NICONET Report 2000-1

Condition and Error Estimates in the Solution of Lyapunov and Riccati Equations [1]

Petko Hr. Petkov[2], Da-Wei Gu[3], Mihail M. Konstantinov[4] and Volker Mehrmann[5]

January 2000

[1]This paper presents research results of the European Community BRITE-EURAM III Thematic Networks Programme NICONET and is distributed by the Working Group on Software WGS. *WGS secretariat*: Mrs. Ida Tassens, ESAT - Katholieke Universiteit Leuven, K. Mercierlaan 94, 3001-Leuven-Heverlee, BELGIUM. This report is also available by anonymous ftp from *wgs.esat.kuleuven.ac.be/pub/WGS/REPORTS/SLWN2000-1.ps.Z*

[2]Department of Automatics, Technical University of Sofia, 1756 Sofia, Bulgaria

[3]Control Systems Research, Department of Engineering, University of Leicester, Leicester LE1 7RH, U.K.

[4]University of Architecture & Civil Engineering, 1 Hr. Smirnenski Blv., 1421 Sofia, Bulgaria

[5]Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany

Abstract

The condition number estimation and the computation of residual based forward error estimates in the numerical solution of matrix algebraic continuous-time and discrete-time Lyapunov and Riccati equations is considered. The estimates implemented involve the solution of triangular Lyapunov equations along with usage of the LAPACK norm estimator. Results from numerical experiments demonstrating the performance of the estimates proposed are presented.

1 Introduction

The conditioning and forward error estimation is an important step in the solution of computational problems. In this paper we consider the estimation of condition numbers and the computation of residual-based forward error estimates pertaining to the numerical solution of matrix algebraic continuous-time and discrete-time Lyapunov and Riccati equations which arise in control theory. For this purpose we use the LAPACK matrix norm estimator [1] allowing to obtain low cost condition and forward error estimates which are usually sufficiently accurate.

The usage of the LAPACK matrix norm estimator to estimate the condition numbers for the nonsymmetric eigenproblem through the solution of matrix Sylvester equations was proposed in [3]. This estimator is used in LAPACK also to find residual-based forward error estimates in the solution of linear systems of equations according to the method presented in [2]. The close idea to implement the matrix norm estimator in the estimation of the conditioning and forward errors in the numerical solution of matrix Sylvester equation was described in [12]. However, this idea was not applied up to the moment to the condition and forward error estimation in the case of matrix Lyapunov and Riccati equations. For these equations it is possible to take advantage of the symmetry, thus reducing significantly the operation cost of the estimation for high order equations. For four types of continuous- and discrete-time Lyapunov and Riccati equations we implement the corresponding condition and forward estimates in eight Fortran 77 subroutines which may be used along different methods for solving Lyapunov and Riccati equations.

The following notation is used in the paper. $\mathcal{R}$ – the field of real numbers; $\mathcal{R}^{m\times n}$ – the space of $m \times n$ matrices $A = [a_{ij}]$ over $\mathcal{R}$; A^T – the transpose of a matrix A; $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ – the maximum and minimum singular value of A; $\|A\|_1$ – the matrix 1-norm of the matrix A; $\|A\|_2 = \sigma_{\max}(A)$ – the spectral norm of the matrix A; $\|A\|_F = (\sum |a_{ij}|^2)^{1/2})$ – the Frobenius norm; I_n – the unit $n \times n$ matrix; $A \otimes B$ – the Kronecker product of matrices A and B; $\mathrm{Vec}(A)$ – the vector, obtained by stacking the columns of A in one vector; ε – the roundoff unit of the machine arithmetic.

In what follows we shall consider the Lyapunov equation

$$A^T X + XA = C, \tag{1}$$

the discrete-time Lyapunov equation

$$A^T XA - X = C, \tag{2}$$

the Riccati equation

$$A^T X + XA + C - XDX = 0, \tag{3}$$

and the discrete-time Riccati equation

$$A^T XA - X + C - A^T XB(R + B^T XB)^{-1} B^T XA = 0 \tag{4}$$

or its equivalent

$$X = C + A^T X(I_n + DX)^{-1} A = 0, \;\; D = BR^{-1}B^T \tag{5}$$

where $A \in \mathcal{R}^{n\times n}$ and the matrices $C,\ D,\ X \in \mathcal{R}^{n\times n}$ are symmetric. In the case of the Riccati equations 3 and 4 we assume that there exists a non–negative definite solution X which stabilises $A - DX$.

The numerical solution of matrix Lyapunov and Riccati equations may face some difficulties. First of all, the corresponding equation may be *ill-conditioned*, i.e., small perturbations in the coefficient matrices $A,\ C,\ D$ may lead to large variations in the solution. As is well known, the conditioning of a problem depends neither on the method used nor on the properties of the computer architecture. So, it is necessary to have a quantitative characterisation of the conditioning in order to estimate the accuracy of solution computed.

The second difficulty is connected with the stability of the numerical method and the reliability of its implementation. The backward numerical stability of the methods for solving the Lyapunov equations is not proved and as it is known [17] that the methods for solving the Riccati equations are generally unstable. This requires to have an estimate of the forward error in the solution computed.

