UMINF 10.03

Any positive residual curve is possible for the Arnoldi method for Lyapunov matrix equations

Carl Christian Kjelgaard Mikkelsen

March 17th, 2010

Abstract In this paper we consider the Lyapunov equation $AX + XA^T + bb^T = 0$, where $A \in \mathbb{R}^{n\times n}$ is negative definite and $b \in \mathbb{R}^n$. The Arnoldi method is an iterative algorithm which can be used to compute an approximate solution. However, the convergence can be very slow and in this paper we show how to explicitly construct a Lyapunov equation with a given residual curve. The matrix A can be chosen as symmetric negative definite and it is possible to arbitrarily specify the elements on the diagonal of the Cholesky factor of $-A$. If the symmetry is dropped, then it is possible to arbitrarily specify $A+A^T$, while retaining the residual curve.

Keywords Lyapunov matrix equations · Arnoldi method · Krylov subspace methods

Mathematics Subject Classification (2000) 15A24, 65F30

1 Introduction

The Arnoldi method can be used to solve the continuous time Lyapunov matrix equation

$$AX + XA^T + BB^T = 0, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is negative definite and $B \in \mathbb{R}^{n\times p}$. In this paper we deal exclusively with the simplest case of

$$p = 1$$

which we emphasize by writing

$$AX + XA^T + bb^T = 0, \quad b \in \mathbb{R}^n. \tag{1.2}$$

Financial support has been provided by the Swedish Foundation for Strategic Research under the frame program grant A3 02:128

Carl Christian Kjelgaard Mikkelsen
Department of Computing Science and HPC2N, Umeå University, SE-90187 Umeå, Sweden
E-mail: spock@cs.umu.se

Equation (1.1) plays an important role in model reduction of linear time invariant dynamical systems [1]. If A is stable, then equation (1.1) has a unique solution X which is symmetric positive semidefinite. Frequently, the matrices A and B are such that $X = X(A,B)$ admits a good low rank approximation.

The Arnoldi method is an iterative method which constructs a sequence $\{X_j\}$ of low rank approximations to X. The method was introduced by Saad [7] who considered the case of $p = 1$ and extended it to the general case of $p > 1$ by Jaimoukha and Kasenally [3]. The convergence of the Arnoldi method is a nontrivial question. Recently, Simoncini and Druskin [9] have derived a set of theorems describing the convergence in terms of the numerical range of A. It is well known that the Arnoldi method may require a large number of iterations before converging, and Kressner [5] has developed a variation specifically designed to combat this problem. The storage requirements are reduced at the cost of an increased number of arithmetic operations.

The contribution of the present paper is to demonstrate that any positive residual curve is possible for the Arnoldi method. We show how to explicitly construct both symmetric and nonsymmetric Lyapunov equations for which the Arnoldi method returns a given residual curve. Theorem 3.2 has been extracted from [6], but the proof has been simplified.

Several authors, including Hu and Reichel [2] as well as Jbilou and Riquet [4], have all derived algorithms having Saad's algorithm [7] as a special case. Therefore, any positive residual curve is possible for their algorithms as well. Simoncini [8] has developed a related algorithm which applies both A and A^{-1} to accelerate convergence. Our analysis does not apply to this algorithm.

The paper is organized as follows. In Section 2 we state the basic Arnoldi method for Lyapunov equations along with a few results which are necessary in order to derive the main theorems, which in turn are presented in Section 3. Section 4 contains MATLAB implementations of our algorithms and we exhibit a few matrices for which the residual curve is rather unusual.

We use MATLAB notation occasionally. Specifically, if A is an n by n matrix, $A(:,j)$ will refer to the jth column of A and $A(1:j,1:j)$ will refer to the upper left j by j corner of A. The notation $e_j^{(n)}$ is used for the jth column of the n by n identity matrix I_n. If the dimension is clear from the context, the superscript will be omitted and we shall write e_j, rather than $e_j^{(n)}$.

Exact arithmetic is assumed throughout the paper.

2 The Arnoldi method for Lyapunov equations

In this section we give a brief description of the basis Arnoldi method for Lyapunov matrix equations. We deal exclusively with the simplest base of $p = 1$.

