{A}_\ell Z^{((k-1)\ell)} \quad \textit{with} \quad \mathcal{A}_\ell = \prod_{k=1}^{\ell}(E+\tau_k A)^{-1}(E-\overline{\tau}_k A), \tag{3.10}$$

   b. $$Z_{k\ell} = [\,Z_{(k-1)\ell},\;\; Z^{(k\ell)}\,]. \tag{3.11}$$

   END FOR

It follows from (2.4), (3.10) and (3.11) that the generalized LR-Smith(ℓ) method is equivalent to the generalized LR-ADI iterate with the cyclically repeated shift parameters $\tau_1,\ldots,\tau_\ell$.

Remark. Note that at every iteration step in Algorithms 1 and 2 the number of columns of the approximate solution factors Z_k and $Z_{k\ell}$ is increased by m and ℓm, respectively. To keep the low rank structure in the Cholesky factors in case of large m and slow convergence, we can replace the iterate by its low rank approximation computed via the updated singular value decomposition, see [1] for details.

4 Numerical example

In this section we present the results of numerical experiments. Computations were carried out using MATLAB 6.5 with machine precision $\varepsilon \approx 2.22 \times 10^{-16}$.

Consider the 2D instationary Stokes equation describing the flow of an incompressible fluid in a domain $\Omega \subset \mathbb{R}^2$: $v_t = \Delta v - \nabla p + f$, $\operatorname{div} v = 0$ with appropriate initial and boundary conditions. The finite differences discretization on a uniform staggered grid with 70×70 grid points leads to the descriptor system $E\dot{x}(t) = Ax(t) + Bu(t)$ of order $n = 14980$ with the matrices

$$E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & 0 \end{bmatrix}.$$

Here A_{11} is the discrete Laplace operator, A_{12} and A_{12}^T are the discrete gradient and divergence operators, respectively. Without loss of generality we may assume that the matrix A_{12} has full column rank. In this case the spectral projections P_l and P_r onto the left and right deflating subspaces corresponding to the finite eigenvalues of $\lambda E - A$ have the following form

$$P_l = \begin{bmatrix} \Pi & -\Pi A_{11} A_{12} (A_{12}^T A_{12})^{-1} \\ 0 & 0 \end{bmatrix}, \qquad P_r = \begin{bmatrix} \Pi & 0 \\ -(A_{12}^T A_{12})^{-1} A_{12}^T A_{11} \Pi & 0 \end{bmatrix},$$

where $\Pi = I - A_{12}(A_{12}^T A_{12})^{-1} A_{12}^T$ is a projection onto the null space of A_{12}^T along the image of A_{12}, see [13]. The matrix $B \in \mathbb{R}^{14980,1}$ is chosen at random.

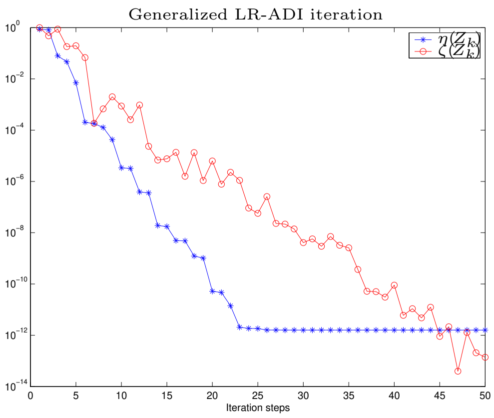

Figure 1: The convergence history of the generalized LR-ADI method with $k_{\max} = 15$.

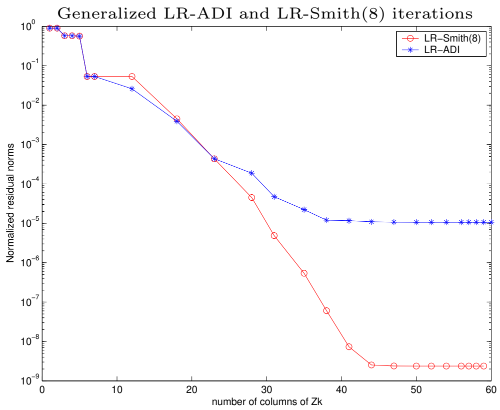

Figure 2: The convergence of the generalized LR-ADI and LR-Smith(8) methods.

In Figure 1 we present the convergence history for the generalized LR-ADI method in terms of the normalized residual norm $\eta(Z_k)$ as in (2.7) and the ratio $\zeta(Z_k) = \|Z^{(k)}\|_F / \|Z_k\|_F$. The number of ADI shift parameters is 15. One can see that the generalized LR-ADI method converges fast and the solution matrix of order 14980 can be approximated quite accurately by a matrix of rank 40. Note that the normalized residual norm $\eta(Z_k)$ stagnates on a relatively small level 10^{-12}, which is caused by round-off errors.

Figure 2 shows the convergence of the generalized LR-ADI method with 8 shift parameters and the generalized LR-Smith(8) method. One can see that for the small number of shift parameters the generalized LR-Smith(8) method gives the better approximate solution factor than the generalized LR-ADI method.

References

[1] A.C. Antoulas, D.C. Sorensen, and S. Gugercin. A modified low-rank Smith method for large-scale Lyapunov equations. *Numerical Algorithms*, 32(1):27–55, 2003.

[2] A.C. Antoulas, D.C. Sorensen, and Y. Zhou. On the decay rate of the Hankel singular values and related issues. *Systems Control Lett.*, 46(5):323–342, 2002.

[3] R.H. Bartels and G.W. Stewart. Solution of the equation $AX + XB = C$. *Comm. ACM*, 15(9):820–826, 1972.

[4] D. Estévez Schwarz and C. Tischendorf. Structural analysis for electric circuits and consequences for MNA. *Int. J. Circ. Theor. Appl.*, 28:131–162, 2000.

[5] S.J. Hammarling. Numerical solution of the stable non-negative definite Lyapunov equation. *IMA J. Numer. Anal.*, 2:303–323, 1982.

[6] J.-R. Li and T. Penzl. Square root method via Cholesky factor ADI. In *Proceedings of the 14th International Symposium on Mathematical Theory of Networks and Systems Perpignan, France*, 2000.

[7] J.-R. Li and J. White. Low rank solution of Lyapunov equations. *SIAM J. Matrix Anal. Appl.*, 24(1):260–280, 2002.

[8] R. März. Canonical projectors for linear differential algebraic equations. *Comput. Math. Appl.*, 31(4-5):121–135, 1996.

[9] D.W. Peaceman and H.H. Rachford. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3:28–41, 1955.

[10] T. Penzl. A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, 21(4):1401–1418, 1999/00.

[11] T. Penzl. Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Systems Control Lett.*, 40(2):139–144, 2000.

[12] T. Stykel. *Analysis and Numerical Solution of Generalized Lyapunov Equations*. Ph.D. thesis, Institut für Mathematik, Technische Universität Berlin, Berlin, 2002.

[13] T. Stykel. Balanced truncation model reduction for semidiscretized Stokes equation. To appear in *Linear Algebra Appl.*, 2004.

[14] E. Wachspress. Iterative solution of the Lyapunov matrix equation. *Appl. Math. Lett.*, 1:87–90, 1988.