The paper is organised as follows. In Section 2 we review some known results which are necessary in the study of the conditioning of equations (1) - (4). In Section 3 we present an efficient method for computing condition number estimates which is based on matrix morm estimator implemented in LAPACK . In Section 4 we propose residual based forward error estimates which implement also the LAPACK norm estimator and may be used in conjuction with different methods for solving the corresponding equation. The software implementation of the condition and error estimates is based entirely on LAPACK and BLAS [16, 8, 7] subroutines. In Section 5 we present the some numerical examples demonstrating the performance of estimates implemented. Some conclusions are made in Section 6.

2 Conditioning of Lyapunov and Riccati equations

Let the coefficient matrices A, C, D in (1) - (4) be subject to perturbations ΔA, ΔC, ΔD, respectively, so that instead of the initial data we have the matrices $\tilde{A} = A + \Delta A$, $\tilde{C} = C + \Delta C$, $\tilde{D} = D + \Delta D$. The aim of the *perturbation analysis* of (1) - (4) is to investigate the variation ΔX in the solution $\tilde{X} = X + \Delta X$ due to the perturbations ΔA, ΔC, ΔD. If small perturbations in the data lead to small variations in the solution we say that the corresponding equation is *well-conditioned* and if these perturbations lead to large variations in the solution this equation is *ill-conditioned*. In the perturbation analysis of the Lyapunov and Riccati equations it is supposed that the perturbations preserve the symmetric structure of the equation, i.e., the perturbations ΔC and ΔD are symmetric. If $\|\Delta A\|$, $\|\Delta C\|$ and $\|\Delta D\|$ are sufficiently small, then the perturbed solution $\tilde{X} = X + \Delta X$ is well defined [15].

Consider first the Riccati equation (3). The *condition number of the Riccati equation* is defined as (see [6])

$$K_R = \lim_{\delta \to 0} \sup \left\{ \frac{\|\Delta X\|}{\delta \|X\|} : \|\Delta A\| \leq \delta \|A\|, \ \|\Delta C\| \leq \delta \|C\|, \ \|\Delta D\| \leq \delta \|D\| \right\}.$$

For sufficiently small δ we have (within first order terms)

$$\frac{\|\Delta X\|}{\|X\|} \leq K\delta.$$

Let $\bar{X}$ be the solution of the Riccati equation computed by a numerical method in finite arithmetic with relative precision ε. If the method is *backward stable*, then we can estimate the error in the solution error

$$\frac{\|\bar{X} - X\|}{\|X\|} \leq p(n)K\varepsilon$$

with some low–order polynomial $p(n)$ of n. This shows the importance of the condition number in the numerical solution of Riccati equation.

Consider the perturbed Riccati equation

$$(A+\Delta A)^T(X+\Delta X) + (X+\Delta X)(A+\Delta A) + C + \Delta C - (X+\Delta X)(D+\Delta D)(X+\Delta X) = 0 \quad (6)$$

and set $A_c = A - DX$. Subtracting (4) from (6) and neglecting the second– and higher–order terms in ΔX (i.e., using a first–order perturbation analysis) we obtain a Lyapunov equation in ΔX:

$$A_c^T \Delta X + \Delta X A_c = -\Delta C - (\Delta A^T X + X\Delta A) + X\Delta D X = 0. \quad (7)$$

Let $\mathrm{Vec}(M)$ denotes the vector, obtained by stacking the columns of the matrix M. Then we have that

$$\|\mathrm{Vec}(M)\|_2 = \|M\|_F$$

and equation (7) can be written in the vectorized form as

$$\begin{aligned}(I_n \otimes A_c^T + A_c^T \otimes I_n)\mathrm{Vec}(\Delta X) &= -\mathrm{Vec}(\Delta C) - \\ &\quad (I_n \otimes X + (X \otimes I_n)W)\mathrm{Vec}(\Delta A) + \\ &\quad (X \otimes X)\mathrm{Vec}(\Delta D)),\end{aligned} \tag{8}$$

where we use the representations

$$\begin{aligned}\mathrm{Vec}(\Delta A^T) &= W\mathrm{Vec}(\Delta A) \\ \mathrm{Vec}(MZN) &= (N^T \otimes M)\mathrm{Vec}(Z)\end{aligned}$$

and W is the so called vec–permutation matrix [10].

Since the matrix A_c is stable, the matrix $I_n \otimes A_c^T + A_c^T \otimes I_n$ is nonsingular and we have that

$$\begin{aligned}\mathrm{Vec}(\Delta X) &= (I_n \otimes A_c^T + A_c^T \otimes I_n)^{-1}(-\mathrm{Vec}(\Delta C) - \\ &\quad (I_n \otimes X + (X \otimes I_n)W)\mathrm{Vec}(\Delta A) + \\ &\quad (X \otimes X)\mathrm{Vec}(\Delta D))\end{aligned} \tag{9}$$

Equation (9) can be written as

$$\mathrm{Vec}(\Delta X) = -[P^{-1}, \ \ Q, \ \ -S] \begin{bmatrix} \mathrm{Vec}(\Delta C) \\ \mathrm{Vec}(\Delta A) \\ \mathrm{Vec}(\Delta D) \end{bmatrix} \tag{10}$$

where

$$\begin{aligned}P &= I_n \otimes A_c^T + A_c^T \otimes I_n, \\ Q &= P^{-1}(I_n \otimes X + (X \otimes I_n)W), \\ S &= P^{-1}(X \otimes X).\end{aligned}$$