Let $A \in \mathbb{R}^{n\times n}$ and let $b \in \mathbb{R}$. The standard Krylov subspace $K_j(A,b)$ is given by

$$K_j(A,b) = \text{span}\{b, Ab, A^2b, \dots, A^{j-1}b\}, \quad j = 1,2,\dots.$$

It is clear that

$$K_j(A,b) \subseteq K_{j+1}(A,b)$$

and

$$K(A,b) = \cup_{j=1}^{\infty} K_j(A,b) \subseteq \mathbb{R}^n$$

is the smallest A invariant subspace containing b. The smallest integer m such that

$$K(A,b) = K_m(A,b)$$

is called the grade of b with respect to A. Now, let $\{v_i\}_{i=1}^m$ be a sequence of orthonormal vectors such that

$$K_j(A,b) = \operatorname{span}\{v_1, v_2, \ldots, v_j\}$$

and let $V_j \in \mathbb{R}^{n\times j}$ be given by

$$V_j = \begin{bmatrix} v_1 \ v_2 \ \ldots \ v_j \end{bmatrix}.$$

Then

$$AV_m = V_m H_m$$

for some matrix $H_m \in \mathbb{R}^{m\times m}$ simply because the columns of V_m span $K(A,b)$. In fact, there is only one choice for H_m, namely

$$H_m = V_m^T A V_m$$

because $V_m^T V_m = I_m$. In addition, H_m must be upper Hessenberg

$$H_m = \begin{bmatrix} h_{11} & h_{12} & \ldots & \ldots & h_{1m} \\ h_{21} & h_{22} & \ldots & \ldots & h_{2m} \\ 0 & h_{32} & h_{33} & \ldots & h_{3m} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & h_{m,m-1} & h_{mm} \end{bmatrix},$$

simply because

$$AK_j(A,b) \subset K_{j+1}(A,b), \quad j = 1,2,\ldots,m-1.$$

Now, consider the Lyapunov equation 1.2. The standard Arnoldi method seeks an approximation of X of the form

$$X_j = V_j Y_j V_j^T, \quad Y_j \in \mathbb{R}^{j\times j}$$

such that the corresponding residual given by

$$R(X_j) = AX_j + X_j A^T + bb^T$$

satisfies the Galerkin condition

$$V_j^T R(X_j) V_j = 0.$$

This condition is satisfied if and only if Y_j solves the reduced order equation

$$H_j Y_j + Y_j H_j^T + \beta^2 e_1^{(j)} {e_1^{(j)}}^T = 0 \tag{2.1}$$

where

$$H_j = V_j^T A V_j \in \mathbb{R}^{j\times j}, \quad \text{and} \quad \beta = \|b\|_2.$$

If A is negative definite, then H_j is negative definite and the reduced order equation (2.1) has a unique solution Y_j. It is straight forward to verify that the Frobenius norm of the residual satisfies

$$\|R(X_j)\|_F = \sqrt{2}|h_{j+1,j}|\|Y_j(:,j)\|_F. \tag{2.2}$$

It is important to realize that X_j depends exclusively on the subspace $K_j(A,b)$. In particular, it is independent of the choice of orthonormal basis! In short, we are free to choose the basis which is most suitable to our analysis. For the sake of stability, the typical implementation of the Arnoldi method uses the Arnoldi algorithm to construct a sequence of orthonormal vectors $\{v_j\}_{j=1}^m$ such that

$$K_j(A,b) = \text{span}\{v_1, v_2, \dots v_j\}, \quad j = 1,2,\dots,m$$

while simultaneously extracting the matrices

$$H_j = V_j^T A V_j, \quad \text{where} \quad V_j = \begin{bmatrix} v_1 \; v_2 \; \dots \; v_j \end{bmatrix}.$$

The reduced order equations are solved using one of the dense methods, say, Bartel-Stewart's method. The norm of the residual is computed using equation (2.2) which does not require the n by n residual matrix to be formed explicitly.

The Arnoldi method enjoys the following finite termination property,

$$\|R(X_j)\|_F > 0, \quad j < m, \quad \text{and} \quad \|R(X_m)\|_F = 0$$

where m is the degree of b with respect to A. This property has mainly theoretical interest, simply because m is often comparable to n and we can rarely afford to complete m iterations.

Now, given a sequence of positive real numbers $\{r_j\}_{j=1}^{n-1}$ we seek a negative definite matrix $A \in \mathbb{R}^{n\times n}$ and a vector $b \in \mathbb{R}^n$ such that $K(A,b) = \mathbb{R}^n$ and the residual curve for the Arnoldi method applied to the Lyapunov equation (1.2) satisfies

$$\|R(X_j)\|_F = r_j, \quad j = 1,2,\dots,n-1, \quad \text{and} \quad \|R(X_n)\|_F = 0.$$

The fundamental problem is the nontrivial relationship between A, and b which define the Lyapunov equations and the matrices V_m, and H_m which determine the residuals $R(X_j)$. We must determine a class of matrices for which this relationship is particularly simple. Ideally, we would like $V_n = I_n$ and $H_n = A$.

