Journal of Computational and Applied Mathematics 236 (2011) 1531–1542

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Block Arnoldi-based methods for large scale discrete-time algebraic Riccati equations

A. Bouhamidi[a], M. Heyouni[b], K. Jbilou[a,*]

[a] Univ. Lille- Nord de France, ULCO, LMPA, 50 rue F. Buisson BP699, F-62228 Calais Cedex, France

[b] ENSAH, Ecole Nationale des Sciences Appliquées d'Al-Hoceima, Université Mohammed Premier, Oujda, Maroc

ARTICLE INFO

Article history:
Received 6 April 2011
Received in revised form 21 June 2011

Keywords:
Block Arnoldi
Block Krylov
Discrete Riccati equations
Extended
Low rank
Newton

ABSTRACT

In the present paper, we present block Arnoldi-based methods for the computation of low rank approximate solutions of large discrete-time algebraic Riccati equations (DARE). The proposed methods are projection methods onto block or extended block Krylov subspaces. We give new upper bounds for the norm of the error obtained by applying these block Arnoldi-based processes. We also introduce the Newton method combined with the block Arnoldi algorithm and present some numerical experiments with comparisons between these methods.

1. Introduction

Algebraic Riccati equations (discrete-time or continuous-time) play a fundamental role in many areas such as control, filter design theory, model reduction problems, differential equations and robust control problems [1–3]. They arise in linear-quadratic regulator problems, H_∞ or H_2-control, model reduction problems and many others; see, e.g., [1,2] and the references therein. For historical developments, applications and importance of algebraic Riccati equations, we also refer to [4,5] and the references therein.

In this paper we propose numerical methods for large discrete-time algebraic Riccati equations (DARE) of the form

$$X = A^T XA - A^T XB(R + B^T XB)^{-1} B^T XA + C^T C, \tag{1}$$

where $A, X \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times s}$, $C \in \mathbb{R}^{s\times n}$ and $R \in \mathbb{R}^{s\times s}$ with $R = R^T > 0$ and $s \ll n$.

Using the following relations

$$(I + U)^{-1} = I - U(I + U)^{-1} \tag{2}$$

and

$$V(I + UV)^{-1} = (I + VU)^{-1} V, \tag{3}$$

we can show (see [6]) that the matrix equation (1), could also be written as

$$X - A^T X(I + GX)^{-1} A - C^T C = 0, \tag{4}$$

* Corresponding author.
E-mail addresses: bouhamidi@lmpa.univ-littoral.fr (A. Bouhamidi), mohammed.heyouni@gmail.com (M. Heyouni), jbilou@lmpa.univ-littoral.fr (K. Jbilou).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.09.017

where $G = B R^{-1} B^T$. Eq. (1) arises from the discrete-time linear quadratic problem

$$\min_u \frac{1}{2} \sum_{k=0}^{\infty} (y_k^T y_k + u_k^T R u_k)$$

subject to the constraints

$$x_{k+1} = A x_k + B u_k,$$
$$y_k = C x_k, \quad k = 0, 1, \ldots$$

where x_k is the state vector of dimension n, u_k is a control vector of $\mathbb{R}^s$, and y_k is the output vector of length s.

The pair (A, B) is said to be d-stabilizable if there exists a matrix L such that $A - BL$ is d-stable which means that all its eigenvalues are in the open unit disk. The pair (A, C) is said to be detectable if the pair (A^T, C^T) is d-stabilizable. Throughout this paper, we assume that these conditions are satisfied. It follows that in such a case the optimal feedback law u_k is given by

$$u_k = -(R + B^T X_d B)^{-1} B^T X_d A x_k; \quad k = 0, 1, \ldots$$

where X_d is the unique d-stabilizing positive semi-definite solution of Eq. (1), (the eigenvalues of the matrix $A - B(R + B^T X_d B)^{-1} B^T X_d A$ are in the open unit disk); see [2] for more details.

During the last decades, many numerical methods for solving the DARE (1) with small and dense matrices have been developed. The standard computational methods are based on the Schur and structure-preserving Schur methods [7–9,3,10], matrix sign function methods [7], Newton-type methods [11–14,2] and the symplectic Lanczos method [15].

Projection methods, on block Krylov subspaces, using the block Arnoldi process, and on matrix Krylov subspaces, using the global Arnoldi process, have also been applied to compute low rank approximate solutions to large and sparse algebraic Riccati equations [16–18]. However, no convergence have been given when using Arnoldi-type methods for solving discrete-time algebraic Riccati equations.

In this paper, we present block projection methods that allow us to compute low rank approximations to the stabilizing solution of (1). We project the initial problem onto a block or onto an extended block Krylov subspace, generated by the pair (A^T, C^T) and we obtain a low dimensional DARE that is solved by a standard algorithm such as the Schur method [8]. We give new theoretical results such as upper bounds for the norm of the error. We also introduce the Newton method associated with the block Arnoldi algorithm used for solving, at each Newton's iteration, the obtained Stein matrix equation.

The paper is organized as follows. In Section 2, we review the block Arnoldi algorithm with some properties. In Section 3, we apply the block Arnoldi method to extract low rank approximate solutions to the DARE (1). The main theoretical results are given in Section 4 where some new bounds for the norm of the error are derived. In Section 5, we will see how to use the Newton–Hewer method in combination with the block Arnoldi algorithm to solve (1). Section 5 is devoted to the extended block Arnoldi for the DARE (1). In the last section, we give some numerical experiments and comparisons between the mentioned three methods.

Throughout this paper, we use the following notations: The 2-norm and the F-norm of matrices will be denoted by $\|.\|_2$ and $\|.\|$ respectively. The scalar product associated with the Frobenius norm is defined by $\langle Z_1, Z_2 \rangle = \mathrm{trace}(Z_1^T Z_2)$. Finally, I_r and $O_{r \times l}$ will denote the identity matrix of size $r \times r$ and the zero matrix of size $r \times l$, respectively.

2. The block Arnoldi algorithm

We first recall the block Arnoldi process applied to the pair (F, G) where $F \in \mathbb{R}^{n \times n}$ and $G \in \mathbb{R}^{n \times s}$. The block projection subspace $\mathcal{K}_m(F, G)$ is given by

$$\mathcal{K}_m(F, G) = \mathrm{Range}([G, F\,G, F^2\,G, \ldots, F^{m-1}\,G]).$$

Note that the subspace $\mathcal{K}_m(F, G)$ is a simple sum of s Krylov subspaces

$$\mathcal{K}_m(F, G) = \overset{s}{\underset{i=1}{+}} \mathbb{K}_m(F, G^{(i)}),$$

where $\mathbb{K}_m(F, G^{(i)})$ is the Krylov subspace generated by F and the i-th column $G^{(i)}$ of the $n \times s$ matrix G. The block Arnoldi algorithm is summarized as (see for example [16,17])

Algorithm 1 The block Arnoldi algorithm

- Compute the QR factorization $G = V_1 R_1$,
- For $j = 1, \ldots, m$ do
  - $W_j = F\,V_j$,
  - for $i = 1, 2, \ldots, j$ do
    - . $H_{i,j} = V_i^T\,W_j$,
    - . $W_j = W_j - V_j\,H_{i,j}$,