If we set

$$\eta = \max\left\{\|\Delta A\|_F/\|A\|_F, \ \|\Delta C\|_F/\|C\|_F, \ \|\Delta D\|_F/\|D\|_F\right\},$$

then it follows from (10) that

$$\|\Delta X\|_F/\|X\|_F \le \sqrt{3}K_F\eta,$$

where

$$K_F = \left\| \left[P^{-1},\ Q,\ S \right] \right\|_2 / \|X\|_F$$

is the condition number of (3) using Frobenius norms. The computation of K_F requires the construction and manipulation of $n^2 \times n^2$ matrices which is not practical for large n. Furthermore, the computation of the condition number of the Riccati equation involves the solution matrix X, so that the condition number can be determined only after solving the equation.

Since the computation of the exact condition number is a difficult task, it is useful to derive approximations of K that can be obtained cheaply.

Rewrite equation (7) as

$$\Delta X = -\Omega^{-1}(\Delta C) - \Theta(\Delta A) + \Pi(\Delta D), \tag{11}$$

where

$$\begin{aligned} \Omega(Z) &= A_c^T Z + Z A_c, \\ \Theta(Z) &= \Omega^{-1}(Z^T X + X Z), \\ \Pi(Z) &= \Omega^{-1}(X Z X) \end{aligned}$$

are linear operators in the space of $n \times n$ matrices, which determine the sensitivity of X with respect to the perturbations in $C,\ A,\ D$, respectively. Based on (11) it was suggested in [6] to use the approximate condition number

$$K_B := \frac{\|\Omega^{-1}\|\|C\| + \|\Theta\|\|A\| + \|\Pi\|\|D\|}{\|X\|}, \tag{12}$$

where

$$\begin{aligned} \|\Omega^{-1}\| &= \max_{Z \neq 0} \tfrac{\|\Omega^{-1}(Z)\|}{\|Z\|} \\ \|\Theta\| &= \max_{Z \neq 0} \tfrac{\|\Theta(Z)\|}{\|Z\|} \\ \|\Pi\| &= \max_{Z \neq 0} \tfrac{\|\Pi(Z)\|}{\|Z\|} \end{aligned}$$

are the corresponding induced operator norms. Note that the quantity

$$\|\Omega^{-1}\|_F = \max_{Z \neq 0} \frac{\|Z\|_F}{\|A_c^T Z + Z A_c\|_F} = \frac{1}{\mathrm{sep}_F(A_c^T, -A_c)}$$

where

$$\mathrm{sep}_F(A_c^T, -A_c) := \min_{Z \neq 0} \frac{\|A_c^T Z + Z A_c\|_F}{\|Z\|_F} = \sigma_{\min}(I_n \otimes A_c^T + A_c^T \otimes I_n)$$

is connected to the sensitivity of the Lyapunov equation

$$A_c^T X + X A_c = -C$$

(see [11]).

Comparing (9) and (11) we obtain that

$$\begin{aligned} \|\Omega^{-1}\|_F &= \|P^{-1}\|_2, \\ \|\Theta\|_F &= \|Q\|_2, \\ \|\Pi\|_F &= \|S\|_2. \end{aligned} \tag{13}$$

Consider now in brief the conditioning of the other matrix equations.

The sensitivity of the Lyapunov equation (1) is determined by the norms of the operators

$$\begin{aligned} \Omega(Z) &= A^T Z + ZA, \\ \Theta(Z) &= \Omega^{-1}(Z^T X + XZ). \end{aligned}$$

In the case of the discrete-time Lyapunov equation (2) the corresponding operators are determined from

$$\begin{aligned} \Omega(Z) &= A^T Z A - Z, \\ \Theta(Z) &= \Omega^{-1}(Z^T X A + A^T X Z). \end{aligned}$$

Finally, it can be shown that the conditioning of the discrete-time Riccati equation (4) is determined by the norms of the operators

$$\begin{aligned} \Omega(Z) &= A_c^T Z A_c - Z, \\ \Theta(Z) &= \Omega^{-1}(Z^T X A_c + A_c^T X Z), \\ \Pi(Z) &= \Omega^{-1}(A_c^T X Z X A_c) \end{aligned}$$

where $A_c = (I_n + DX)^{-1} A$.

3 Condition estimation

The quantities $\|\Omega^{-1}\|_1$, $\|\Theta\|_1$, $\|\Pi\|_1$ arising in the sensitivity analysis of Lyapunov and Riccati equations can be efficiently estimated by using the norm estimator, proposed in [13] which estimates the norm $\|T\|_1$ of a linear operator T, given the ability to compute Tv and $T^T w$ quickly for arbitrary v and w. This estimator is implemented in the LAPACK subroutine `xLACON` [1], which is called via a reverse communication interface, providing the products Tv and $T^T w$.