Let $W_j(A,b) \in \mathbb{R}^{n\times j}$ be the matrix given by

$$W_j(A,b) = \begin{bmatrix} b \; Ab \; A^2b \; \dots A^{j-1}b \end{bmatrix}. \tag{2.3}$$

Then $W_n(A,b)$ is nonsingular if and only if $K(A,b) = \mathbb{R}^n$. We have the following lemma.

Lemma 2.1 *Let $A \in \mathbb{R}^n$ be an upper Hessenberg matrix which is unreduced, i.e.*

$$a_{j+1,j} \neq 0, \quad j = 1,2,\dots,n-1,$$

and let $W_j(A,e_1)$ be given by equation (2.3). Then $W_m(A,e_1)$ is nonsingular and

$$K_j(A,e_1) = span\{e_1,e_2,\dots,e_j\}, \quad j = 1,2,\dots,n$$

where e_i is the ith column of the n by n identity matrix.

Proof The lemma follows immediately from the sparsity pattern of A. If $''\times''$ denotes an element which is not necessarily zero and if $''*''$ denotes an element which is nonzero, then

$$H_m = \begin{bmatrix} \times & \times & \dots & \dots & \times \\ * & \times & \dots & \dots & \times \\ 0 & * & \times & \dots & \times \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & * & \times \end{bmatrix} \quad \Rightarrow \quad W_n(A,e_1) = \begin{bmatrix} 1 & \times & \dots & \dots & \times \\ 0 & * & \dots & \dots & \times \\ 0 & 0 & * & \dots & \times \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & * \end{bmatrix}.$$

In short, by limiting ourselves to equations of the type

$$AX + XA^T + e_1^{(n)} {e_1^{(n)}}^T = 0,$$

where $A \in \mathbb{R}^{n\times n}$ is upper Hessenberg and

$$a_{j+1,j} \neq 0, \quad j = 1,2,\dots,n-1,$$

then we may chose

$$V_j = \begin{bmatrix} e_1 \ e_2 \ \dots \ e_j \end{bmatrix}$$

and the reduced order equation are merely

$$A_j Y_j + Y_j A_j^T + e_1^{(j)} {e_1^{(j)}}^T = 0$$

where

$$A_j = A(1:j,1:j)$$

is the upper left j by j corner of A. In addition, the expression for the Frobenius norm of the jth residual simplifies to

$$\|R(X_j)\|_F = \sqrt{2}\,|a_{j+1,j}|\|Y_j(:,j)\|_F.$$

Now, the matrix Y_j depends exclusively on the entries of A_j. In particular, it is independent of the entry $a_{j+1,j}$. It follows that if

$$\|Y_j(:,j)\|_F \neq 0,$$

then we may have

$$\|R(X_j)\|_F = r_j,$$

simply by choosing

$$a_{j+1,j} = \frac{r_j}{\sqrt{2}\,\|Y_j(:,j)\|_F} > 0,$$

and this choice will not affect the previous residuals. The following lemma is now relevant.

Lemma 2.2 *Let $H \in \mathbb{R}^{n \times n}$ be a stable matrix and let Y be the solution of the Lyapunov equation*

$$HY + YH^T + e_1^{(n)} e_1^{(n)^T} = 0.$$

If H is upper Hessenberg and if

$$h_{j+1,j} \neq 0, \quad j = 1,2,\dots,n-1,$$

then the last column of Y is nonzero.

Proof The matrix Y is symmetric positive semidefinite with

$$\text{Ran} Y = K(H, e_1)$$

simply because H is stable and the inhomogeneous term is symmetric positive semidefinite. However, by Lemma 2.1

$$K(H, e_1) = \mathbb{R}^n$$

and since

$$\text{Ran} Y = K(H, e_1)$$

we see that Y is nonsingular. It follows, that every column of Y is nonzero.

3 The main results

In this section we show how to construct Lyapunov equations for which the Arnoldi method has a given residual curve.