Consider for definiteness the case of continuous-time Riccati equation. With respect to the computation of

$$\|\Omega^{-1}\|_F = \|P^{-1}\|_2 = \frac{1}{\mathrm{sep}_F(A_c^T, -A_c)}$$

the use of `xLACON` means to solve the linear equations

$$Py = v,\ P^T z = v,$$

where

$$P = I_n \otimes A_c^T + A_c^T \otimes I_n,\ \ P^T = I_n \otimes A_c + A_c \otimes I_n,$$

v being determined by `xLACON`. This is equivalent to the solution of the Lyapunov equations

$$\begin{aligned} A_c^T Y + Y A_c &= V \\ A_c Z + Z A_c^T &= V \\ \mathrm{Vec}(V) = v,\ \mathrm{Vec}(Y) &= y,\ \mathrm{Vec}(Z) = z. \end{aligned} \tag{14}$$

The solution of Lyapunov equations can be obtained in a numerically reliable way using the Bartels–Stewart algorithm [5], which first reduces the matrix A_c to Schur triangular form via orthogonal similarity transformations and then solves recursively the triangular Lyapunov equation. Note that in (14) the matrix V is symmetric, which allows a reduction in complexity by operating on vectors v of length $n(n+1)/2$ instead of n^2.

An estimate of $\|\Theta\|_1$ can be obtained in a similar way by solving the Lyapunov equations

$$\begin{aligned} A_c^T Y + Y A_c &= V^T X + XV \\ A_c Z + Z A_c^T &= V^T X + XV. \end{aligned} \tag{15}$$

To estimate $\|\Pi\|_1$ via `xLACON`, it is necessary to solve the equations

$$\begin{aligned} A_c^T Y + Y A_c &= XVX \\ A_c Z + Z A_c^T &= XVX, \end{aligned} \tag{16}$$

where the matrix V is again symmetric and we can again work with shorter vectors.

The estimation of $\|\Omega\|_1,\ \|\Theta\|_1,\ \|\Pi\|_1$ in the case of the other equations is done in a similar way.

The accuracy of the estimates that we obtain via this approach depends on the ability of `xLACON` to find a right–hand side vector v which maximises the ratios

$$\frac{\|y\|}{\|v\|},\ \frac{\|z\|}{\|v\|}.$$

when solving the equations

$$Py = v,\ P^T z = v.$$

As in the case of other condition estimators it is always possible to find special examples when the value produced by `xLACON` underestimates the true value of the corresponding norm by an arbitrary factor. Note, however, that this may happen only in rare circumstances.

To avoid overflows, instead of estimating the condition number K_B of the continuous-time Riccati equation an estimate of the reciprocal condition number

$$\frac{1}{\tilde{K}_B} = \frac{\widetilde{\mathrm{sep}}_1(\bar{A}_c^T, -\bar{A}_c)\|\bar{X}\|_1}{\|C\|_1 + \widetilde{\mathrm{sep}}_1(\bar{A}_c^T, -\bar{A}_c)(\|\tilde{\Theta}\|_1\|A\|_1 + \|\tilde{\Pi}\|_1\|D\|_1)}$$

is determined. Here $\bar{A}_c$ is the computed matrix A_c and the estimated quantities are denoted by $\tilde{}$. The same approach can be used in the condition estimation of the other equations.

4 Error estimation

A posteriori error bounds for the computed solution of the matrix equations (1) - (4) may be obtained in several ways, see for instance [9]. One of the most efficient and reliable ways to get an estimate of the solution error is to use practical error bounds, similar to the case of solving linear systems of equations [2, 1] and matrix Sylvester equations [12].

Consider again the Riccati equation (3).

Let

$$R = A^T\bar{X} + \bar{X}A + C - \bar{X}D\bar{X}$$

be the exact residual matrix associated with the computed solution $\bar{X}$. Setting $\bar{X} := X + \Delta X$, where X is the exact solution and ΔX is the absolute error in the solution, one obtains

$$R = (A - D\bar{X})^T\Delta X + \Delta X(A - D\bar{X}) + \Delta X D \Delta X.$$

If we neglect the second order term in ΔX, we obtain the linear system of equations

$$\bar{P}\mathrm{Vec}(\Delta X) = \mathrm{Vec}(R),$$

where $\bar{P} = I_n \otimes \bar{A}_c^T + \bar{A}_c^T \otimes I_n,\ \bar{A}_c = A - D\bar{X}$. In this way we have

$$\|\mathrm{Vec}(X - \bar{X})\|_\infty = \|\bar{P}^{-1}\mathrm{Vec}(R)\|_\infty \le \|\ |\bar{P}^{-1}|\ |\mathrm{Vec}(R)|\ \|_\infty.$$

As it is known [2] this bound is optimal if we ignore the signs in the elements of $\bar{P}^{-1}$ and $\mathrm{Vec}(R)$.

In order to take into account the rounding errors in forming the residual matrix, instead of R we use

$$\bar{R} = fl(C - A^T\bar{X} - \bar{X}A - \bar{X}D\bar{X}) = R + \Delta R,$$

where

$$|\Delta R| \leq \varepsilon(4|C| + (n+4)(|A^T|\,|\bar{X}| + |\bar{X}|\,|A|) + 2(n+1)|\bar{X}|\,|D|\,|\bar{X}|) =: R_\varepsilon.$$

Here we made use of the well known error bounds for matrix addition and matrix multiplication [14].

In this way we have obtained the overall bound

$$\frac{\|X - \bar{X}\|_M}{\|\bar{X}\|_M} \leq \frac{\|\,|P^{-1}|\,(|\mathrm{Vec}(\bar{R})| + \mathrm{Vec}(R_\varepsilon))\|_\infty}{\|\bar{X})\|_M}, \tag{17}$$

where $\|X\|_M = \max_{i,j}|x_{ij}|$.

The numerator in the right hand side of (17) is of the form $\|\,|P^{-1}|r\|_\infty$, and as in [2, 12] we have

$$\begin{array}{rcccc} \|\,|\bar{P}^{-1}|r\,\|_\infty & = & \|\,|\bar{P}^{-1}|D_R e\|_\infty & = & \|\,|\bar{P}^{-1}D_R|e\|_\infty \\ & = & \|\,|\bar{P}^{-1}D_R|\,\|_\infty & = & \|\bar{P}^{-1}D_R\|_\infty \end{array}$$

where $D_R = \mathrm{diag}(r)$ and $e = (1,1,...,1)^T$. This shows that $\|\,|P^{-1}|r\|_\infty$ can be efficiently estimated using the norm estimator `xLACON` in LAPACK, which estimates $\|Z\|_1$ at the cost of computing a few matrix-vector products involving Z and Z^T. This means that for $Z = \bar{P}^{-1}D_R$ we have to solve a few linear systems involving $\bar{P} = I_n \otimes \bar{A}_c^T + \bar{A}_c^T \otimes I_n$ and $\bar{P}^T = I_n \otimes \bar{A}_c + \bar{A}_c \otimes I_n$ or, in other words, we have to solve several Lyapunov equations $\bar{A}_c^T X + X\bar{A}_c = V$ and $\bar{A}_c X + \bar{A}_c^T = W$. Note that the Schur form of $\bar{A}_c$ is already available from the condition estimation of the Riccati equation, so that the solution of the Lyapunov equations can be obtained efficiently via the Bartels-Stewart algorithm. Also, due to the symmetry of the matrices $\bar{R}$ and R_ε, we only need the upper (or lower) part of the solution of this Lyapunov equations which allows to reduce the complexity by manipulating only vectors of length $n(n+1)/2$ instead of n^2.

The error estimation in the solution of (1), (2) and (4) is obtained in a similar way.

The software implementation of the condition and error estimates for the equations (1)-(4) is based entirely on LAPACK and BLAS [16, 7] subroutines. It consists of eight routines written in Fortran 77 which may be used along with different methods for solving the corresponding equations.

5 Numerical examples

In this section we present four examples which demonstrate the performance of the estimates implemented in the solution of families of Lyapunov and Riccati equations whose conditioning vary very much. All computations were carried out on a PC with relative machine precision $\varepsilon = 2.22 \times 10^{-16}$.

In order to have a closed form solution, the test matrices in the Lyapunov and Riccati equations are chosen as

$$A = TA_0T^{-1}, \;\; C = T^{-T}C_0T^{-1}, \;\; D = TD_0T^T,$$

where $A_0,\ C_0,\ D_0$ are diagonal matrices and T is a nonsingular transformation matrix. The solution is then given by $X = T^{-T}X_0T^{-1}$ where X_0 is a diagonal matrix whose elements are determined simply from the elements of $A_0,\ C_0,\ D_0$. To avoid large rounding errors in constructing and inverting T, this matrix is chosen as $T = H_2SH_1$ where H_1 and H_2 are elementary reflectors and S is a diagonal matrix,

$$\begin{aligned} H_1 &= I_n - 2ee^T/n, \;\; e = [1, 1, ..., 1]^T, \\ H_2 &= I_n - 2ff^T/n, \;\; f = [1, -1, 1, ..., (-1)^{n-1}]^T, \\ S &= \mathrm{diag}(1, s, s^2, ..., s^{n-1}), \;\; s > 1. \end{aligned}$$

Using different values of the scalar s, it is possible to change the condition number of the matrix T with respect to inversion, $\mathrm{cond}_2(T) = s^{n-1}$.

The matrices A, C, D are computed easily with high precision. The numerical solution of the corresponding equations, however, may present a difficult task for the methods which are of current use, since the diagonal structure of these equations is not recognized by these methods.

Example 1 Consider the solution of a family of Lyapunov equations of sixth order, constructed such that

$$\begin{aligned} A_1 &= \mathrm{diag}(-1 \times 10^{-k}, -2, -3 \times 10^k), \\ C_1 &= \mathrm{diag}(2 \times 10^k, 4, 6 \times 10^{-k}). \end{aligned}$$

The solution X_0 is given by

$$X_1 = \mathrm{diag}(10^{2k}, 1, 10^{-2k}).$$

The condition number of thes equations increases quickly with the increasing of k and s.