Theorem 3.1 *Let $n > 1$ be an integer and let $r_j > 0$ for $j = 1,2,\dots,n-1$. Then there exists a symmetric negative definite matrix $A \in \mathbb{R}^{n \times n}$ such that the residual curve for the Arnoldi method applied to*

$$AX + XA^T + e_1^{(n)} e_1^{(n)^T} = 0$$

satisfies

$$\|R_j\|_F = r_j, \quad j = 1,2,\dots,n-1, \quad \textit{and} \quad \|R_n\|_F = 0.$$

Proof Consider Algorithm 1. We claim that the algorithm is well defined and that there are no divisions by zero. Let P_k be the statement that

"A_k is symmetric negative definite and tridiagonal with $a_{j+1,j} > 0$ for $j < k$"

and let

$$\Omega = \{k \in \{1,2,\dots,n\} : P_k \text{ is true}\}.$$

We claim that

$$\Omega = \{1,2,\dots,n\}.$$

First we notice that $1 \in \Omega$, because

$$A_1 = -l_{11}^2 < 0$$

Algorithm 1 The symmetric negative definite case

Input: $l_{jj} > 0, j = 1,2,\ldots,n$.
Output: $A = A_n = -L_n L_n^T < 0$.

for j=1,2,. . . ,n **do**
  Define
$$L_j := \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ & \ddots & \ddots & \\ & & l_{j,j-1} & l_{jj} \end{bmatrix}.$$
  Define $A_j := -L_j L_j^T$.
  if $j < n$ **then**
    Solve
$$A_j Y_j + Y_j A_j^T + e_1^{(j)} {e_1^{(j)}}^T = 0,$$
    with respect to Y_j.
    Define
$$l_{j+1,j} := -\frac{r_j}{\sqrt{2} l_{jj} \|Y_j(:,j)\|_2}.$$
  end if
end for

is symmetric negative definite. Since A_1 is 1 by 1 there are no off-diagonal elements and there is nothing further to show. Now, assume that $1 < k \leq n$, and

$$\{1,2,\ldots,k-1\} \subseteq \Omega.$$

We claim that $k \in \Omega$. It is clear, that

$$A_k = -L_k L_k^T$$

is a symmetric negative definite and tridiagonal matrix, regardless of the values of

$$l_{j+1,j}, \quad j = 1,2,\ldots,k-1,$$

simply because L_k is a nonsingular, lower triangular and bidiagonal matrix. Since we are assuming

$$\{1,2,\ldots,k-1\} \subseteq \Omega,$$

it only remains to be shown that

$$a_{k,k-1} > 0.$$

By assumption, $k-1 \in \Omega$. Therefore A_{k-1} is a (symmetric) negative definite matrix. In particular, A_{k-1} is stable and Y_{k-1} is well defined. By Lemma 2.2 the last column of Y_{k-1} cannot be zero. Hence, $l_{k,k-1}$ is well defined (no division by zero) and

$$l_{k,k-1} = -\frac{r_{k-1}}{\sqrt{2}\, l_{k-1,k-1} \|Y_{k-1}(:,k-1)\|_2} < 0.$$

Therefore, the matrix L_k is well defined and with $A_k = -L_k L_k^T$ we have

$$a_{k,k-1} = -l_{k,k-1} l_{k-1,k-1} = \frac{r_{k-1}}{\sqrt{2}\|Y_{k-1}(:,k-1)\|_2} > 0.$$

It follows that $k \in \Omega$ and by the well ordering principle

$$\Omega = \{1, 2, \dots, n\}.$$

We now claim that if the Arnoldi method is applied to the Lyapunov equation

$$AX + XA^T + e_1^{(n)} {e_1^{(n)}}^T = 0,$$

where $A = A_n$ is given by Algorithm 1, then the residuals satisfy

$$\|R(X_j)\|_F = r_j, \quad j = 1, 2, \dots, n-1, \quad \text{and} \quad \|R(X_n)\|_F = 0.$$

Now, because $A = A_n$ is upper Hessenberg and

$$a_{j+1,j} > 0, \quad j = 1, 2, \dots, n-1$$

it follows from Lemma 2.1 that we may choose

$$V_j = \begin{bmatrix} e_1 & e_2 & \dots & e_j \end{bmatrix}$$

and then the corresponding reduced order equations are given by

$$A_j Y_j + Y_j A_j^T + e_1^{(j)} {e_1^{(j)}}^T = 0, \quad j = 1, 2, \dots, n-1,$$

where

$$A_j = A(1:j, 1:j) \in \mathbb{R}^{j \times j}.$$

By Lemma 2.2 the last column/row of Y_j is nonzero. The Frobenius norm of the residual is given by

$$\|R(X_j)\|_F = \sqrt{2} |a_{j+1,j}| \|Y_j(:, j)\|_2.$$

In our case we have deliberately chosen

$$a_{j+1,j} = \frac{r_j}{\sqrt{2}\|Y_j(:, j)\|_2} > 0$$

from which follows

$$\|R(X_j)\|_F = r_j, \quad j = 1, 2, \dots, n-1.$$

The fact that

$$\|R(X_n)\|_F = 0$$

is the finite termination property of the Arnoldi method.