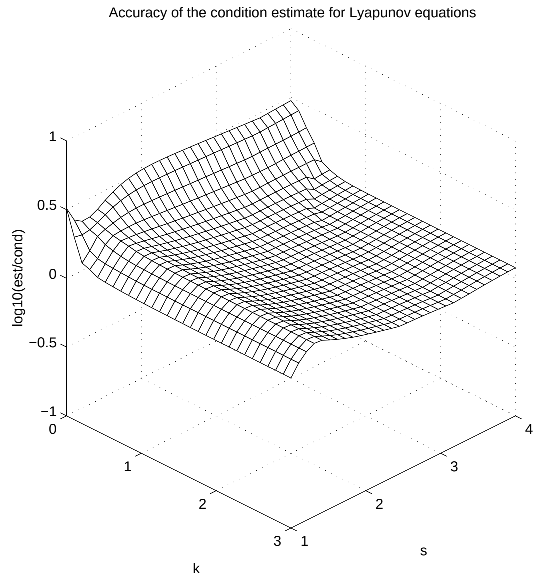

Figure 1: Accuracy of the condition number estimate for a family of Lyapunov equations

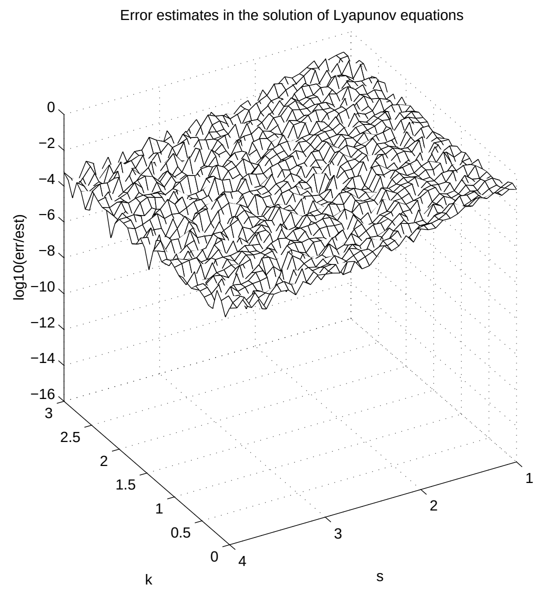

Figure 2: Accuracy of the forward error estimate for a family of Lyapunov equations

In Figure 1 we show the ratio of the condition number estimate and the value of the actual condition number as a function of k and s. This ratio remains close to 1 even for large k and

s when the condition number is of order 10^{11}. In Figure 2 we show the ratio of the actual forward error in the solution and the estimate of this error for the same values of k and s. In all cases the actual error is less than the residual based estimate. With the increasing of the condition number the forward error estimate becomes pessimistic, but the difference between both quantities is at most in the last four decimal digits.

Example 2 Consider a family of discrete-time Lyapunov equations of sixth order obtained for matrices A_0, C_0 whose diagonal blocks are chosen as

$$\begin{aligned} A_1 &= \mathrm{diag}(1 - 10^{-k}, 0, 1/2), \\ C_1 &= \mathrm{diag}(10^{-k}, 10^{k}, 10^{-k}). \end{aligned}$$

As in the continuous-time case these equations become ill-conditioned with the increase of k ans s.

The solution of the discrete-time Lyapunov equation is done by the algorithm proposed in [4].

The results related to the condition number estimate, shown in Figure 3, demonstrate the good performance of this estimate for different k and s (the ratio of the estimate and the true condition number remains close to 1).

The results related to the forward error estimate, presented in Figure 4, show that for the given discrete-time Lyapunov equations the error estimate may be pessimistic. In any case, however, we are sure that the actual forward error in the solution is less than the estimate obtained.

Example 3 Consider a family of Riccati equations, constructed as described above with

$$A_0 = \mathrm{diag}(A_1, A_1), \quad C_0 = \mathrm{diag}(C_1, C_1), \quad D_0 = \mathrm{diag}(D_1, D_1),$$

where

$$\begin{aligned} A_1 &= \mathrm{diag}(-1 \times 10^{-k}, -2, -3 \times 10^{k}), \\ C_1 &= \mathrm{diag}(3 \times 10^{-k}, 5, 7 \times 10^{k}), \\ D_1 &= \mathrm{diag}(10^{-k}, 1, 10^{k}), \\ X_1 &= \mathrm{diag}(1, 1, 1). \end{aligned}$$

The solution is given by

$$X_0 = \mathrm{diag}(X_1, X_1), \quad X_1 = \mathrm{diag}(1, 1, 1).$$

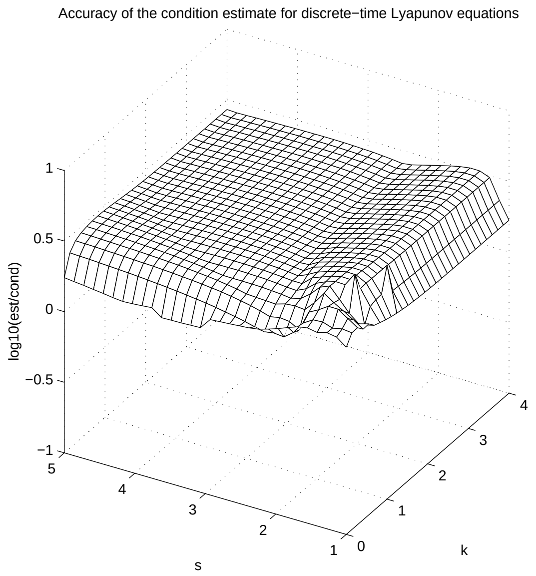

Figure 3: Accuracy of the condition number estimate for a family of discrete-time Lyapunov equations