Remark 1: We have deliberately imposed the condition $a_{j+1,j} > 0$ rather than the weaker condition of $a_{j+1,j} \neq 0$. Therefore, if the Arnoldi algorithm is used to compute a basis for the Krylov subspace $K(A, e_1)$, then $v_j = e_j$ for all j.

Nonsymmetric matrices can be constructed as well and in this case we have complete control over $A + A^T$.

Theorem 3.2 *Let $n > 1$ be an integer and let $r_j > 0$ for $j = 1,2,\dots,n-1$. Let $\{\lambda_j\}_{j=1}^n \subset (-\infty,0)$. Then there exists a negative definite matrix $A \in \mathbb{R}^{n\times n}$ such that the residual curve for the Arnoldi method applied to*

$$AX + XA^T + e_1^{(n)} e_1^{(n)^T} = 0, \quad e_1 = (1,0,0,\dots,0)^T$$

satisfies

$$\|R_j\|_F = r_j, \quad j = 1,2,\dots,n-1, \quad \textit{and} \quad \|R_n\|_F = 0$$

and $A + A^T$ satisfies

$$A + A^T = diag\{\lambda_1,\lambda_2,\dots,\lambda_n\}.$$

Proof Consider Algorithm 2. We claim that the algorithm is well defined and there are no divisions by zero.

Algorithm 2 The nonsymmetric negative definite case

Input: $\lambda_j < 0$, $j = 1,2,\dots,n$.
1: **for** $j = 1,2,\dots,n$ **do**
2: Define
$$a_{jj} = \frac{1}{2}\lambda_j.$$
3: Define $A_j \in \mathbb{R}^{j\times j}$ by
$$A_j := \begin{bmatrix} a_{11} & -a_{21} & & \\ a_{21} & a_{22} & \ddots & \\ & \ddots & \ddots & -a_{n,n-1} \\ & & a_{n,n-1} & a_{nn} \end{bmatrix}.$$
4: **if** $j < n$ **then**
5: Solve
$$A_j Y_j + Y_j A_j^T + e_1^{(j)} e_1^{(j)^T} = 0$$
with respect to $Y_j \in \mathbb{R}^{j\times j}$.
6: Define
$$a_{j+1,j} := \frac{r_j}{\sqrt{2}\|Y_j(:,j)\|_2}.$$
7: **end if**
8: **end for**

Let P_k be the statement

"A_k is negative definite and tridiagonal with $a_{j+1,j} > 0$ for $j < k$"

and let

$$\Omega = \{k : P_k \text{ is true}\}.$$

We claim that

$$\Omega = \{1,2,\dots,n\}.$$

First, we notice $1 \in \Omega$, because

$$A_1 = \frac{1}{2}\lambda_1 < 0$$

is negative definite. Since A_1 is 1 by 1 there are no off-diagonal elements and there is nothing further to show. Now, suppose $k \geq 2$ and assume

$$\{1,2,\dots,k-1\} \subseteq \Omega.$$

We claim that $k \in \Omega$. By assumption, $k-1 \in \Omega$ which implies A_{k-1} is stable and Y_{k-1} is well defined. It follows from Lemma 2.2, that the last column of Y_{k-1} is nonzero. Therefore, $a_{k,k-1}$ is well defined (no division by zero) and

$$a_{k,k-1} = \frac{r_{k-1}}{\sqrt{2}\,\|Y_{k-1}(:,k-1)\|_2} > 0.$$

It is clear from the definition that A_k is tridiagonal and

$$A_k + A_k^T = 2\mathrm{diag}\{a_{11}, a_{22}, \dots a_{kk}\} = \mathrm{diag}\{\lambda_1, \lambda_2, \dots, \lambda_k\}$$

is obviously negative definite. Therefore $k \in \Omega$ and

$$\Omega = \{1,2,\dots,n\}$$

by the well ordering principle.

We now claim that if the Arnoldi method is applied to the equation

$$AX + XA^T + e_1 e_1^T = 0$$

where $A = A_n$ is the matrix returned by Algorithm 2, then the residuals satisfy

$$\|R(X_j)\|_F = r_j, \quad j = 1,2,\dots,n-1, \qquad \|R(X_n)\|_F = 0.$$

Since A is upper Hessenberg and

$$a_{j+1,j} > 0$$

the reduced order equations are merely

$$A_j Y_j + Y_j A_j^T + e_1^{(j)} {e_1^{(j)}}^T = 0$$

where

$$A_j = A(1:j, 1:j).$$

The corresponding residual is given by

$$\|R(X_j)\| = \sqrt{2}\,|a_{j+1,j}|\|Y_j(:,j)\|_2,$$

but we have deliberately chosen

$$a_{j+1,j} = \frac{r_j}{\sqrt{2}\,\|Y_j(:,j)\|_2},$$

which implies

$$\|R(X_j)\| = r_j, \quad j = 1,2,\dots,n-1.$$

The fact that

$$\|R(X_n)\|_F = 0$$

is just the finite termination property of the Arnoldi method.