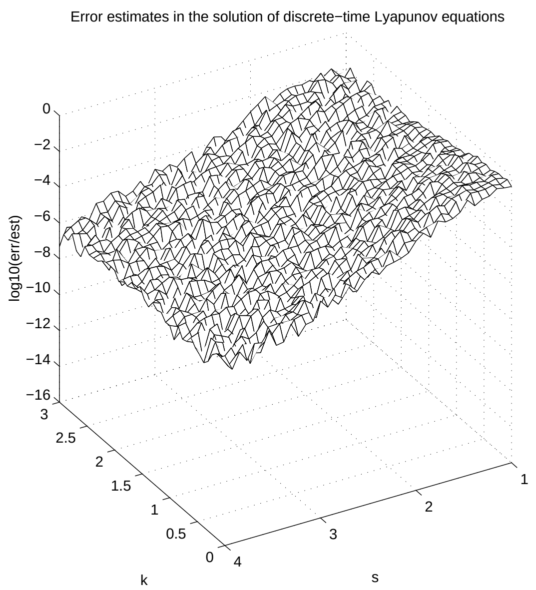

Figure 4: Accuracy of the forward error estimate for a family of discrete-time Lyapunov equations

The conditioning of these equations deteriorates with the increase of k and s.

These equations are solved with the routines described in [18].

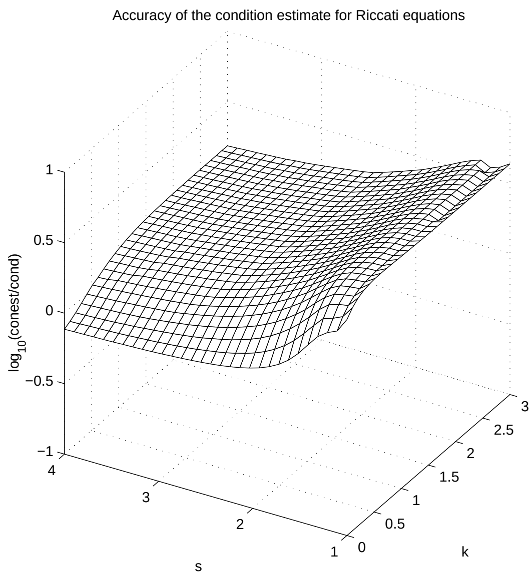

Figure 5: Accuracy of the condition number estimate for a family of Riccati equations

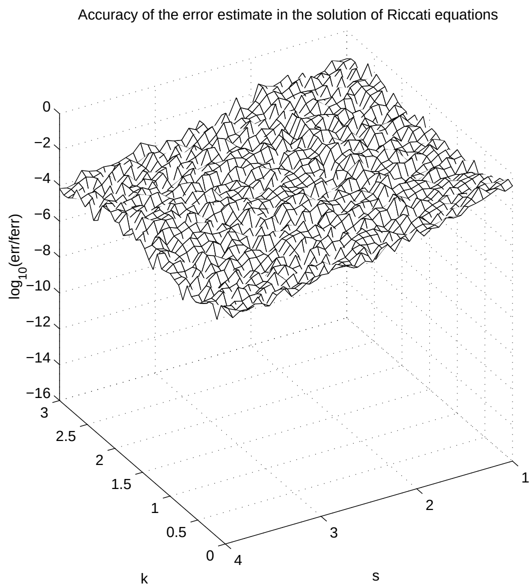

Figure 6: Accuracy of the forward error estimate for a family of Riccati equations

In Figure 5 we show the ratio of the condition number estimate and the exact condition number and in Figure 6 we show the ratio of the exact forward error in the solution and its estimate as functions of k and s. Both estimates produce acceptable results in this case.

The next example illustrates the potential pessimism in the forward error estimate for the disceter-time Riccati equation.

Example 4 Consider a family of 6-th order discrete-time Riccati equations whose matrices A_0, C_0, D_0 are chosen as

$$A_0 = \mathrm{diag}(A_1, A_1), \;\; C_0 = \mathrm{diag}(C_1, C_1), \;\; D_0 = \mathrm{diag}(D_1, D_1),$$

where

$$\begin{aligned} A_1 &= \mathrm{diag}(0, 1, 2), \\ C_1 &= \mathrm{diag}(10^k, 1, 10^{-k}), \\ D_1 &= \mathrm{diag}(10^{-k}, 10^{-2*k}, 10^{-k}). \end{aligned}$$

The conditioning of these equations deteriorates with the increase of k and s.

The accuracy of the condition number estimate for the discrete-time Riccati equations is shown in Fig. 7. As for the continuous-time Riccati equations, the condition number estimate is close to the true condition number. However, as shown in Fig.8, the forward error estimate becomes very pessimistic for large k and s.

6 Conclusions

The results obtained in the paper show that it is possible to use successfuly the LAPACK matrix norm estimator in the condition and forward estimation for the Lyapunov and Riccati matrix equations arising in control. The numerical experiments show that the condition estimates are allways of the same order as the true condition numbers. However, the forward error estimates may be pessimistic just as in the solution of linear systems of equations. It should be pointed out that, theoretically, the forward error estimates may underestimate the true errors in the solution of the Riccati equations due to the neglecting of the higher order terms in the analysis. Such phenomenon was never observed in practice which shows that the forward error estimates are sufficiently reliable.