The following corollary is an immediate consequence of Theorem 3.2.

Corollary 3.1 *Let $n > 1$ be an integer and $r_j > 0$ for $j = 1,2,\ldots,n-1$. Let $C \in \mathbb{R}^{n\times n}$ be symmetric negative definite. There there exists a negative definite matrix $A \in \mathbb{R}^{n\times n}$ such that the residual curve for the Arnoldi method applied to*

$$AX + XA^T + e_1e_1^T = 0$$

satisfies

$$\|R(X_j)\|_F = r_j, \quad j = 1,2,\ldots,n-1, \quad \textit{and} \quad \|R(X_n)\|_F = 0$$

while

$$A + A^T = C.$$

Proof Since C is symmetric negative definite it can be written as

$$C = U\Lambda U^T,$$

where

$$U^TU = I_n \quad \text{and} \quad \Lambda = \text{diag}\{\lambda_{11}, \lambda_{22}, \ldots, \lambda_{nn}\} < 0.$$

By Theorem 3.2 there exists a negative definite matrix A_0, such that the residual curve for the Arnoldi method applied to

$$A_0X + XA_0^T + e_1^{(n)}{e_1^{(n)}}^T = 0,$$

satisfies

$$\|R(X_j)\|_F = r_j, \quad j = 1,2\ldots,n-1, \quad \|R(X_n)\|_F = 0$$

while $A_0 + A_0^T = \Lambda$. Now, define

$$A = UA_0U^T, \quad b = Ue_1^{(n)}.$$

Then, $A + A^T = C$ and since the residual curve it not affected by the orthonormal transformation U the proof is complete.

4 Numerical experiments

In this section we state a short MATLAB implementation of Algorithms 1 and 2 and we exhibit a couple of matrices for which the residual curve is rather unusual.

The two algorithms are straight forward and implementations can be found in Figure 4.1 and Figure 4.2.

The Arnoldi method is defined whenever A is negative definite. However, if the exact solution of the corresponding Lyapunov equation does not admit a good low rank approximation, then there is no hope of rapid convergence.

Example 1: Consider the residual curve

$$r_j = 1, \quad 1,2,\ldots,n-1, \quad r_n = 0. \tag{4.1}$$

Fig. 4.1 A MATLAB implementation of Algorithm 1

```
function [A, L]=my_snd(r,l)

n=length(l); A=zeros(n,n); L=zeros(n,n); Q=zeros(n,n); Q(1,1)=1;

for j=1:n
  L(j,j)=l(j);
  A(1:j,1:j)=-L(1:j,1:j)*L(1:j,1:j)';
  if (j<n)
    Y=lyap(A(1:j,1:j),Q(1:j,1:j));
    L(j+1,j)=r(j)/(sqrt(2)*L(j,j)*norm(Y(:,j),2));
  end
end
```

Fig. 4.2 A MATLAB implementation of Algorithm 2

```
function A=my_nd(r,lambda)

n=length(lambda); A=zeros(n,n); Q=zeros(n,n); Q(1,1)=1;

for j=1:n
  A(j,j)=lambda(j)/2;
  if (j<n)
    Y(1:j,1:j)=lyap(A(1:j,1:j),Q(1:j,1:j));
    A(j+1,j)=r(j)/(sqrt(2)*norm(Y(1:j,j),2));
    A(j,j+1)=-A(j+1,j);
  end
end
```

We choose $n = 500$ and used Algorithm 1 to construct a symmetric negative definite matrix A_1 with a unit lower triangular Cholesky factor L_1. We solved the equation

$$A_1 X_1 + X_1 A_1^T + e_1 e_1^T = 0 \tag{4.2}$$

using the MATLAB function 'lyap'. The Frobenius norm (relative) residual was

$$\|A_1 \hat{X}_1 + \hat{X}_1 A_1^T + e_1 e_1^T\|_F \approx 1.0922 \cdot 10^{-10}.$$

We computed the singular values for $\hat{X}_1$ using the MATLAB function 'svd', see Figure 4.3. It is clear that $\hat{X}_1$ can be approximated accurately with a low rank matrix. However, in view of the residual curve, the Arnoldi method cannot hope to retrieve a good low rank approximation for X_1. Finally, we applied our own implementation of the Arnoldi method to equation (4.2) and compared the computed residual curve

$$r'_j \approx 1, \quad 1,2,\ldots,n-1, \quad r'_n \approx 0$$

to the target given by equation (4.1). The result is displayed in Figure 4.4. We see that the (relative) difference between the computed residual and the target is less than $2.5 \cdot 10^{-9}$ for all but the nth iteration, where the *computed* residual does not assume the target value of zero.