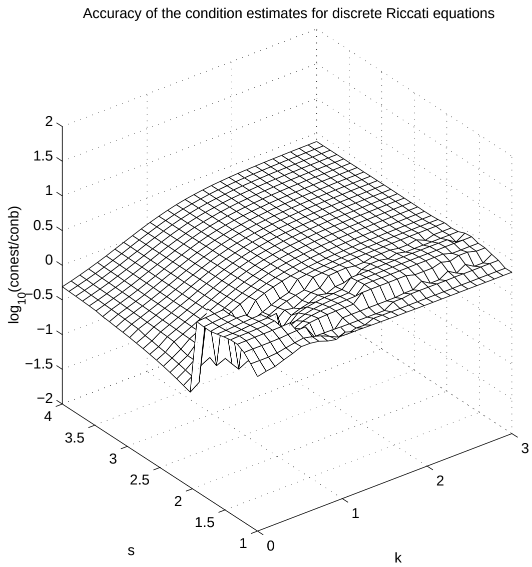

Figure 7: Accuracy of the condition number estimate for a family of discrete-time Riccati equations

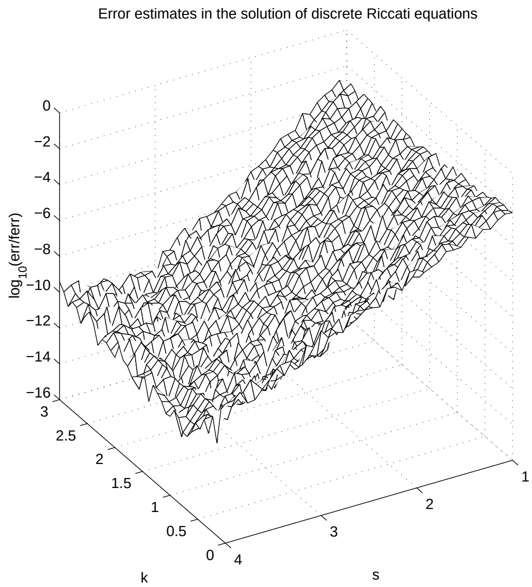

Figure 8: Accuracy of the forward error estimate for a family of discrete-time Riccati equations

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, PA, second edition, 1995.

[2] M. Arioli, J.W. Demmel, and I.S. Duff. Solving sparse linear systems with sparse backward error. *SIAM J. Matrix Anal. Appl.*, 10:165–190, 1989.

[3] Z. Bai, J. Demmel, and A. Mckenney. On computing condition numbers for the nonsymmetric eigenproblem. *ACM Trans. Math. Software*, 19:202–223, 1993.

[4] A.Y. Barraud. A numerical algorithm to solve $A^TXA - X = Q$. *IEEE Trans. Automat. Control*, AC-22:883–885, 1977.

[5] R.H. Bartels and G.W. Stewart. Algorithm 432: Solution of the matrix equation $AX + XB = C$. *Comm. ACM*, 15:820–826, 1972.

[6] R. Byers. Numerical condition of the algebraic Riccati equation. *Contemp. Math.*, 47:35–49, 1985.

[7] J.J. Dongarra, J. Du Croz, I. Duff, and S. Hammarling. A set of Level 3 Basic Linear Algebra Subprograms. *ACM Trans. Math. Software*, 16:1–17, 1990.

[8] J.J. Dongarra, J. Du Croz, S. Hammarling, and R.J Hanson. An extended set of FORTRAN Basic Linear Algebra Subprograms. *ACM Trans. Math. Software*, 14:1–17, 1988.

[9] A.R. Ghavimi and A.J. Laub. Backward error, sensitivity and refinement of computed solutions of algebraic Riccati equations. *Numer. Alg. Appl.*, 2:29–49, 1995.

[10] H.V. Henderson and S.R. Searle. The vec-permutation matrix, the vec operator and Kronecker products: A review. *Lin. Multilin. Alg.*, 9:271–288, 1981.

[11] G. Hewer and C. Kenney. The sensitivity of the stable Lyapunov equation. *SIAM J. Cont. Optim.*, 26:321–344, 1988.

[12] N. J. Higham. Perturbation theory and backward error for $AX - XB = C$. *BIT*, 33:124–136, 1993.

[13] N.J. Higham. FORTRAN codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation (Algorithm 674). *ACM Trans. Math. Software*, 14:381–396, 1988.

[14] N.J. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia, PA, 1996.

[15] C. Kenney and G. Hewer. The sensitivity of the algebraic and differential Riccati equations. *SIAM J. Cont. Optim.*, 28:50–69, 1990.

[16] C.L. Lawson, R.J. Hanson, D.R. Kincaid, and F.T. Krogh. Basic Linear Algebra Subprograms for FORTRAN usage. *ACM Trans. Math. Software*, 5:308–323, 1979.

[17] P.H. Petkov, N.D. Christov, and M.M. Konstantinov. *Computational Methods for Linear Control Systems.* Prentice-Hall, Hemel Hempstead, Herts, UK, 1991.

[18] P.H. Petkov, M.M. Konstantinov, and V. Mehrmann. DGRSVX and DMSRIC: Fortan 77 subroutines for solving continuous-time matrix algebraic Riccati equations with condition and accuracy estimates. Technical Report SFB393/98-16, Fak. für Mathematik, TU Chemnitz, Chemnitz, FRG, May 1998.