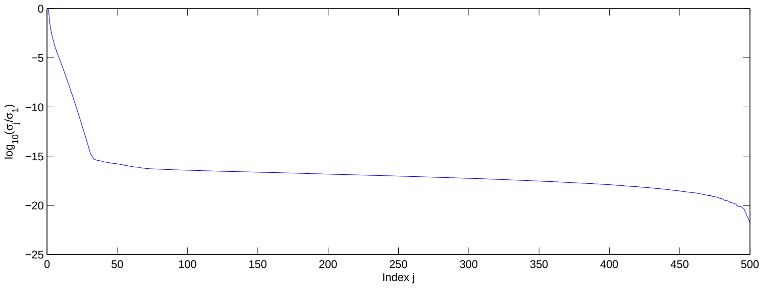

Fig. 4.3 The decay of the singular values for the computed solution of the symmetric Lyapunov equation (4.2).

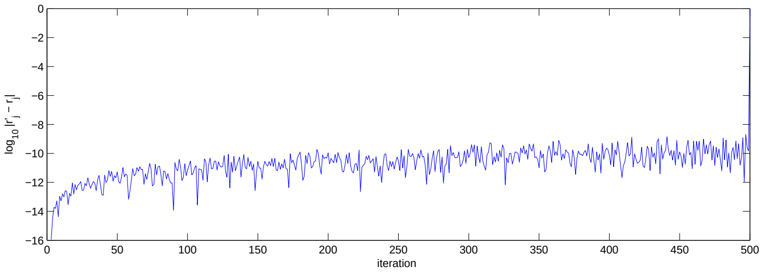

Fig. 4.4 Comparison between the computed, $\{r'_j\}$ and the desired residual curve $\{r_j\}$ for a symmetric Lyapunov equation generated by Algorithm 1 corresponding to a constant residual curve.

Example 2: Consider the residual curve

$$r_j = j, \quad 1,2,\ldots,n-1, \quad r_n = 0. \tag{4.3}$$

We chose $n = 500$ and used Algorithm 2 to construct a nonsymmetric negative definite matrix A_2 for which $A_2 + A_2^T = I_n$. We solved the equation

$$A_2X_2 + X_2A_2^T + e_1e_1^T = 0 \tag{4.4}$$

using the MATLAB function 'lyap'. The Frobenius norm (relative) residual was

$$\|A_2\hat{X}_2 + \hat{X}_2A_2^T + e_1e_1^T\|_F \approx 4.0491 \cdot 10^{-10}.$$

We computed the singular values for $\hat{X}_2$ using the MATLAB function 'svd', see Figure 4.5. It is clear that $\hat{X}_2$ can be approximated accurately with a low rank matrix. However, in view of the residual curve, the Arnoldi method cannot hope to retrieve a good low rank approximation for X_2. Finally, we applied our own implementation of the Arnoldi method to equation (4.4) and compared the computed residual curve

$$r'_j \approx j, \quad 1,2,\ldots,n-1, \quad r'_n \approx 0$$

to the target given by equation (4.3). The result is displayed in Figure 4.6. We see that the relative difference between the computed residual and the target is less than

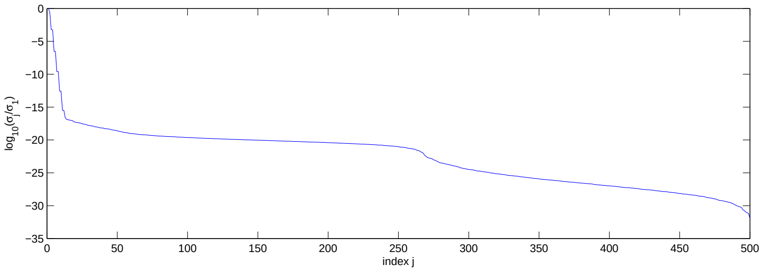

Fig. 4.5 The decay of the singular values for the computed solution of the nonsymmetric Lyapunov equation (4.4).

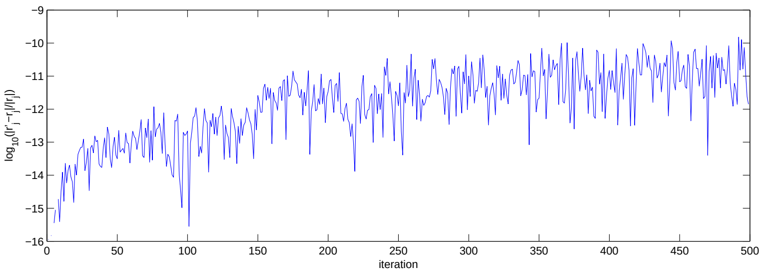

Fig. 4.6 Comparison between the computed, $\{r'_j\}$ and the desired residual curve $\{r_j\}$ for the nonsymmetric Lyapunov equation generated by Algorithm 2 corresponding to a linear residual curve.

$2.0 \cdot 10^{-10}$ for all but the nth iteration, where the *computed* residual does not assume the target value of zero.

Remark 2: In the jth iteration of both Algorithm 1 and Algorithm 2 we solve a Lyapunov equation of dimension j using a dense method and $O(j^3)$ arithmetic operations. Therefore the total number of arithmetic operations is $O(n^4)$, limiting the construction of test equations to small values of n. It is entirely possible that the workload can be reduced, but our goal is merely to show that any positive residual curve is possible.

5 Conclusion

We have shown that if $\{r_j\}_{j=1}^{n-1}$ is a sequence of positive real numbers, then it is possible to find either a symmetric or a nonsymmetric negative definite matrix A for which the Arnoldi method applied to

$$AX + XA^T + e_1 e_1^T = 0$$

produces the residual curve

$$\|R(X_j)\|_F = r_j, \quad j = 1, 2, \ldots, n-1 \quad \text{and} \quad \|R(X_n)\|_F = 0.$$

In each case there is a certain amount of freedom in the construction. If A is symmetric, then it is possible to specify the main diagonal of the Cholesky factorization of $-A$ arbitrarily. If A is nonsymmetric, then we may specify $A+A^T$ arbitrarily.

Our MATLAB codes can be used to generate equations for which the Arnoldi method will not be successful. The robustness of other methods can be measured against these equations.

Our analysis emphasizes that the Arnoldi method should be used with caution. Theoretically, it is applicable whenever A is negative definite, but there is no guarantee that the convergence will be monotone or that it will even converge before the nth iteration. In addition, we have seen that it is entirely possible that the exact solution admits a good low rank approximation, but the residual curve is nondecreasing. We conclude that the method does not exploit the low rank phenomenon.

Simoncini and Druskin [9] have bounded the convergence rate in terms of the numerical range of the matrix A. It would be interesting to determine conditions on A and b for which the convergence of the Arnoldi method is both rapid and monotone.

References

1. Antoulas, A.C.: Approximation of Large-Scale Dynamical Systems, *Advances in Design and Control*, vol. 6. SIAM (2005). ISBN: 0-89871-529-6
2. Hu, D.Y., Reichel, L.: Krylov subspace methods for the Sylvester equation. Linear Algebra Appl. **172**, 283–313 (1992)
3. Jaimoukha, I.M., Kasenally, E.M.: Krylov subspace methods for solving large Lyapunov equations. SIAM J. Numer. Anal. **31**(1), 227–251 (1994). DOI http://dx.doi.org/10.1137/0731012
4. Jbilou, K., Riquet, A.J.: Projection methods for large Lyapunov equations. Linear Algebra Appl. **415**(2–3), 344–358 (2006)
5. Kressner, D.: Memory-efficient Krylov subspace techniques for solving large-scale Lyapunov equations. In: Proceedings of the IEEE International Symposium on Computer-Aided Control Systems, pp. 613–618. San Antonio, Texas (2008)
6. Mikkelsen, C.C.K.: Numerical methods for large Lyapunov equations. Ph.D. thesis, Purdue University, West Lafayette, Indiana, USA (2009)
7. Saad, Y.: Numerical solution of large Lyapunov equations. In: Signal Processing, Scattering and Operator Theory, and Numerical Methods (Amsterdam, 1989), pp. 503–511. Birkhäuser, Boston, MA (1990)
8. Simoncini, V.: A new iterative method for solving large-scale Lyapunov matrix equations. SIAM J. Sci. Comput. **29**(3), 1268–1288 (2007). DOI http://dx.doi.org/10.1137/06066120X
9. Simoncini, V., Druskin, V.: Convergence analysis of projection methods for the numerical solution of large Lyapunov equations. SIAM J. Numer. Anal. **47**(2), 828–843 (2009)